Leveraging the value of Crowdsourced Geographic Information to detect Cultural Ecosystem Services

Gianfranco Antonino Gliozzo

University College London

Thesis submitted for the Degree of Doctor of Engineering (EngD)

January 2018

Author's declaration

I, Gianfranco Antonino Gliozzo, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Within ecological research and environmental management, there is current focus on demonstrating the links existing between human well-being and nature conservation. There is a need for better understanding how and why people value certain places over others. At the same time, there is a lack of consolidated methodologies, and limited experimentation in the detection of places connected to the immaterial benefits we get from nature. Those benefits are termed Cultural Ecosystem Services (CES). This research analyses the potential of Crowdsourced Geographic Information (CGI) to support the detection of CES with large scale insights derived from the analysis of digital cultural practices. CGI is produced through social media, in situations where individuals choose to share content. Therefore, a CGI project is often the expression of a community of interest and different projects have different supporting communities with different demographics and cultural profiles. The research combines multiple projects pertaining to three different categories of CGI to avoid focusing only on a community or on a digital cultural practice. Using ecological and social considerations, this thesis contributes to the evaluation of such projects as potential analytical tools for CES research. The degree of appreciation of a specific place is derived from the number of people creating, sharing, or refining the information about it. The sequence of decisions and actions that leads to the sharing of information leaves digital proxies of spatial preferences, with people sharing specific information considering the place not only “worth visiting” but also “worth sharing”. Using south Wales and London as case studies, we demonstrate how the analysis of CGI can be included in methodologies used to detect CES. These results highlight how the inclusion of CGI, can be very effective in addressing some of the current priorities in conservation. It could potentially be used for better prioritisation, planning and management of natural and cultural resources towards a more sustainable development.

Impact statement

This work is primarily addressed at the enhancement and optimisation of public policy using a combination of technologies and disciplines.

This work advances the comprehension not only of the way people and nature are related, but also of the ways in which new kinds of data and technologies must be implemented to elicit knowledge. We here highlight the limitations and potential developments of overall conceptualisations created to explain how humans and nature are inextricably connected. This work advocates the centrality of culture who plays a pivotal role in facilitating the regeneration of human-nature relationships.

For academia we analyse and develop in a rigorous and coherent way the use of geographic information produced by people to elicit their preferences regarding places. This leads to the development of a methodology that uses and advances knowledge, methods, and tools taken from computer science and geography to serve ecology and public policy design.

The methodology proposed can obtain in weeks very effective evaluations that are usually obtained after several months if not more than a year.

Outside from academia the impact on public policy design can be more relevant in places where more data created by people is recorded in either digital or analogue format. This is also coherent with studies that demonstrated how less economically developed communities are more reliant on the benefits of the local natural environment.

Academia might use this work as a contribute to enhance methods but also overall conceptualisations. Public policy might adopt and develop the methods who are able to deliver inexpensive and fast assessments. Both academia and public policy can gain advantage of large amounts of data without falling in the oversimplification of reality often advocated by technology focused approaches.

Acknowledgments

This work is the result of years of dedication and hard work who have been inspired supported and motivated by several people.

Firstly, I want to thank my parents and my sister who taught me the values of hard working, perseverance, and equipped me with an inquisitive mind.

A driving force in inspiring an academic degree has been given to me by my early education, experiences, and challenges in the University of Catania under Professor Paolo La Greca.

The experiences in the Netherlands gave me the possibilities to widen my horizons thanks to the kind hosting of Professor Gerrit-Jan Carsjens in Wageningen and the exceptionally talented, resourceful, and dedicated lecturers of the MSc GIMA.

This led me to London where Sebastiano played a fundamental role helping me to get a grip on the place.

Professor Muki Haklay invited, nurtured, sustained endlessly, inspired, drove, and borne the burden of my personality with endless patience and dedication. Without any of his contributions this work was probably not done. I thank him also for the people he has chosen for his research group, they are amongst my dearest friends.

I must thank ZSL and Nathalie Pettorelli that acted as industrial sponsor of this work.

I am grateful also to the communities and companies behind the projects I used for this work. I am especially grateful to GiGL's people who kindly sat with me several times to discuss about their precious data holdings.

My deepest and biggest obligation goes to my wife Maria Grazia. She supported, inspired, tolerated, cared for me in so many ways all along the final years of this work. Her love and her presence help me to strive to be a better human. She cast a new light, a new life, a new purpose, and happiness in my being.

List of abbreviations and acronyms

AONB	Area of Outstanding Natural Beauty
API	Application Programming Interface
BBOX	Bounding box
BNG	British National Grid
CES	Cultural Ecosystem Services
CGI	Crowdsourced Geographic Information
CICES	Common International Classification Of Ecosystem Services
CM	Collaborative Mapping
Cr	Contributor ratio
CSV	Comma Separated Values
ES	Ecosystem Services
FEGS	Final Ecosystem Goods and Services
GBIF	Global Biodiversity Information Facility
GiGL	Greenspace information for Greater London CIC
GIS	Geographic Information Systems
GPS	Global Positioning System
HERCULES	HERitage in CULtural landscapES
ICT	Information and Communication Technologies
InVEST	Integrated Valuation of Ecosystem Services and Tradeoffs
IPBES	Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services
IT	Information Technology
JSON	JavaScript Object Notation
LISA	Local Indicator of Spatial Association
LNHS	London Natural History Society

LSOA	Lower Layer Super Output Areas
MAES	Mapping and Assessment of Ecosystems and their Services
MEA	Millennium Ecosystem Assessment
MAUP	Modifiable Area Unit Problem
MENE	Monitor of Engagement with the Natural Environment
NBN	National Biodiversity Network
NC	Natural Capital
NNR	National Nature Reserves
OA	Output Area
OECD	Organisation for Economic Cooperation and Development
OPENness	Operationalisation of Natural Capital and Ecosystem Services
OS	Ordnance Survey
OSM	OpenStreetMap
PES	Payment for Ecosystem Services
POI	Point Of Interest
PPGIS	Public Participation GIS
SSSI	Site of Special Scientific Interest
TEEB	The Economics of Ecosystems and Biodiversity
TSV	Tab Separated Values
UGC	User Generated Content
UK NEA	United Kingdom National Ecosystem Assessment
UK NEA FO	United Kingdom National Ecosystem Assessment, Follow-On phase
VGI	Volunteered Geographic Information
Vp	Volunteer productivity
XML	eXtensible Markup Language

Table of Contents

Table of Figures

Table of Tables

1 Introduction[1]

1.1 Nature and society

The engagement of people with nature is inherently bound to human nature. Quality of life and a sense of fulfilment have been linked to contact with nature (Max-Neef 1991, Seymour 2016). Civilisation as defined in the Oxford dictionary[2] comes from the Latin term for city (civitas), and human population is increasingly concentrated in urban environments, rising from 43% to 52% at a global scale between 1990 and 2012[3]. This trend is just one of the factors that contributed to the creation of a dichotomy human-nature that relies only on human social/economical constructed perceptions (Demeritt 2002). The increasingly urbanised population and an economic system focused on consumerism has contributed in the creation of a cultural disengagement of humans from nature. A consequence of this disengagement is that open spaces and natural environment are seen as *"open land"* for the planning of infrastructures, the disposal of waste, and marginally the place of some enjoyment of the natural environment. The value of nature *per se* has been supported by natural scientists and several conceptualisations and policies at global and national level have been proposed. But the disengagement from nature and the estimation of *"open land"* as a resource for economic development often undermined such initiatives.

The relation among human activities, the environment, and development has a long tradition of concerns, studies, and policies at every level. The fact that the environment has been seen as a resource to exploit for economic development is also witnessed by the report *The Limits to Growth* (Meadows et al. 1972). Landers (2016) citing Daily et al. (1997) explains how, following a more individualistic and productive perspective on human life there has been the development of an "environmental blind spot". Another important report sketched the concept of a different relation between environment and development. The so-called Brundtland report (WCED 1987) proposed the concept of sustainable development as the development that meets social, economic, and ecological targets without being detrimental to any of them. Subsequently Agenda 21 (UN 1993), an agenda for the twenty-first century, was designed as an implementation plan for sustainable development linking social compatibility, economics growth, and ecological protection. In Agenda 21 there is a strong emphasis on the local scale of

[1] Several sections of the thesis are adaptations from "Gliozzo, G., N. Pettorelli, and M. Haklay. 2016. Using crowdsourced imagery to detect cultural ecosystem services: a case study in South Wales, UK. Ecology and Society 21." And "Boakes, E. H., G. Gliozzo, V. Seymour, M. Harvey, C. Smith, D. B. Roy, and M. Haklay. 2016. Patterns of contribution to citizen science biodiversity projects increase understanding of volunteers' recording behaviour. Scientific Reports 6:33051."

[2] The stage of human social development and organisation which is considered most advanced

[3] http://wdi.worldbank.org/table/3.12

intervention to achieve results at a global scale. Together, those studies and policies are converging towards a stronger emphasis on the need to involve people to seek knowledge and spur action in a coherent attempt to promote bottom-up approaches (Schmidt et al. 2006). In a later stage, cultural development and cultural assets have been included in the agenda for sustainable development (United Cities and Local Governments - Committee on culture 2008). Together with environmental responsibility, social equity and economic viability, cultural vitality has been named the fourth pillar of sustainable development (Hawkes 2001). While it mostly focuses on cultural production, it enhances the role of local situated knowledge as a fundamental element of identity and regards it as a value that must not be treated as a commodity or a consumer good (United Cities and Local Governments - Committee on culture 2008). Several initiatives to combine social, economic, and ecological components in development fell short of expectation and therefore a novel approach was needed to combine all those concepts. Relevant trade-offs between economic development and ecosystems needs better accounting systems. The Ecosystem Service approach was then developed (Potschin et al. 2016a). The Sustainable Development Goals (SDG) (United Nations 2015) created a consensus around goals and targets to achieve sustainable development. The assessment methods to support policy making needs to be ready for the challenge of the Millennium.

1.2 The Ecosystem Services approach

The Millennium Ecosystem Assessment (MEA) focused on an anthropocentric valuation of nature (Millennium Ecosystem Assessment 2005), with a view to develop a framework for accounting for nature and the benefits it provides to humans. The MEA uses widely the concept of Ecosystem Services (ES) to describe all aspects of the natural world that contribute directly to the enhancement of human well-being. The ES approach aims at sustainable development rather than focusing only on the intrinsic value of nature. The physical constituents of the natural world and their processes constitute the Natural Capital (NC) from which ecosystem services are either extracted, derived, or yield spontaneously (Costanza and Daly 1992). The ES approach stresses the role of flows of benefits (Hinterberger et al. 1997). The genesis of the ES concept is not a scientific outcome but has been mainly coined as a pedagogical concept designed to raise public interest for biodiversity conservation (Peterson et al. 2010). The idea that ecosystems provide benefits to society is deep rooted and Potschin et al. (2016a) traces it back up to the 19th century. Its communicative performance has been boosted by its extensive use in the Millennium Ecosystem Assessment. According to the Scopus abstract and citation database there is a continuous growth trend in citations since 1993 while from 2005, the year of the MEA publication, the yearly growth trend is in double figures. The definition of ES includes the health and well-being benefits that humans derive from nature. These benefits include:

> "provisioning services such as food, water, timber, and fiber; regulating services that affect climate, floods, disease, wastes, and water quality; cultural services that provide recreational, aesthetic, and spiritual benefits; and supporting services such as soil formation, photosynthesis, and nutrient cycling" *(*Millennium Ecosystem Assessment 2005 p. v).

The ES approach has eased interactions between natural scientists, economists, planners, and decision makers using an anthropocentric economic-utilitarian perspective (Carpenter et al. 2009). The broader paradigm embraces concepts that fall outside the disciplinary boundaries of natural and environmental sciences, where the focus is on biophysical elements. While environmental social and economic sciences are the first and explicit target of the conceptualisation, their integration is still lacking especially on the social sciences side (Abson et al. 2014). The most recent conceptualisation to frame the relationship between humans and their environment is still heavily based on ES but it is now stressing a parity between humans and nature, people and nature (Mace 2016, Potschin et al. 2016a) emphasising the oneness of socio-ecological or social-ecological systems[4]. Within the research into ES, there is no closed set of tools or approaches, and the field is highly dynamic and open to debate and innovation (Fish et al. 2016c). The ES approach implies a source, or a provision point of the benefit, a flow, and a receptor or beneficiary area where the benefits are realised. A more detailed analysis such as in The Economics of Ecosystem and Biodiversity (TEEB) report (de Groot et al. 2010) stresses how the complexity of the natural world implies a quite more complex chain of elements involved from ecosystems to beneficiary. In the TEEB report, therefore, the definition is more precisely focusing also on indirect benefits since ES are defined as *"the direct and indirect contributions of ecosystems to human well-being."* (de Groot et al. 2010 p. 19). The shape and strengths of the spatial relationships among so many elements can differ widely (Haines-Young and Potschin 2009a, Fisher et al. 2009, Potschin and Haines-Young 2016a, Costanza et al. 2017).

1.2.1Cultural Ecosystem Services

The integration of the social aspect in the ecosystems is especially challenging when we deal with cultural ecosystem services (CES). The most anthropocentric ES are cultural ecosystem services, which relate to human non-material engagement with nature as they provide recreational, aesthetic, and spiritual benefits. Cultural ecosystem services are focusing on cultural preferences. The relationship between CES and biophysical elements is weak; therefore, preference-based valuation approaches are favoured over

[4] In (Potschin et al. 2016b) we have the term *'socio-ecological'* used 29 times, *'socioecological' 2* times, and *'social-ecological'* 138 times

those based on biophysical elements (Pascual et al. 2010). Non-consumptive use of natural resources also characterises CES (Millennium Ecosystem Assessment 2003). This implies a unique way of representing the spatial relationship between source flow and receptor. More specifically, the flow can be non-material in terms of information (Bagstad et al. 2013a), or "inversely" material in the sense that users move towards places they perceive as worth visiting (Costanza 2008): the beneficiary moves towards the services and not the other way around. ES that are characterised by biophysical elements such as amount of water, nutrients in soil, temperature, and moisture are measurable in a straightforward way. CES are more difficult to quantify because they are a collection of human cultural perceptions that provide benefits that have no physical component. Nonetheless, material elements when related to cultural practices trigger such perceptions. Estimation methodologies rely mainly on proxies. Fluxes, numbers of visits, and expenditure by tourists are the quantifiable elements used to measure the effects of tourism (Ghermandi et al. 2009, Bateman et al. 2013a). While people's perceptions are achieved through participation in nationwide surveys or through participatory mapping approaches. In current practice, the investigation of CES has implied the integration of landscape ecology, landscape aesthetics, social sciences, and environmental or natural sciences (Schaich et al. 2010, Bagstad et al. 2013a, Hernández-Morcillo et al. 2013, Plieninger et al. 2013). But this pool of disciplines can be widened at least following the disciplines mentioned by Chan and Satterfield (2016 p. 344) including "psychology, education, anthropology, sociology, planning, environmental health, medicine, diverse humanities, etc."

1.2.2 Cultural Ecosystem Services geographical extent

When surveys have provided the spatial extent and location of places, the precise identification of places has followed the interpretation of descriptions. This has hindered the ability to locate these places accurately. Participatory mapping techniques can provide more precise indication on location but this methodology is applied to limited spatial units. Some studies have investigated CES along with other ES, analysing trade-offs. Those results remain heavily dependent on long-lasting traditions and on the cooperation of heterogeneous and complex research teams, as in Bateman et al. (2013a). Such approaches strongly depend on the methodologies and data needed for both natural and social sciences. While there is a strong focus on outreach to a wider audience, this happens rarely because of the use of technical or disciplinary vocabularies and approaches. Focusing on the UK, Bateman et al. (2013a) used a survey, the Monitor of Engagement with the Natural Environment (MENE) (Herbert and Brobyn 2011). MENE is the result of a very wide survey involving 45,000 interviews at a yearly pace.

Preferences and spatial locations have been in this case analysed through interviews. The delimitation of a spatial entity from textual descriptions can be inaccurate.

Novel approaches and methodologies are needed to overcome some or all of these limitations, keeping in mind that CES are not directly bounded to geophysical elements but they are embedded in human perceptions. While ES can be considered themselves *boundary objects* (Star and Griesemer 1989, Abson et al. 2014) their investigation therefore requires a wider perspective that embraces different disciplines (Barnaud and Antona 2014, Chan and Satterfield 2016).

1.2.3 On diversity in the detection of Cultural Ecosystem Services

Since “Cultural services comprise a novel category of services that captures many of the non-use or passive use values of ecological resources” and “most cultural services are still regulated by custom and usage, or by traditional taboos, rights and obligations. Nevertheless, they are directly used by people, and so are amenable to valuation by methods designed to reveal people’s preferences” (Perrings 2006 pp. 10–11).

To grasp people's or populations’ preferences, we need to combine both qualitative and quantitative research and to investigate with methodologies that are pertaining to domains not traditionally bound to natural sciences. Benefits from cultural services are non-material and their enjoyment is more characterised by access than by ownership: they are non-use goods. Cultural goods are also mentioned in the UK’s national ecosystem assessment (2011a) in particular in the chapter devoted to cultural services the need to cover knowledge gaps is stressed and innovative approaches are suggested “in future, their collective and non-monetary value will need to be understood using a range of participatory and deliberative techniques, such as multi-criteria analysis, that require both quantitative and qualitative methods.” (Church et al. 2011, Fish et al. 2011). To better understand ecosystem services, we need to use valuation methods that are both inclusive and participative. In the attempt to create an economic evaluation of key cultural benefits Mourato et al. (2010) identified as important elements that increase the knowledge of nature, the improvement of natural capital and consequently stress the importance of outdoor learning with a case study mentioning explicitly environmental citizen science. The overall concept of citizen science embraces a wide range of scientific endeavours that include scientists and citizens in the generation, execution, and management of research enquiries (ECSA 2015). The environmental element of citizen science is mostly focused on data collection of environmental observations. It is also called biological recording mostly in the UK (Powney and Isaac 2015). This implies a tight relation with the surrounding environment. Regarding outdoor learning, "benefits are consumed in the form say of the enjoyment of current amenities" and “future opportunities are enhanced and the output can be construed"..."as an investment in human capital" (Mourato et al. 2010). Similar conclusions are also underlined by Hobbs and White (2012). Learning and enjoying while in a natural or cultural environment are strictly related (Fischer and Eastwood 2016) also called cultural practices (Fish et al. 2016a, 2016b). A wider perspective on the need to open up to a better understanding of social-ecological systems is needed. The need to understand spatial-social practices suggests embracing new approaches with a stronger dialogue with other disciplines and new technologies (Abson et al. 2014, Barnaud and Antona 2014, Kenter 2016a). A static analysis of the spatial extent of social-ecological systems based on biophysical as well as census elements is deemed to fail because census spatial units are not designed to understand the relations between society and nature and due to the new ways in which spatial practices has been evolving. Two main factors are shaping the new social and spatial morphologies. Individual mobility in combination with the advancements in Information Technologies (IT). Mobility allows people to get to environmental spaces while social technologies convey messages of presence absence and preferences that are the target of studies aiming at CES.

1.3 Shrinking space clustering individuals: places matter more than ever

In recent decades, communication and transportation technologies are reshaping social and human geographies. Martinotti (1993, 1996) described new kinds of populations that live and *use* metropolitan urban space. The ease of commuting and the increased Information and Communication Technologies (ICT) are providing new tools to manage our relationships; thus, the space of *local* has stretched significantly. Consequently, places play an even more significant role in human interactions. Mobile communication popularity has even surpassed television, the most pervasive media ever before (Katz and Aakhus 2002). Mobile communication helps increasing the number of interactions providing, in comparison to co-presence, lower quality alternatives but faster ways of communication. Contrary to some provocative visions as depicted by Mitchell (1996), places still play an important role in human interactions (Höflich 2005), but even more the increased freedom of movements and the flexibility in communication allow people to selectively choose *places worth being in*. In fact, the quality and the importance of the interaction are the criteria used by people to choose amongst the several networking opportunities. As Mitchell himself suggests in his article *The revenge of place* (2005 p. 45), "the special qualities of particular places are more important, not less so, in the digital electronic era". Human interactions now can span from asynchronous communication (e-mail) to the more effective way of interaction that remains the co-presence and eye contact (Mitchell 2001). Those elements from a social point of view are the basis of the rise of the so-called *network society* (Castells 2010). Human connectedness is increasingly less tightly related to co-presence and the possibility to keep being connected allow us a wider freedom on mobility (Katz and Aakhus 2002). We have an increasingly *mobile society* (Castells et al. 2004). The role of Geography at a global and at a local scale rather than dismissed is enhanced by advances in ICT with several disciplinary analyses joining Mitchell's insights suggesting the *revenge of geography* (The Economist 2003) or as Kaplan (2015) from a political geography perspective writes "rather than eliminating the relevance of geography, globalization is reinforcing it". The investigation of places worth being in is the objective of CES research. While mobility allows people to choose where to be, researchers need to find ways to identify places people prefer. In this effort, researchers interested in ecology and sustainability need to open up their research methodologies and embrace new data sources, methods, paradigms, and technologies. The new ICT technologies provide us an enormous amount of data that help analysing people's preferences.

1.4 The opportunity offered by large amounts of data

ICT and networking technologies generate great amounts of data of different types, diverse origins, and at different speeds. The exponential growth of data created at the beginning of the century raised the prospect that the quantity of data could overcome the network capacity. The terms *exaflood* or *data deluge* (Swanson 2007, Markoff 2009, The Economist 2010) were coined to express this concern. The great volume of data, however, was perceived as a great opportunity and the term 'Big Data' has been coined to identify large volumes of data (Diebold 2012). A more comprehensive definition also includes variety and velocity as elements that characterise Big Data (Laney 2001). Variety refers to the inclusion of sensors and citizen originated information while velocity refers to the incessant flow of real time produced data. Wider and more comprehensive definitions have been drawn. Kitchin (2013) uses an expanded definition and his approach focuses on the potential of Big Data for cultural geography research. The variety of data in Laney's paper refer to diverse IT systems, standards, and semantics. In 2001 there were the first signs of another element that in the last decade increased the variety of data, namely the creation of user generated content.

1.4.1Web 2.0 and the socially constructed exaflood

In the last decade, the above-mentioned tendency for personal wider mobility became even stronger with the technological advancements applied over the global interconnected network of computers (the internet). New ICT led to the creation of an interactive platform called Web 2.0 (O' Reilly 2005). Web 2.0 is characterised by the ability offered to users to contribute to websites' content. This seamless two-way communication has been also used by companies to outsource simple tasks to an enormous number of individuals. With the advent of Web 2.0, online collaboration and social networking has been simplified. Organisations started building on the possibility to engage external people in executing micro tasks. Howe (2006) called this phenomenon *crowdsourcing*. Tapscott and Williams (2008) analysed in fine details the various kinds of mass collaboration enabled by Web 2.0. Another way that enabled companies to improve their productivity has been the ability to communicate more easily both internally and externally questioning employees and clients about improvements. A similar activity called collective intelligence (Lesser et al. 2012) has been implemented in IBM. Volunteered content shared online includes also social media. All the online collaboration endeavours have specific features. The technological innovations allowed the creation of information platforms that, providing services to communities, allowed them to pursue the aim of their project. User generated content not only contributes to the collection of big amount of data, but it enhances its diversity also in terms of intrinsic characteristics of its own kind. Specific characteristics that are found on all collaborative projects are the long-tail effect (Anderson 2006), enabling global diffusion for smaller and niche endeavours, and participation inequality (Nielsen 2006, Haklay 2016a) exemplified by rules similar to 90-9-1 where 90% of website users do not contribute to its content, 9% provide very limited content while 1% do most of the job. In terms of productivity, the 1% of contributors produces the 90% of content, the remaining 10% is produced by the remaining 99% of users. Those rules describe the relation between contributors and contribution. The term *produsage* (Bruns 2007) has been suggested for the process enabled by the Web 2.0 that results in the creation of what Brums called users-led content. While the term and the following use of *produser* for the users-producers of the web content are intriguing, given the above mentioned 90-9-1 rule it is more correct to distinguish between contributors and users of user generated content. Studies over the use of social media or photosharing projects created an accurate analysis of the audience involved (Duggan and Brenner 2013, Li et al. 2013) and can lead to interesting blending of those sources of information for a better understanding of people’s preferences. The new kind of data, its large amount and its potential use drove to the rise of novel approaches to scientific investigation.

1.4.2 Data intensive scientific research and crowdsourced information

The large amount of data produced by researchers, sensors, and users when published can lead to different research methodologies. The urge to promote a *big science* whose main feature consists in the publication of very large amounts of data has been promoted since the 1960s (de Solla Price 1963) and has been reinforced with the advent of the digitalisation of information and the Web 2.0 (Costello 2009). Henceforth a large amount of scientific data started to be published online. Together with the information produced directly in a digital format, to augment the available corpus of information, also the digitalisation of existing documents and data increased the information available. This allow scientists to extend to the past the innovative research methods and technologies that are emerging. The overall concept of big science leads to the definition of new methodologies for a data intensive computational science. John Gray from Microsoft research even defined a 'fourth paradigm' for scientific research (Hey 2009). Large amounts of environmental observations, remote sensing data and other kind of data produced during the scientific investigation were often mentioned in works referring to big science or the fourth paradigm (Hampton et al. 2013, Woodward et al. 2014, Raffaelli et al. 2014). But still the vast amount of data generated by users was largely ignored with minor opening towards citizen science[5]. There is still a focus on ecological studies that overlooks the opportunities offered by approaches that tend to link research and policy, as suggested in the first part of this chapter. Ecology and Cultural Ecosystem Services research can greatly benefit from Web 2.0 produced data if we also consider that part of this data has attached geographic information.

1.4.3 Spatially referred crowdsourced information

When social content is associated with a location or when the location itself is the objective of the social action new possibilities are open to understand people's spatial preferences. After Goodchild (2007a), the most used term that encapsulates all the projects that take advantage of crowdsourced geographic information are called VGI, which stands for Volunteered Geographic Information. This term is usually utilised interchangeably with Crowdsourced Geographic Information (CGI). Not only do users collect a considerable amount of information, but often they collect information without geographical content. For some projects, a geographic reference is added by someone else as a supplementary tag. At present, social geographical information is collected under several schemes and purposes. For some Web 2.0 projects the collection of geographic information is the primary target. Into this category can fall photo collections aiming at landscapes or land cover, collections of ratings and comments for touristic or leisure purposes, geosocial games, environmental observations, and the creation of geographic information for mapping and cartographic purposes. Some other times a spatial reference is a valued but not necessary element, for example when someone adds a location to a Wikipedia article. The combination of Web 2.0 collected information can open several options for the comprehension of the human-spatial relations. The

[5] But still not seen as a source of social-ecological information but as another source of data over the natural environment. See Panel 2 in (Hampton et al. 2013)

blending of a photosharing platform such as Flickr and a social media as Twitter carried out by Li et al. (2013) revealed insights to population profiling and characteristics of places. Similar geodemographic studies have been performed in the UK by Adnan and Longley (2013, 2016) and Adnan et al. (2013). The rise of social media technologies can enable existing communities to expand their collaboration online but can also contribute to the creation of new kinds of communities of interest related to the possibility to share and enhance digital content. When the digital content has a geographic reference, we can elicit spatial knowledge from this crowdsourced geographic information. The domain analysing crowdsourced geographic information is in rapid growth with a rising number of studies targeting several aspects of it. Some books and reports collected the most updated outcomes of the field (Sui et al. 2013a, Haklay et al. 2014, Arsanjani et al. 2015c, Capineri et al. 2016) but still more contributes can be added to the debate as this work will demonstrate. The possibilities offered by the new research methods and large amounts of data have already been explored and their use raises also new questions related to the role of theories (Elwood et al. 2013) will be at the core of CGI research (Wilson and Graham 2013). Amongst the several possibilities in which Web 2.0 enabled communities to collect spatial information as explored by See et al. (2016), environmental citizen science in the form of volunteer-led biological recording, is the only one explicitly relating humans, nature, and research. When spatial information is recorded when collecting data, we have Geographical Citizen Science (Haklay 2010a, 2013). While Geographical Citizen Science looks like a social-ecological system in a nutshell, in reality, volunteer-led biological recordings have not been used to elicit social-ecological systems. Amongst the various spatially enabled social media we can list other two families of projects: photosharing projects, whose aim is the collection and sharing of digital images, and collaborative mapping, which aims explicitly to the creation of digital geographic information for mapping and cartography. As seen before, the overall environmental concerns at a local and global scale are facing a lack of engagement and motivation for a sustainable development. The attempts to overcome the dichotomy between environment and development have been widely unsuccessful. The Ecosystem Service approach, then, focusing on the enhancement of the human-nature relationship can take advantage of the spatial information collected through crowdsourced geographical information.

1.5 Contribution: Crowdsourced geographical information and Cultural Ecosystem Services

We have seen that trends in different disciplines are evolving and that the value that can be elicited from their interactions and integrations is still underdeveloped. From one side, technological affordances widened and not all their potential has been employed or employed consistently or with rigour. The technological achievements related to size variety and heterogeneity of digital information are increasingly grounded. There is Web 2.0-created non-material content with explicit geographic reference and lineage. The elicitation of value from such sources for social/spatial analysis is therefore potentially subject to high levels of automation and can be performed at unprecedented spatial scales[6]. From another side, there is an increasing need to influence, inform and support policy and public engagement to preserve an environment favourable to human wellbeing. The most challenging aspect of conservation is related to the detection and valuation of non-material benefits that humans derive from ecosystems. Technological affordances and environmental policy making are connected then and the bonds between humans and nature are witnessed by their actions as they are also recorded in digital format. A new research perspective can therefore explore the relation between places and cultural ecosystem services using large, varied, and user generated digital information. We can infer that not only volunteer-led biological recording initiatives, but all place-based activities are closely related to cultural ecosystem services in a way that is usually overlooked on mainstream CES research. Following Dunkel (2015 p. 177), "the photograph that someone takes of a place, and the photographs they choose to upload may both reflect, in some way, their perceptions of the place". The connection between place-based activities like environmental citizen science and CES is given by the fact that the enjoyment of nature, the inspiration and the aesthetic enjoyment that have been listed amongst the constituents of cultural ecosystem services, are also among the main constituents of the motivation behind the participation in place-based activities (Hobbs and White 2012). Creating and sharing digital information about places invokes such feelings, and arguably generates CES. Sharing pictures, collecting environmental information, creating the digital map of an area, documenting the story of a place: all fall in this category. This link between policy designed concepts and individual actions can lead to very fruitful results. Crowdsourced geographic information is therefore strongly related to cultural ecosystem services. While all the other ecosystem services provide physical elements that can be somehow measured, valued, priced, cultural ecosystem services with their inner emotional nature are the drivers for the enthusiastic lovers of nature that are eager to reiterate their enjoyment, protecting, observing, investigating, sharing, mapping nature, and taking part in citizen science projects. In this work, the opportunity offered by new social morphologies will be matched with the new methodological and technological opportunities that help characterizing places. This work explores the potential use of vast amounts of user generated content to support cultural ecosystem services aimed research. This thesis will test the design of "data mills"

[6] To expand also temporal scales, we need digitalisation of past knowledge

that try to extract meaningful information from the exaflood, the data deluge. The "digital footprints, or data shadows or by-products of human/machine interactions" (Capineri 2016) will be analysed to demonstrate if crowdsourced geographic information effectively can be used to identify Cultural Ecosystem Services.

1.6 Research issues

The links between the domains and technologies as outlined above, while bearing an interesting potential, should be thoroughly investigated. An approach blindly reliant on technology and data alone could lead to erroneous or inconclusive results. To this day, the data quality assessments for Crowdsourced Geographic Information are tailored to evaluate its fitness for the purpose of using it for cartography and mapping. As our approach is almost unprecedented, we need then to create a solid foundation based on the knowledge of the processes that drive the creation of CGI data. We cannot rely on evaluations made using quality criteria that are extraneous to our purpose. Our intuition connects CGI and CES, but we should evaluate how CGI can contribute to CES research. This can be achieved also creating a deep understanding of concepts and methodologies used to map CES to date. The exploration of the two domains that we are connecting also passes through the search for early attempted contaminations and it will also drive us towards the better definition of our research methodology. Below we express some of the issues in the form of research questions that we can group in four levels. The first level aims at the understanding of the intrinsic characteristics of the individual domains and technologies previously mentioned. The second level aims at exploring transdisciplinary studies that already attempted approaches similar to ours. The third level aims at the design of the current study. The fourth level characterises the result of our analytical phase. Below are the levels and research questions:

INTRINSIC CHARACTERISTICS OF THE INDIVIDUAL DOMAINS:

Technological affordances: *What is the potential, nature, and limitations of UGC?* CGI is a subset of UGC and therefore it inherits from it some characteristics. The identification of them will help shaping the most appropriate approach understanding how community dynamics shape and influence data. The identification of embedded limitations and biases will allow us to find the potential, to interpret and to design the most appropriate way to use CGI for CES research.

Experiences and best practices: *What are examples of uses of CGI to elicit social-territorial bonds?* The call to open up to new methodologies and disciplines stresses the fact that beyond ecology other disciplines may have been developed studies to characterise cultural benefits from places. Therefore, lessons can be learnt from those studies in terms of data already used, methods, tools used or developed, spatial coverages and results. Those elements will inform all the stages of our research, from data selection and preparation, methodology design and will help in contextualizing and evaluating results.

How CES research is conducted and what are the best practices? We are approaching this work with the perspective of the ES approach. Therefore, we need to be aware of the most relevant advancements on valuations experiences and disciplinary frameworks developed to stress how nature contributes with non-material benefits to human wellbeing. This will enable us to create a common ground in terminological and ontological aspects with researchers in the field, identify open issues and limitations to the current practice.

EXAMPLES OF CROSS CONTAMINATIONS ALREADY DONE

How CGI can fit in ES research? The previous questions created the foundation for our innovation. With this question begins our contribution. Once a disciplinary overview has been completed separately, both in CES and CGI domains, we should envisage how CGI can effectively support CES research. Knowing the peculiar features of both domains, this will enable us to identify which limitations of current practice in CES research can be complemented by the design of an approach who takes advantage of CGI peculiar features.

What are the early uses of CGI for ES and CES? There are already some examples of the use of CGI for CES research. How to identify and characterise them is paramount to learn from them and to develop our study. A review of methods, approaches, and results will help define our approach towards CGI, data preparation, and analytical methods.

DESIGN OF THE METHODOLOGY

Is there the possibility to design a methodology that allows us to compare the different CGI sources? We are going to investigate distinct types of CGI. We envisage that, given the different levels of coordination and the multiplicity of platforms and engagements, even similar projects will show different individual features. To compare the projects and to keep this comparison consistent with best practices in both CGI and CES research, we will need to identify a minimal set of tests that allows us to create a comparative framework. This will enable us to consistently characterise each project inside the group of CGI projects that we are going to analyse.

CHARACTERISATIONS AND FINDINGS

The methodology designed will be directed towards the characterisation of CGI data we are using. We are going to see how similar CGI projects behave, the relation between them, which aspects of engagement for places they can be the proxy for and how the outcomes of the analytical phase can help us in implementing, if possible, automated methods to support CES research. Therefore, the following research issues will arise from the analysis of the CGI resources:

Do individual CGI projects inside the families work similarly? This issue is very relevant because it will not only allow us to characterise the projects, but it will also show us how much they are complementary and interchangeable. Local patterns of success may occur but also some CGI projects may end their lifespan. Knowing the degrees of similarities of the different projects we will be able to evaluate the impact on CES evaluations of localised absence of a project in space or time.

What are the aspects of places that can be elicited using CGI? Human experience of the environment is multifaceted; the MEA definition of CES includes already several of them. We do not expect that all of them can be detected using a generic CGI. We expect to find a range of CES benefits addressed by every individual project. Some CES aspect may be elicited through some specific methodologies applied to specific CGI projects while some aspects may still be undetectable using CGI.

Is it possible to implement a fully automated large-scale study to elicit CES? The evaluation of CGI resources and their ability to elicit CES for large scale studies is highly reliant on the possibility to implement tools able to derive estimations of CES with reduced or no human supervision and intervention. Our work will test the CGI resources in small scale areas to reduce computational challenges[7], and will be able to draw recommendations on the potential for a large-scale implementation of CES estimation using CGI resources.

What are the limitations that must be considered evaluating our work? We must be aware that not all aspects of CES will be detectable through CGI; it may happen that no aspect of CES can be elicited by some CGI sources. We may end up detecting only spatial aggregations due to social phenomena with no ecological component at all.

What are the potential developments? Our evaluations may lead to the definition of several developments. Depending on the results of our analysis we can provide useful guidance for several communities. We may develop recommendations for CGI communities, find bugs, or suggest another potential use of their work that may lead to widening the community. As a result, we could write guidelines for the development of tools for large scale evaluations of CGI to detect CES. A kind of *Big Data* for CES. We may also find out the components of success of specific areas and related activities, deriving suggestions for site managers of conservation areas. We could also find out areas appreciated by the public but less known to conservation organisations. Finally, we could evaluate how a specific CGI project may be designed to gather information for better detection of CES.

[7] Usually overcome in short time by the quick improvements of hardware performances

1.7 Thesis structure

The next chapter, will explore the current literature on Ecosystem Services and Culture Ecosystem Services focusing on their mapping.

Chapter Three will be a review of the studies that used crowdsourced geographic information stressing when they tried to elicit people's preferences.

Chapters Four and Five will explain the methodology, data, and places chosen as well as the research questions that will be targeted.

Chapters Six to Eight will show the three case studies developed respectively in south Wales using photosharing crowdsourced information, in London using citizen science, and in south east Wales using CGI projects where collaboration is explicitly encouraged through the enhancement of unique articles of features to describe information.

Chapter Nine will discuss the main results of this work while chapter Ten will draw the conclusions and underline the potential of the study.

Literature review section

In this section we introduce the target domain of our work in chapter 2. And then in chapter 3 we show the practices and technologies we propose to support the identified research domain.

2 Ecosystem Services and Cultural Ecosystem Services

2.1 Introduction

The research domain gear towards the use of the Ecosystem Services (ES) approach who values the benefits that nature provides to human wellbeing. The effort to accommodate environment and development, while also stressing the link between human well-being and the natural environment, led to the development of the Ecosystem Services approach. This urged the development of a wide range of assessment and valuation methods (in section 2.2.1) that nonetheless are subject to critiques and limitations 2.2.2.

Cultural Ecosystem Services (CES), that are the focus of this work, are, amongst the classes of ES, the most anthropocentric and the less studied (section 2.4). The benefits they provide to human wellbeing are characterised by non-materiality therefore their analysis and evaluation needs to be pursued beyond the typical domains of life sciences (section 2.4.1). Most recently a bespoke framework (in section 2.4.2) and the adoption of a wide range of evaluations (section 2.4.3) are widening the horizon. The scope of the ES approach is to provide assessment methods to support policy decisions. Methodologies to analyse the spatial component of ES therefore have been widely developed (as in 2.5) but when dealing with CES the relation between material spaces, benefits and beneficiary is utmost complicated (section 2.5.2). Several approaches have been attempted as shown in (2.6) starting from the ones more based on habitat and land use seen as proxies or as components in models (section 2.6.1) that are vastly used for large scale projects which do not engage with residents and stakeholders.

To date, the most successful methodologies to detect CES are based on small scale mixed methods approaches. They engage people and often use methodologies borrowed from social and landscape centric studies (section 2.6.2). There is an emerging research trend favoured by the social and technological developments that have been shown in the previous chapter. People engagement for the natural environment is detected through the secondary and interpretative analysis of crowdsourced geographic data. This new or only moderately exploited approach has some interesting but inconsistent early studies (section 2.6.3). In the next chapter 3 we explore in more details the practices, technologies, and relevant applications that are behind the production of Crowdsourced Geographic Information.

2.2 The Ecosystem Services approach

The conceptualisation of Ecosystem Services dates back to the late ’70 (Westman 1977), cited in Gómez-Baggethun et al. (2010). It has gone through several refinements: One of the earliest and most cited definition is:

> “Ecosystem services are the conditions and processes through which natural ecosystems, and the species that make them up, sustain and fulfil human life.” (Daily 1997 p. 3)

Some conceptual clarification is needed since often the term ecosystem service is used to denote every ecological process while it is intended differently in the conceptualisation used to promote and shape a better management of ecosystems. Therefore, a better distinction between functions and services is needed:

> "The terms ‘ecosystem service’ and ‘ecosystem function’ are largely interchangeable… although ecosystem services can be defined as ecological processes that benefit people, whereas ecosystem functions can be considered as all ecological processes regardless of whether they are beneficial to humanity" (Luck et al. 2003 p. 331).

The need to build on this clear-cut definition to better drive monitoring efforts drove Pettorelli et al. (2017) to develop an in depth ontological distinction between ecosystems processes, functions, and services starting from their relations with the overall concept of biodiversity. In this comprehensive classification, it is clear how ES are the one providing benefits to humans. On the same issue Costanza et al. (2017), summing up twenty years of research in the field, stress how the existence of beneficiaries of a specific ecosystem function make it a service. The large-scale adoption of the ecosystem services conceptualisation followed the Millennium Ecosystem Assessment (MEA) (Millennium Ecosystem Assessment 2003, 2005). Its definition links human well-being and ecosystems aiming at creating an evaluation framework for a better management of ecosystems. This in the effort to fill a gap in scientific assessment ability that emerged at the end of the twentieth century (Potschin et al. 2016a). The full definition of Ecosystem Services used in the MEA is:

> “Ecosystem services are the benefits people obtain from ecosystems. These include provisioning services such as food and water; regulating services such as regulation of floods, drought, land degradation, and disease; supporting services such as soil formation and nutrient cycling; and cultural services such as recreational, spiritual, religious and other nonmaterial benefits." (Millennium Ecosystem Assessment 2003 p. 3)

Noteworthy is that some definitions call ES processes that delivers goods or benefits, while others call ES a benefit in itself. This is why further frameworks have been developed. In Figure 2-1 the relation between the ES conceptualisation and the components of human well-being according to the Millennium Ecosystem Assessment are presented.

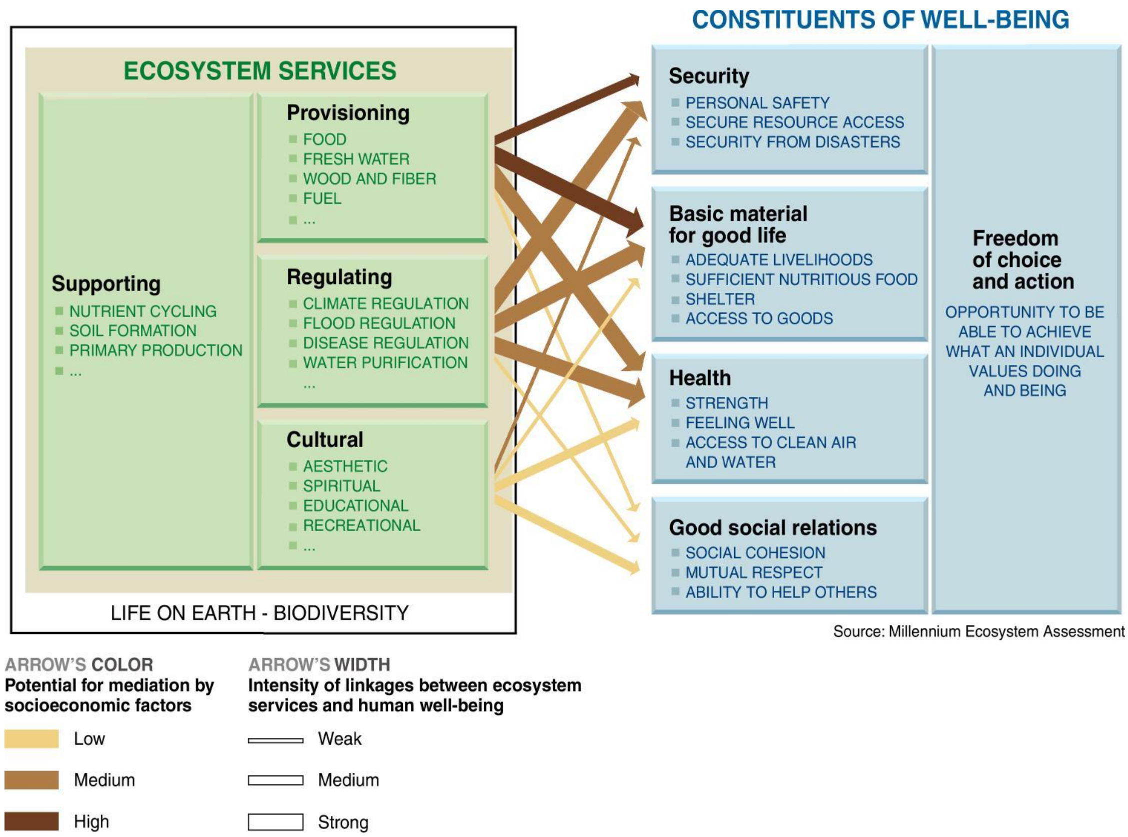

Figure 2-1: Linkages between Ecosystem Services and Human Well-being (Millennium Ecosystem Assessment 2005 p. vi)

Another relevant definition is the one in the TEEB (The Economics of Ecosystems and Biodiversity) where ES have been initially defined as the benefits that people obtain from ecosystems (European Communities 2008). The current version on the TEEB website[8] mentions ES as direct and undirect contributions and conflates goods and services. Fisher at al. (2009, 2011) have chosen a specific definition of ES[9]:

> "ecosystem services are the aspects of ecosystems utilized (actively or passively) to produce human well-being. The key points are that 1) services must be ecological phenomena and 2) that they do not have to be directly utilized." (Fisher et al. 2009 p. 645)

This definition is more focused on actual use of ecosystems and excludes symbolic and cultural aspects of the human experience of nature. There have been several classifications for ES (Costanza 2008, Haines-Young and Potschin 2011, 2016) depending on the discriminating factor chosen and the conceptualisation for ES chosen. Haines-Young and Potschin (CICES V4.3 2012) developed a Common International Classification of Ecosystem Services (CICES). In the ES approach a pivotal role is played by the concept of human well-being, which is defined as:

> "Human well-being has multiple constituents, including basic material for a good life, freedom and choice, health, good social relations, and security. Well-being is at the opposite end of a continuum from poverty, which has been defined as a

[8] http://www.teebweb.org/resources/glossary-of-terms/
[9] Derived from (Boyd and Banzhaf 2007)

> “pronounced deprivation in well-being.” The constituents of well-being, as experienced and perceived by people, are situation-dependent, reflecting local geography, culture, and ecological circumstances" (Millennium Ecosystem Assessment 2003 p. 3)

The inclusion of the environment in the well-being evaluation is coherent with the Human-Scale Development Matrix (H-SDM) developed by Manfred Max Neef (1991) where relationship harmony and distance from nature are considered amongst the factors to evaluate for human development. This definition of well-being stresses locality and contingency, this should be enhanced with more anthropocentric evaluations:

> “Ecosystem services are intrinsically anthropocentric, and resulting arguments for conserving nature to ensure for their continued delivery are in addition to, not in place of, ethical and scientific ones.” (Fisher et al. 2008 p. 2051)

Human well-being, human benefits, ethical and scientific values, everything turns around human needs and evaluations. This stresses the need to incorporate in the ecological tradition focused on natural sciences new analytical methodologies to underline the importance of human-nature interaction. Therefore according to the ES perspective, the explanations that fit this framework are not exclusively natural-ecological functions but also human knowledge of natural processes, feasible uses or ongoing consumption of processes and products of the natural world (Spangenberg et al. 2014). This meets the Natural Capital (NC) cascade process, as described by (Haines-Young and Potschin 2009a, 2011) that describes how the biophysical structures and processes deliver benefits hence well-being to humans. The cascade approach (Figure 2-2) start from natural capital on the left-hand side and then follows a cascade process up until the definition of the delivered benefit and its value. This approach is not based on the relevance for any specific group of receptors. It is based on a nature centric conceptualisation. There might be a barrier between the function and service boxes. The barrier might express the fact that not all ecological functions produce services that contribute to human well-being. We think that this conceptualisation and diagram express the potential of ecosystems to produce benefits.

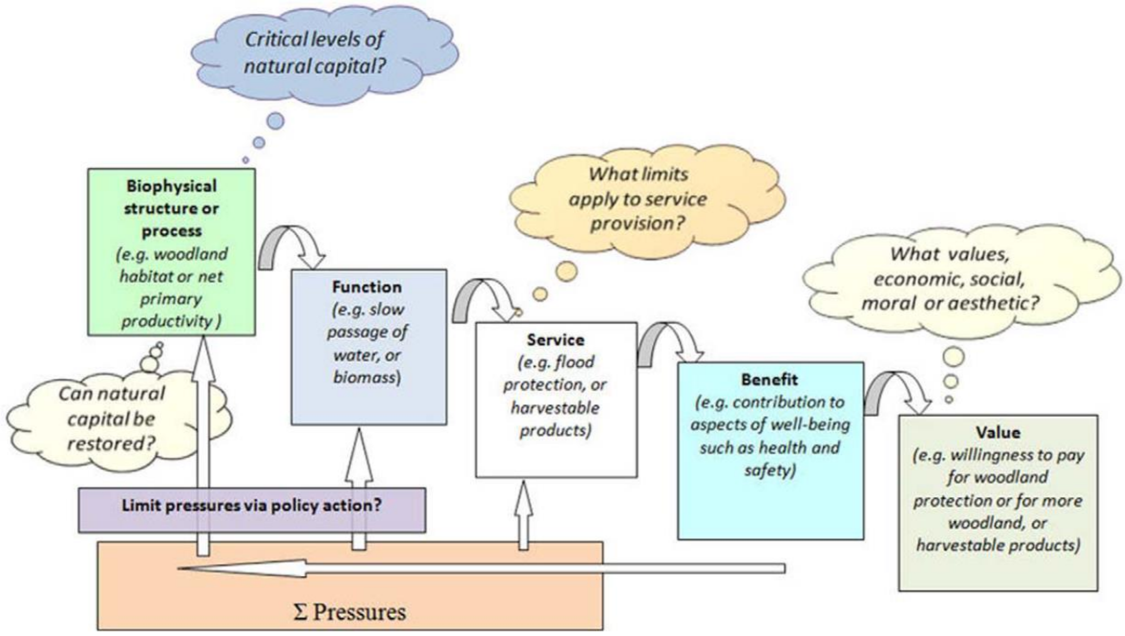

Figure 2-2: The ecosystem service cascade model from (Potschin and Haines-Young 2011)

According to Costanza et al. (2017 p. 5) this conceptualisation represents "oversimplification of a complex reality and an unnecessary complication of what is essentially a very straight-forward definition". They propose a more complex diagram where the interactions between social, human, built, and natural capital are more explicitly included as in Figure 2-3 whose layout stresses multifaceted interactions that arose ES. It is not NC in itself that arise ES but meaningful interactions between natural, built, human, and social capital.

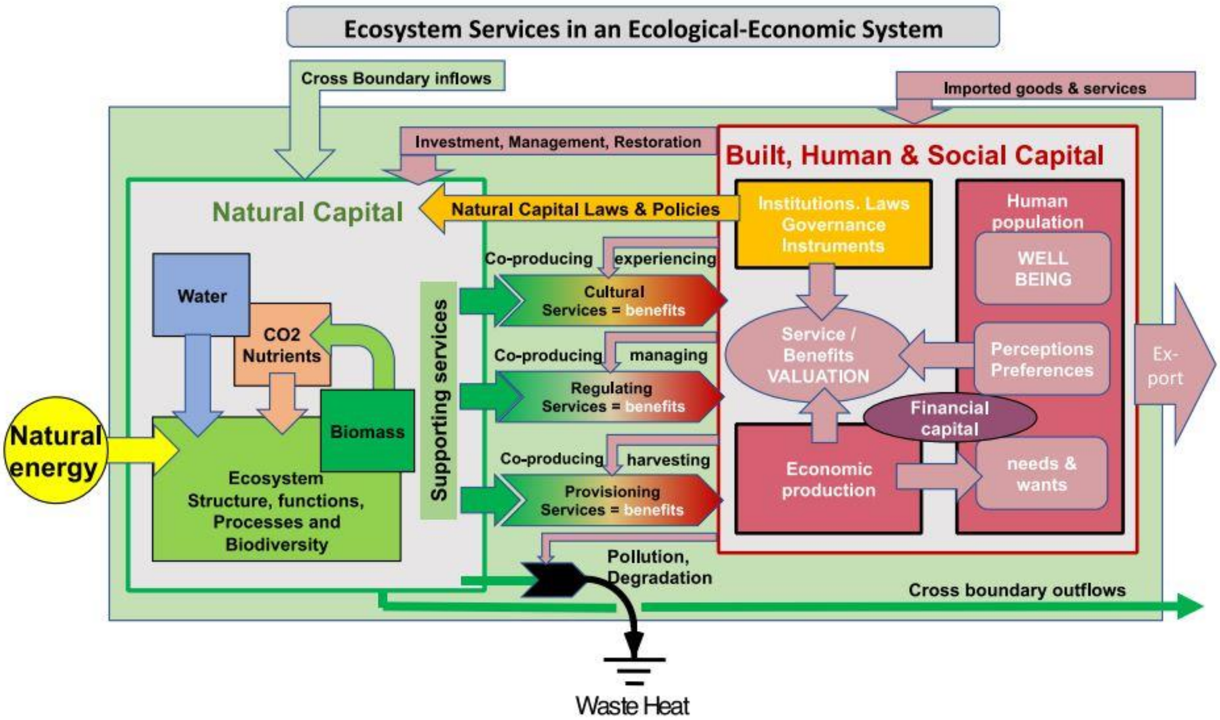

Figure 2-3: The interactions needed to produce ES (Costanza et al. 2017 p. 6)

This approach portraits ES as interactions between NC and built, social and human capital, therefore they exist only through interactions with humans. This lead to the importance of stakeholder involvement for all ES assessments. The essential role played by the interaction with humans means that ES exist only when there are receptors, or only where the ecosystem function is accessible to humans as Daily poses:

> "For human beings to realize most of the aesthetic, spiritual, and economic benefits of biodiversity, natural ecosystems must therefore be accessible." (Daily 1997 p. 5)

Accessibility as well is a main discriminant between the ecosystem functions that play the role of ecosystem services, providing benefits for human well-being, and ecosystem functions that are not providing benefits to human wellbeing.

Another relevant conceptualisation is the beneficiary approach, proposed by Boyd and Banzhaf (2007) and then developed by Landers (Landers and Nahlik 2013, Landers et al. 2016) for the USA National Ecosystem Services Classification System (NESCS) (EPA 2015). The beneficiary approach underlines the role of the Final Ecosystem Goods and Services (FEGS) as the ones that directly reach beneficiaries. In a similar conceptualisation Fisher et al. (2009) stresses the distinction between ecosystem services and benefits. While keeping the distinction between ecosystem functions and services which allow them to also characterise the ES through the benefits that contributes to human well-being. FEGS include also services that physically never get to beneficiaries but that are valued by them. This approach does not ignore all intermediate ES but creates a framework for a more effective measurement and better-informed debates focusing on ecosystem functions that are actually delivering clearly defined benefits. Subjectivity and dependency on context are at the forefront of this conceptualisation that tends to mould ecosystem services in a more economist-friendly discourse mimicking the marginal analysis methods. The FEGS and NESCS approach stresses the importance of the identification of beneficiaries or users of the ES making them more easily identifiable, avoiding multiple counting for the same ES and producing a more compelling methodology for economic accounting. While this approach embeds several advantages compared to more natural capital focused approaches it still needs a localised analysis of users and may underestimate FEGS that do not have currently beneficiaries associated to them. To add clarity in CICES (CICES V4.3 2012) there is an explicit reference to final ecosystem services. They are neatly distinguished from goods and benefits (together are named ES products). The CICES classification created a hierarchical four-tiered (Sections, Divisions, Groups, Classes) tree structure for all ES. The top level is called Section and includes the three classes Provisioning, Regulation, Maintenance, and Cultural. Another approach that embeds the identification of beneficiaries on its accounting system is the one that stresses the existence and analysis of Social-Ecological Systems (SES) (Loft et al. 2016). Glaser et al. (2012 p. 4) defined social-ecological systems as

> "complex adaptive systems, consisting of a bio-geophysical unit and its associated social actors and institutions. The spatial or functional boundaries of the system delimit a particular ecosystem and its problem context."

This approach is based on ecosystem functions and actors and its core analysis focuses on socio-ecological structures and practices that are context and time dependent. Those elements are deeply rooted in social structures and practices and have a considerable non-material component. To use the social-ecological analytical framework transdisciplinarity is strongly advocated. This is the reason behind the use of the boundary object concept for ES (Star and Griesemer 1989). The identification of boundary objects enhances effective ways of cooperation not only between disciplinary communities (Abson et al. 2014) but mainly enhances the deliberative process that greatly contributes in the identification of social-ecological entities and functions (Schröter et al. 2014, Kenter et al. 2016, Schröter and van Oudenhoven 2016). The use of boundary objects helped in the creation of common frameworks to integrate the different perspectives. This is also effective in bolstering collaboration and deliberation. Potschin and Haines-Young (2016b) show how different assessment frameworks lead to different analytical methodologies. Their cognitive and cultural roots and the way in which local settings assign value to ecosystem functions and goods are the overarching background for the ES perspective. This means that data used to model ecosystems gain different relevance in different context according to the perceptions of people as will be shown in the following section.

2.2.1Ecosystem Services assessments values and valuation methods

As noted, the ES perspective is suggesting a way to assess and make accountable the role that nature plays for human well-being. Costanza et al. (2017) stress how "being more explicit about the value of ES and NC can help society make better decisions in the many cases in which trade-offs exist." (Costanza et al. 2017 p. 8). The development of assessment methods and procedures follows a dialogue with stakeholders since local and situated knowledge should be taken in consideration. But ES assessments are also based on transdisciplinary discourses. Disciplines have their specific body of knowledge and languages. Hence the need to develop frameworks that allows the integration between the different perspectives. The wide range of physical and cultural elements that are included in values and valuations methods using the ES approach require valuation methods that are multidimensional, comprehensive and include a plurality of values (Gómez-Baggethun et al. 2016, Costanza et al. 2017). Therefore, a classification of values and a characterisation of valuation methods have been proposed. Following the literature, we assume three main categories in ES values, ecological, sociocultural, and economic values (Gómez-Baggethun et al. 2016, Oteros-Rozas et al. 2017). Connected to this categorisation two main valuation methods groups have been proposed. One is based on the criteria used to elicit value and another is based on the way the elicited value is measured. Therefore, we have valuations based on the methodological procedure focused either on biophysical assessment or on human principles and preferences. The second classification is based on the distinction between monetary and non-monetary valuations. It is important to stress how values, valuation methods and resulting ES are not streamlined since valuation methods can contribute to several ES.

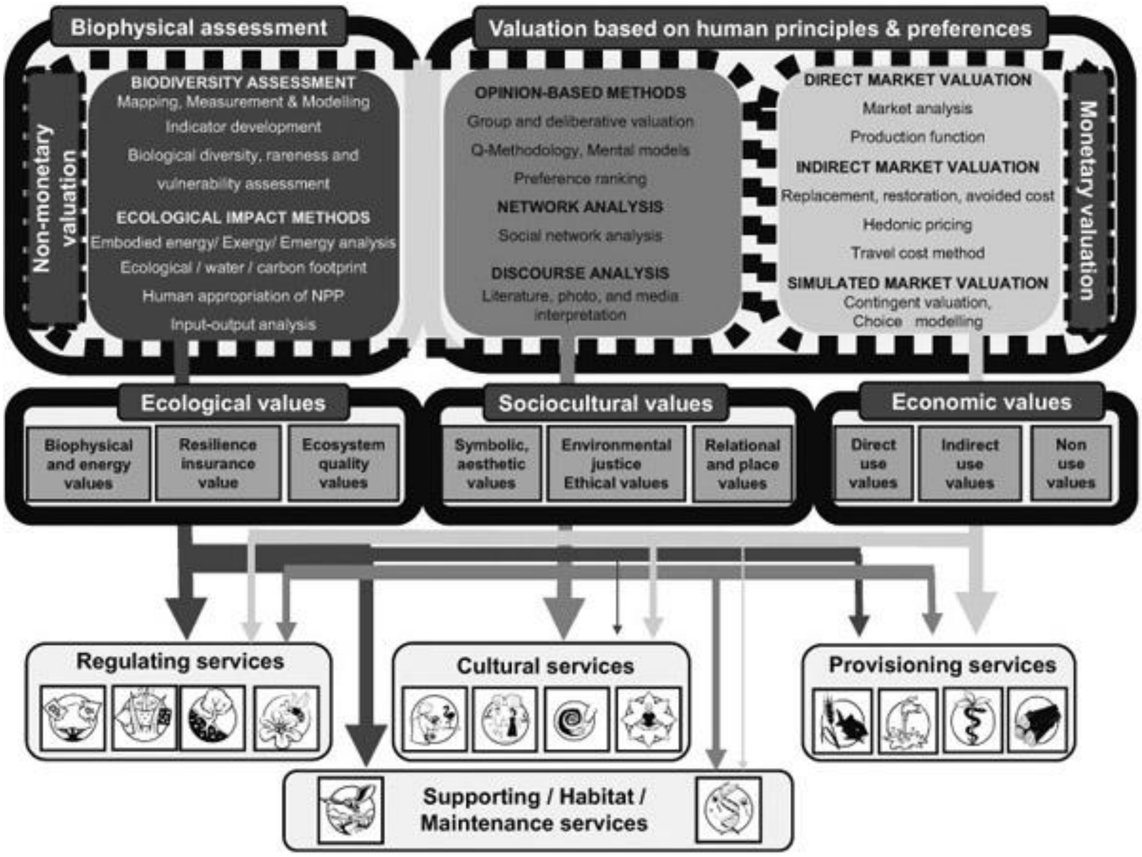

Figure 2-4: Ecosystem services value types and valuation methods. From Gómez-Baggethun et al. (2016 p. 103)

In Figure 2-4 the framework we are describing as proposed in (Gómez-Baggethun et al. 2016) linking values, valuation methods and ES. Ecological and economic values are at the basis of the most developed and well-known methodologies. Such studies are heavily based on biophysical elements and in the case of economic values include the longer lasting disciplinary hybridisation between ecological and economic studies having originated monetary valuations used in environmental impact assessments and especially in cost-benefit analysis. As for the social-ecological framework (Loft et al. 2016), sociocultural values and their associated valuations are a very heterogenous set of methodologies often mediated through social, psychological, and cultural disciplinary studies. Sociocultural valuations include five categories: deliberative valuation, psycho-cultural valuation, social valuation, qualitative valuation, and subjective assessment[10]. Costanza et al. (2017) in their review stress how the existence of public goods in most of the ES require the evaluation methods to be more inclusive and open up to include shared values (Kenter et al. 2015). Another approach towards ES assessment is the use of models. This mainly derives from the literature on modelling natural systems and they developed from the biophysical assessment as in Figure 2-4. This equates only on estimating the source of ES. Developing on such experiences several tools have been developed to assess ES hence including also considerations on flows and receptors. Tools have been collected, compared, and analysed by various studies (Bagstad et al. 2013b, Vorstius and Spray 2015, Burgess et al. 2016, Kienast and Holfenstein 2016, Maes et al. 2016). Noteworthy is the effort of the Ecosystem Knowledge Network (EKN), a network of institutions interested in the implementation of a ES based agenda. One member of the network, the Joint Nature Conservation Committee (JNCC), published an online Tool Assessor[11] and a report (Howard et al. 2016) analysing the status of the art of ES support systems. Several metrics have been developed to compare the 12 tools. Most of them do not have a GIS component (7) while most of them consider spatial data. The report mentions studies that connects geospatial data publicly available like land use and land cover datasets to ES. Having obtained an overview of the overall ES approach and its main valuation methods, we have seen how the field is characterised by multiple views and internal debates. In the following section we show how the conceptualisation and some of its frameworks and methods give rise to contentious debates.

2.2.2Critiquing ES

Several critical elements are associated with the ES approach. Most of them are summarised by Mace (2016), and we focus on some of them, to better understand the ES approach.

HUMAN – NATURE DUALISM: From a general perspective over all the ES framework Cooper et al. (2016) stress how the vision and most striking the “cascade model” (Haines-Young and Potschin 2009a, 2011) places human and nature at two opposite ends. In doing so it stresses the artificial dualism that it is intended to confront. Cooper et al. (2016) also criticise the utilitarian perspective used by the ES approach arguing that aesthetic appreciation is not always concerned to evaluate natural elements as resources. Yet, we have to remember that the ES approach has been endorsed not

[10] Valuation methodologies are also gaining raising interest because apart from helping shaping management policies they are increasingly used in legal litigation processes involving environmental damage.

[11] http://ecosystemsknowledge.net/resources/tools/tool-assessor

because it rightly represents the human-nature relation, but it constitutes a boundary object that mimic the mindsets of the sections of society that less value nature.

NATURAL CAPITAL AND ES: One of the less evident but often misguided elements of the ES approach comes from the confusion between the concept of Ecosystem Services and Natural Capital. As Mace (2016) explains, while capital is the physical component, the service unfolds, and is the flow of benefits. Therefore a (natural) capital that does not produce a flow is not related to a (ecosystem) service. There is not a one to one relation between the two concepts. While every ES spawns from NC, not all NC generates ES. The ES perspective does not see NC as an antagonistic conceptualisation. The ES perspective stresses the origin of services thus providing a way to estimate the potential of NC as a source of ES even when they are currently locally not valued or perceived as valuable stocks. Focusing on services we can also stress the undervalued potential of the capital. Alternatively, we can understand threshold values associated with the production of services from the capital, informing and improving management practices.

SOCIAL INEQUALITIES: Mace (2016) recalls Daw (2011) when discussing the social biases in the provision and enjoyment of ES. Social and economic rather than cultural barriers can lead to exclude substantial portions of populations from enjoying a specific ES (private woodlands, access rights in hunting reserves). Costanza (2016, 2017) also stresses that the focus has been on the balance between nature and society rather than the balance between different sectors of society, social equity is also usually missing in ES assessments.

MONETARY VALUES DO NOT FIT ALL: Another critique is associated with the anthropocentric and economic perspective, and the tendency to render values through monetary market values. A paper from Silvertown (2015) especially epitomised this criticism. To place this critique more in perspective we must stress how economic valuation of ES mainly considers marginal values (Badura et al. 2016) as global assessments are considered ineffective in accounting for total value of ecosystems (Costanza et al. 1997). Marginal values are used to draw programmes evaluations to evaluate policy scenarios, forecast changes and trends as explained in (Bateman et al. 2011). Silvertown (2015) stressed only on one of the several methods available to account for nature's benefits in the ES approach as the only one used. Hence, he elicited from this assumption a focus on economics, use values, neoliberism and market values driving to the only instrument of Payments for Ecosystem Services (PES). Doing so he overlooked not only the Millennium ecosystem assessment but also the vast majority of ES focused studies that stress a variety of values. The responses to Silvertown (Potschin et al. 2016c, Schröter and van Oudenhoven 2016, Wilson and Law 2016), apart from reducing Silvertown's assumptions to the small niche of ES research to whom they are better aimed at, agreed on some limitations and potential development of the approach.

The concern of Silvertown is in some way shared by the signatories of the Antwerp declaration on ES ("The Antwerp Declaration" 2016). As also stressed in similar ways in Gómez-Baggethun et al. (2016) and Costanza et al. (2017), economic value often based on willingness to pay, is just one of the three main value categories used in the ES approach. Moreover, the respondents stressed how Silvertown suggested an approach aiming at valuing conservation and nature for its intrinsic value. But this demonstrated to be unpractical to influence policy in the previous decades.

SUSTAINABLE DEVELOPMENT AND ES: Another critique focuses on the perceived disjunction between the ES approach and the aim of sustainable development. This has been stressed in literature (Potschin et al. 2016a) and is one of the basic aims that is supported by the Antwerp declaration ("The Antwerp Declaration" 2016) as well as in other relevant works (Bennett et al. 2015, Costanza et al. 2017). Sustainability is the whole concept that stresses economic, ecological, and social elements while the elements of social justice have been overlooked in some ES focused studies, this is why the Antwerp declaration reiterates the often-repeated call to promote collaboration across disciplinary boundaries and *"intelligently pluralistic approaches"* (Costanza 2016 p. 19). The social component of sustainable development forces also to broaden valuations designed for ES including not only individual but also social wellbeing (Costanza et al. 2017) towards an economy that may be focused more on human wellbeing beyond Gross Domestic Product (Costanza et al. 2016). The ES approach is set to support the mainstream economic system with another set of values and evaluations but with its focus only on well-being and ecosystems it is missing to voice ecosystem friendly forms of development. What is termed as natural capital is intended as a capital only in ecosystem terms while it can be proposed also as a capital for economic development. As an example, the cultural elements that contribute to shape the local identity, the traditional management practices, the traditional knowledge systems, sense of place, could play a role in developing and rediscovering products, management techniques, local cultures that can be themselves source of non-consumptive forms of economic development. The ES valuations can help understand the thresholds of those alternative and sustainable ways of promoting development.

In the following section, we deepen our understanding of the conceptualisation focusing on non-material benefits from nature. The overall objective of this work is to contribute to the detection of them.

2.3 The local, the social, the temporal: Cultural Ecosystem Services

The ES approach therefore is the attempt to link more tightly science and action. It has been an attempt to express scientific environmental concerns in policy-relevant ways

(Carpenter et al. 2009). ES are difficult to express and measure and their influence has always been underestimated since decisions are often based on market values. Following Carpenter et al. we can underline that

> ” Most decisions are based on market prices, but for many ecosystem services no markets exist, and decision makers have no clear signal for the value of the services. Understanding the true social value of non-marketed ecosystem services depends on the ways that services are used by different stakeholders.” (Carpenter et al. 2009 p. 1308)

Ecosystem services focused research is therefore local grounded and population-targeted. This is also stressed by the fact that ethnic groups as well as economic development of societies influences the perception of nature or of some elements of it (Agyeman 1990, Clarke and Agyeman 2011, Kenter et al. 2014b) as well as the perception of landscapes (Natural England 2009)[12]. Moreover, the difficulty in covering the trade-offs amongst different ES lead to the conclusion that

> “We focus on the need for networked, place-based long-term social–ecological research, opportunities to learn from existing programs, and needs for improved monitoring” (Carpenter et al. 2009 p. 1309).

Social-ecological complex networked and local grounded systems are therefore the analysis units to focus on to provide meaningful insights for decision makers. Social and local elements play then a pivotal role in the ES approach. The term local refers to perceived identity of places. It stresses how ES have a relevant cultural component. Cultural Ecosystem Services play a pivotal role in the ES conceptualisation. They are the most anthropocentric and challenging category of ES. They are associated with sociocultural values (Gómez-Baggethun et al. 2016) and are amongst the less investigated. As Milcu et al. (2013) reported few studies were devoted explicitly to CES, and when performed they mostly focused on the easiest to measure (recreation and ecotourism) rather than focusing on the ones most important to people. A further conceptual understanding was strongly advised, and the work of the UK NEA helped developing a new framework as it will be explained in section 2.4.2. First, we look at the concept of CES in the next section.

2.4 What are CES

Daily et al. (1997) do not use the term CES but mentions ecosystem as capable of providing, amongst others the "provision of aesthetic beauty and intellectual stimulation that lift the human spirit" (Daily et al. 1997 p. 4). Following (Daily 1997, Daily et al. 1997) mention a deep appreciation of natural ecosystems. The arts, religions, and traditions as well as several activities that contribute to the aesthetic beauty and intellectual and

[12] in particular Box 12,2 authored by Hazel Wong in (Chan et al. 2011).

spiritual stimulation provide the evidence to these aspects. In the MEA CES are defined as:

> “the nonmaterial benefits people obtain from ecosystems through spiritual enrichment, cognitive development, reflection, recreation, and aesthetic experiences” (Millennium Ecosystem Assessment 2005 p. 40)

And then a list of examples aggregated in the following categories: Cultural diversity, spiritual and religious values, knowledge systems, educational values, Inspiration, aesthetic values, social relations, sense of place, cultural heritage values, recreation, and ecotourism. In the ontological clarification developed by Pettorelli et al. (2017) it is explicitly stated that the cultural aspect of ES is human centred therefore they exclude the existence of cultural ecosystem functions. The CICES classification (CICES V4.3 2012) identified two divisions for CES (one for physical and intellectual interactions and another for symbolic interaction) and from there 11 classes here including also existence and bequest. The UK NEA adopted a definition for ES that is derived directly from the MEA. The definition of CES in the follow-on version is:

> "cultural ecosystem services are the contributions ecosystems make to human well-being in terms of the identities they help frame, the experiences they help enable and the capabilities they help equip."(Fish et al. 2016b p. 212)

Also in this case, then, clarifications are needed. CES in the cited literature and elsewhere have been labelled in many ways: non-material, non-use, non-consumptive, intangible, all definitions look inversely associated with action and matter therefore it conveys the impression that CES look less identifiable and measurable compared to other ES. We now turn to some of those attributes attached to CES to better understand their characteristics. As expressed by Gee and Burkhard (2010) and cited by Fish et al. (2016b p. 210)

> “[t]he physical environment is a mere bedrock of perception; intangible value is assigned by adding cognitive and imaginative overlays.”

Interestingly several studies proved that CES

> "are intimately linked and the lens through which many other ES derive meaning" (Chan and Satterfield 2016 p. 343).

CES play a pivotal role in the ES conceptualisation since it is through human cultural knowledge and associated decisions that an ecosystem good is considered a benefit. Even the most material ES is imbued with cultural judgments and knowledge systems. Or as put by Fish et al. when describing deliberative techniques to elicit ES:

> "ecosystem services in all their variety are effectively being characterised as objects of collective, shared and social values, and therein, as services of cultural concern." (Fish et al. 2016b p. 215).

In the very similar conceptualisation advocated by Costanza et al another characteristic of CES is stressed, in fact most of them "consist of a mix of private and public goods.

Public goods are non-excludable and multiple users can simultaneously benefit from using them" (Costanza et al. 2017 p. 8). Therefore, CES focused research must be most inclusive and consider thoroughly social and shared values.
In economic valuation terms (Badura et al. 2016), CES can pertain to both use and non-use values. More precisely while in economic valuation terms use values are more relevant. CES are usually associated to direct non-extractive and all options of non-use values. But CES still play a role in use, indirect, and option values. From an economic valuation perspective CES are also challenging since while the mainstream economics tends to value in monetary terms CES, like culture, are predominantly of non-economic character (Fish et al. 2016b). But CES are spawn from almost every interaction with nature even the ones intended to harvest the other ES. Chan and Satterfield (2016) make a compelling case of the interwoven relationship between skills, experiences, learning and emotions related to a traditional fishing practice of native Americans. Another peculiar characteristic of CES is its continuous development. While the need for nutrients, climate regulations and similar ES is given and constant in time[13], social and cultural constructs are more dynamic and change over time, they can even depreciate (Chan and Satterfield 2016). What is today inspiring and worth conserving for sustainable development, might tomorrow be perceived as outdated and in need of redevelopment. Therefore, another important accountable variable to include in the evaluations that deals with CES is "time." CES are specific, local, and can be also temporary. A recent literature review and bibliometric analysis by Schaich et al. (2010) concluded that cultural ecosystem services have been assessed only marginally and therefore propose to link ecosystem services research with cultural landscape research to fill the knowledge gaps and to stress more how human agency has a long tradition of co-evolution with natural ecosystems. According to Kienast and Helfenstein (2016) the term landscape service is more appropriate and might be used for several purposes. Willemen et al. (2012) explains that landscape service is used to identify the ES provided by a landscape. While landscapes provide a relevant amount of CES they cannot be considered covering all of them. The term landscape is semantically related to visual metrics that while are relevant for place-based approaches they are not all encompassing. The European Landscape Convention defined landscapes as a place of perceived identity thanks to the interaction of humans and nature (Council of Europe 2000). When a so wider perspective is used than landscapes can be considered as bearing all meanings of CES. Given that this is not the most common use of the term landscape, as suggested by Frank and Walz (2017), landscape metrics cannot be considered as the only tools to derive CES. Yet, as Scaich et al. (2010) stressed, the role of a cultural landscape research agenda for CES

[13] Unless scientific discoveries change the way we understand, value, and measure their contribute to our wellbeing.

is important. In the following, while largely relying with landscapes we will not use the term landscape services, since we will not focus only on the visual elements. In terms of locality CES are strongly related to local settings. In economic terms mainly, non-use values may decay differently with distance (Bateman et al. 2006). Bateman et al. (2013b) modelled CES as affected by distance decay but they are differently decaying for users rather than non-users. Another intriguing element that characterises CES is the relation between how spiritual values are represented in natural sciences centred documents and how they are represented by individuals that perceive this relation in studies considering spirituality, arts, and religion. An interesting perspective starting from a western culture background is analysed in Cooper et al. (2016). Spiritual bounds to nature are often part of a bundle of feelings and moral duties where the need for care and protection are inextricable. Similar considerations drove Cooper et al. (2016) to the conclusion that CES are a very peculiar case of ES where the artificial interpretation of human bonds with nature cannot be represented by the separation conceptualised by mainstream ES approach and the cascade model more explicitly. The varied nature of CES and its interpretation of social, cultural, and spiritual values and benefits makes their understanding and analysis a very complex endeavour. Their analysis needs to go well beyond the disciplinary borders of natural sciences as we detail in the coming section.

2.4.1Interdisciplinarity and CES

Interdisciplinarity is strongly recommended for the correct understanding of the ES approach as noted above. As we have seen, this is especially relevant for CES. They are often named social-ecological systems even when they do not refer to the specific framework as described in Loft et al. (2016). But even engaging social science analytical approaches do not suffice; a multidisciplinary approach is strongly recommended.

> “Research in this hybrid field requires integrated theoretical approaches and methods that go beyond simply employing existing concepts from natural and social sciences in a multidisciplinary approach” (Loft et al. 2016 p. 91)

Russell et al. (2013 p. 475) widen the field of investigation with a call for engagement with several disciplines

> "There is a pervasive, visceral understanding that our nonmaterial connections to ecosystems provide rich benefits, but systematically analyzing these connections must be done at the interface of disciplines as disparate as behavioral science, philosophy, art, medicine, anthropology, history, and ecology"

Then in the case of CES a more open dialogue is needed not only with other scientific fields but the debate on human well-being and cultural engagement should cross the borders with arts in general and humanities as the explorative works developed by Coates at al. (2014) for the UK NEA. The call for wider disciplinary collaborations and ontological clarifications to understand CES has been developed during the development of the UK NEA. In the following section we briefly introduce their main outcomes.

2.4.2The new framework

Fish et al. (2016b) stressing the peculiarities of CES in comparison with the other ES explain the framework developed during the follow-up phase of the UK National Ecosystem Assessment (UK NEA FO) (Church et al. 2014, Fish and Church 2014). Developing on the previous approach (Church et al. 2011) that for CES tended to characterise *environmental settings*[14] they developed a new ontological perspective. CES, like almost all the hard to monetise values, can be elicited through deliberative and non-monetary evaluations (Kenter 2016a). Therefore, in the framework proposed for the NEA FO, CES are intended not as preformed but they are

> "relational processes and entities that people actively create and express through interactions with ecosystems." (Fish et al. 2016b p. 211).

The definition of CES used in UK NEA has already been reported in 2.4. The UK NEA FO most notably ditched the reference to the non-materiality of CES[15]. In the latest version non-materiality is considered only a characteristic of cultural benefits. A visual representation of the proposed framework is reported in Figure 2-5. A very detailed explanation of this ontological overview is beyond the scope of this work and is widely covered in the UK NEA FO report (Church et al. 2014) and in (Fish et al. 2016b). We here stress the main concepts that explain it and that can be useful for this work. Overall, following the necessary ontological classifications that we have underlined, it distinguishes between values, services, goods, and benefits. The mutually exchanged arrows underline how this framework is designed to stress the relational nature of CES. In doing so this diagram and the conceptualisation behind it express an ontological matching with the similarly interaction based conceptualisation proposed by Costanza et al. (2017)[16]. Cultural goods are here to be interpreted as arising from the interactions between the main blocks of services, benefits, and biophysical domain. The services themselves make up the physical and the non-material component of CES while the benefits are neatly separated and remain non-material. There are four main blocks in the diagram: the services themselves are characterised by the relation between a physical component, named *environmental spaces*, and the non-material component, named

[14] "Are the locations and places where humans interact with each other and nature that give rise to the cultural goods and benefits that people obtain from ecosystems" (UK National Ecosystem Assessment 2011b)

[15] In the previous version of UK NEA, the definition for CES was " The nonmaterial benefits people obtain from ecosystems through spiritual enrichment, cognitive development, reflection, recreation, and aesthetic experience, including, e.g. knowledge systems, social relations, and aesthetic values."(UK National Ecosystem Assessment 2011b)

[16] In essence Costanza et al. (2017) propose a multifaceted view on ES and valuations that more closely resemble the CES focused sections of (UK National Ecosystem Assessment 2011a, 2014) rather than the cascade model.

cultural practices. While both those components shape the biophysical domain, it in turns provides materials and opportunities for them. A neat characterisation of benefits as suggested by Fisher et al. (2009) is evident. The cultural benefits that contribute to human well-being are enabled by CES, but they also shape the physical and non-material components of CES.

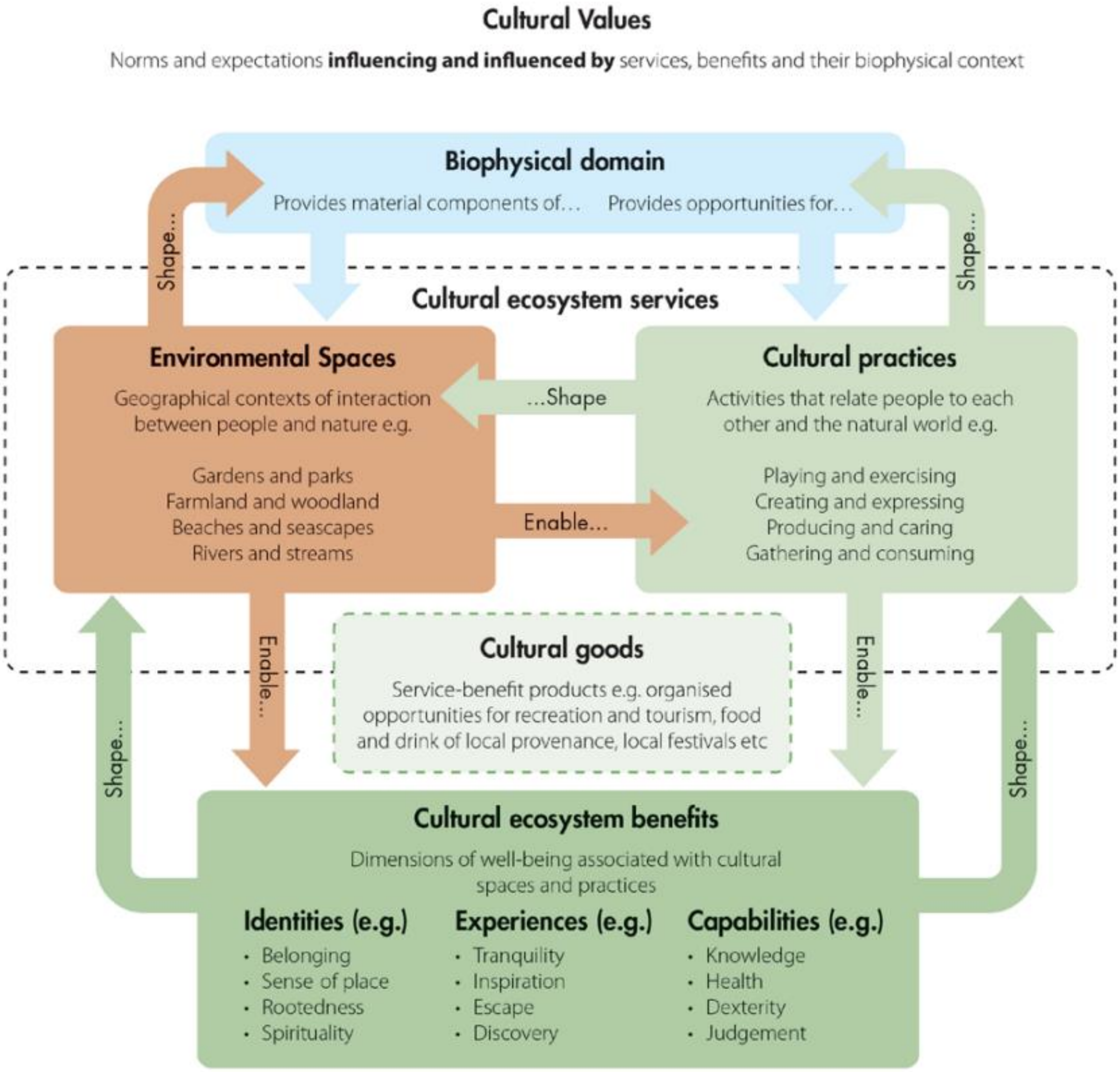

Figure 2-5: The UK-NEA framework for CES from (Church et al. 2014)

The physical components in this framework are confined not only to the environmental spaces but also to some of the cultural goods that, apart from some non-material component, can be attached to several physical elements when identity and belonging can play a role. Consequently, land cover focused maps, often and successfully used to identify and analyse the other ES, fail to grasp CES. For CES the biophysical perspective and the economic perspective needs to explore more methodologies to identify and evaluate CES. The ontological clarification shown above is helpful in describing the various components of CES, but pushing out the benefits from the definition of CES will imply the use of another term that embraces the services and the benefits. For the rest of this study, we will adhere to the conceptual framework in Figure 2-5(Church et al.

2014, Fish and Church 2014, Fish et al. 2016b). In the next section we will focus more on the valuation methods used specifically for CES.

2.4.3CES evaluations deliberative and non-monetary

The goal of CES valuations is to represent more complex shared social values, thus diverting from utilitarianism and individual gains, as expressed in neoclassical economics. ES valuations can even fail when they are used for commodification of nature (Gomez-Baggethun and Ruiz-Perez 2011). It can also cause the resentment by people that feels that the other non-economically evaluated values are disregarded. As Chan and Satterfield report:

> "The valuation in monetary terms triggers concerns about commodification, in part because dominant socio-political trends and the particular institutional settings for ES (ES markets and payments for ES) push simultaneously toward valuation and commodification " (Chan and Satterfield 2016 p. 345)

To overcome the limitations of monetary valuations methods, deliberative non-monetary valuations have been developed and adopted from a wide range of disciplinary domains. Studies from arts and humanities, social psychology, sociology, anthropology, geography, development studies, and ecological economics contributed to the creation of a diverse toolkit to analyse the relation between society and nature. From a management point of view, when benefits are not mediated through markets, Chan and Satterfield (2016) stress how evaluations are more difficult and how locality is very important for CES. All the value setting methods require participation of interested stakeholders and several levels and methodologies for engagement can be identified (Fish et al. 2016d). Kenter (2016a) shows a range of deliberative, analytical-deliberative, psychological, and interpretive approaches to value the environment. These methodologies are well developed and rely on long lasting traditions of social and ecological sciences. When dealing with shared and cultural values the deliberative process is often seen not only as a value elicitation process but is considered a value forming process in itself (Irvine et al. 2016) for which a specific methodology can be developed (Kenter et al. 2014a). Spatial bound methods and psychometric methods, while valid on their own, can complement interpretive approaches since CES often comes in bundles with cultural heritage and biodiversity (Kenter 2016b). Interpretive methods are very effective when dealing with the elicitation of local situated knowledge. They can be easily integrated with more deliberative methods. While providing support from existing cultural representations

> "desk-based interpretive methods can provide a rapid and relatively low-cost analysis of values for different ecosystem services. For example, cultural history studies or analysis of creative literature can reveal particular cultural representations and values associated with places." (Kenter 2016a p. 283).

To recognise the importance of multiplicity of values, we need to use multiplicity of valuation methods. It is also important to stress that different valuation methods can be more suitable for distinct stages of the policy creation process[17]. At an initial survey stage Kenter (2016a) suggests using, amongst other methods, public participation geographical information system - PPGIS, (social) media analysis and existing datasets. The combination of those kinds of data in evaluating new possibilities to support interpretive studies will be the focus of this work. In the following sections we will describe methods and approaches to map ES and more specifically CES.

2.5 Spatial extents

To detect, evaluate, and compare ES in all their elements there is the need to define their spatial extent. Maps and cartographic methods play a very important role in the analysis and communication of ES related knowledge.

2.5.1Spatial characterisation of ES

As noted, the cascade or stock flow receptor model (Haines-Young and Potschin 2009a) is among the common framing of ES and they were used to develop the CICES classification (Haines-Young and Potschin 2011, CICES V4.3 2012). Having described ES as the benefits originated in ecosystems that contributes to human well-being we see that a spatial characterisation of ES can be applied to both the source, the flow, and the receptor. What distinguishes ES from Natural Capital is that a service is delivered while the capital is there regardless of the delivery to some receptor, it may or may not provide a service to people. Therefore, mapping ES is not simply mapping natural resources but mapping stocks of natural resources that are contributing to human well-being. This distinction is clearer when we consider the beneficiary approach of Boyd and Banzhaf (2007) and Landers (2013, 2016). The ES approach considers the full supply chain from stock to beneficiary (Maes et al. 2016). Considering this approach, the spatial and temporal dimension of ES are very complex because the stakeholders involved and those reached by the ES are not necessarily only current and local (Fish et al. 2016d). Diversity in spatial scales and approaches are also advocated by Brown and Fagerholm (2015) especially when dealing with CES. The spatial complexity is reflected in the difficult selection of the spatial extension but also of the analysis units. Palomo et al. (2014) suggesting an approach to conservation that disregards borders between conservation areas and the rest of the territory, noted how the ES approach is particularly apt to embrace such a vision.

[17] We would like to report here how Costanza et al. (2017) to pursue sustainable outcomes through the use of the ES approach, advocate the use for all ES of the valuation methods here proposed to target CES.

Several studies have been performed on ES assessments and they have followed mainly the three operational approaches proposed by Potschin and Haines-Young (2013, 2016b). According to their classification we have a *habitat or land cover* approach, a *systems or model based* approach and a *place-based* approach. The transdisciplinarity and the involvement of stakeholders increases along the classification.

Following the first approach the most common way to map ES is to map ecological units derived from habitats. Most of the time habitat maps are derived from land use or land cover maps. For Kienast and Holfenstein (2016) this approach can be seen as a static simplification of the model based approach. Medcalf et al. (2012) explored the use of land cover data to link cover to habitats and then derive ES, and Dales et al. (2014) developed maps for all ES in England using a 1km² grid. In both cases, some satisfactory result has been achieved for almost all ES. But the mapping of CES ended up being the most critical as will be explained below.

The second approach, the system or model based approach, models the functioning of the ES to compute some explanatory variables and maps. Changing parameters to the model there is the possibility to draw alternative scenarios and test alternative hypotheses. Bateman et al. (2013b) demonstrate this approach in a study for the development of UK NEA derived maps for all ES. They used a more complex approach, compared to Dales et al. (2014), with a more heterogenous set of data and added the influence of transportation networks to compute the movement of receptors. Land use data was still playing a pivotal role in their study. Bateman et al. (2013b) performed an evaluation of the influence of ES-generated values through the creation of six different scenarios of land use change up to the year 2060. Their study stressed the importance of ES-aware planning decisions in increasing the overall values for ES when the evaluation embraces several ES. The design of the scenarios considered the effects of climate change. The study provided outcomes that were very clear and were presented using maps that covered the whole UK, with a spatial resolution of 2km². The detection of spatial entities was not the aim of these maps, which were the geographical support in combining distinct levels of spatial information to understand the influence of locality on overall policies. Kienast and Holfenstein (2016) proposed an evaluation method for model based projects that can be used for the comparison of studies.

The last approach proposed by Potschin and Haines-Young (2013, 2016b) is the place based approach. The place based approach is participatory and targets social-ecological services. This approach engages local communities and may lead to the co-creation of the conceptual, the analytical, and operational frameworks. This approach has not only the advantage of being truly participative and inclusive from an assessment point of view but also contributes to the elicitation of values for all bundles of ES. Hence it creates a more solid basis for the operationalisation of the assessment and the choice of policies.

Place based assessments are usually applied to limited areas as well as landscapes. Kienast and Holfenstein (2016) explained their evaluation metrics also comparing two local based-projects, Willemen et al. (2012) and Fagerholm et al. (2012). Several studies followed the place-based approach that traces back its implementation to landscape focused research and cultural geography (Schaich et al. 2010, Fagerholm et al. 2012, Frank et al. 2012, Bieling et al. 2014, Queiroz et al. 2015b). The careful selection of stakeholders as proposed by Fish et al. (2016d) belongs to this approach. We can now focus more on the features that characterise spatially CES and the methodologies used to derive their spatial extent.

2.5.2 Spatial characterisation of CES

One element that especially characterises CES is their spatial extent. As Hernandez-Morcillo et al (2013) stated the quality of CES indicators is highly enhanced when explicit mapping is included in the deliberative process. Given that CES are not necessarily bound to biophysical elements but to local situated knowledge of people, their spatial extent is inherently not confined to conservation areas. A too strict management of conservation areas led by wrong assumption on nature can disrupt CES and led to the impoverishment of ecosystems (Martinez 2006). As for the mapping of ES a common approach is to map natural features, and therefore natural capital, rather than considering markets and beneficiaries (Cowling et al. 2008, Haines-Young and Potschin 2009b). Bagstad et al. (2013a) present a study on the spatial dynamics of ES where the specificity of CES is emphasised by some unique features. Their carrier type is information, some of their sources and sinks are unlimited, and their flow is not depleted by biophysical features. Dealing with recreation, which is the only CES they analysed, they use a preference based approach for CES and they stress the complexity of flow analysis. As mentioned earlier Fisher at al. (2009, 2011) do not consider CES as properly fitting the definition they have chosen for ES

> "we define ecosystem services to be about ecological phenomena (e.g. not cultural services which we see as very valuable benefits derived from ecosystems and services)" (Fisher et al. 2009 p. 644)

therefore, they produced a visualisation of the spatial extent of ES that targets only biophysical elements (Figure 2-6). But all of it is in contrast with their own assertion of the utility derived from non-use of biodiversity listing the *existence values* (Fisher et al. 2009, 2011).

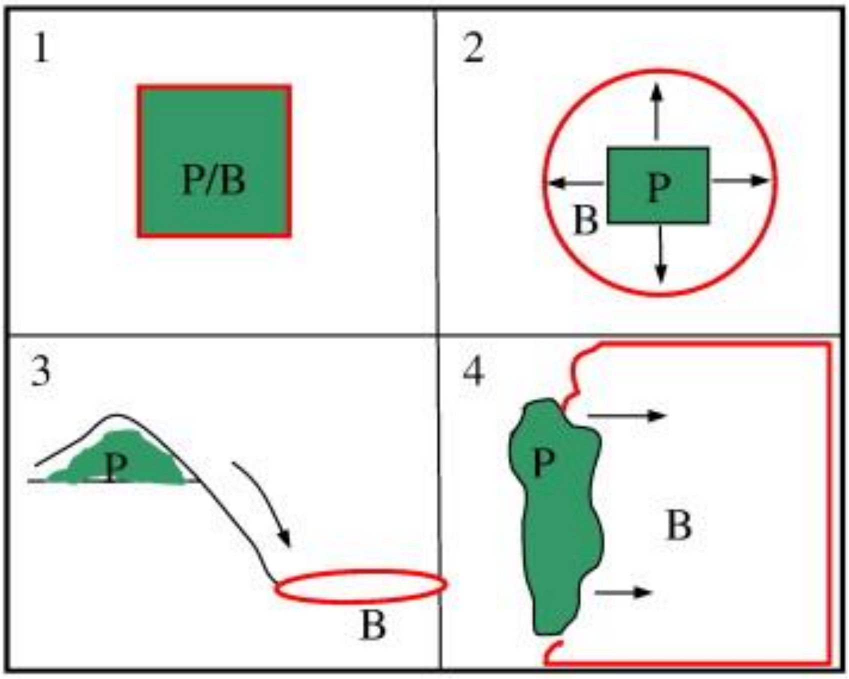

Figure 2-6:Spatial relationship between service Provision (P) and service benefit (B) areas. from Fisher et al. 2009 p. 650

Going beyond this conceptualisation here we focus on the possibility of mapping the spatial extent of CES. The UK National Ecosystem Assessment (UK National Ecosystem Assessment 2011a, 2014) used the term *environmental spaces* to denote the spatial component of CES[18]. Studies of CES vary widely in terms of spatial analysis and mapping techniques. Costanza (2008) proposing a classification of ES based on their spatial characteristics allocates CES in two groups. It is the only ES that is here characterised by two spatial characteristics. The first one is *Global non- proximal* and the second one is *User movement related* where we have the flow of people to unique natural features. Ambrose-Oji and Pagella (2012), while focusing on the assessment of woodlands, made a very clear analysis that considered spatial aspects in a way that can be used in other CES analysis. In their analysis they inherently consider the spatial characteristic *user movement* as defined by Costanza. Most of the studies mentioned by Ambrose-Oji and Pagella (2012) deal with a bundle of ES rather than CES only. Here, receptors might never have had any physical connections to the ecosystem that ignited their cultural benefit. Therefore, Ambrose-Oji and Pagella (2012) referred to the *ecosystem service footprint,* which is not a unique, well-bounded spatial entity, but rather, its shape and extent depend on stakeholders' spatial position, education, preferences, and age. The cultural ecosystem service footprint corresponds therefore to the environmental spaces of the adopted framework (2.4.2). Fish et al. (2016a) reported how active cultural practices can be age specific and influenced by the physical capabilities

[18] In 2011 the spatial extent of CES was named "environmental settings" and the change of terminology is semantically backed

of people. López-Santiago et al. (2014) found out that the appreciation of cultural aspects of a landscape is strongly related to "their rootedness and identification with the place, environmental knowledge, recreational needs, and economic dependence on the place" (López-Santiago et al. 2014)[19] therefore they designed three specific groups of stakeholders, which included also non-local people. The age specific, familiarity and subjective appreciation of different parts of the landscape has also been documented in a study precursor of the UK NEA in terms of CES (The Research Box 2009). Similar conclusions were drawn in other studies as in Plieninger et al. (2013). People who experience the aesthetic enjoyment associated with a woodland live in its vicinity. But this enjoyment is also experienced by those who are willing to move toward the woodland and by people accessing representations of the woodland through the media. The "ecosystem service footprint" for CES overlaps with the environmental spaces as defined by Fish et al. (2016b). Developing from Fisher at al. (2009, 2011) Ambrose-Oji and Pagella (2012) propose a difference nuance of the conceptual model to draw the spatial extent of CES that we provide in Figure 2-7.

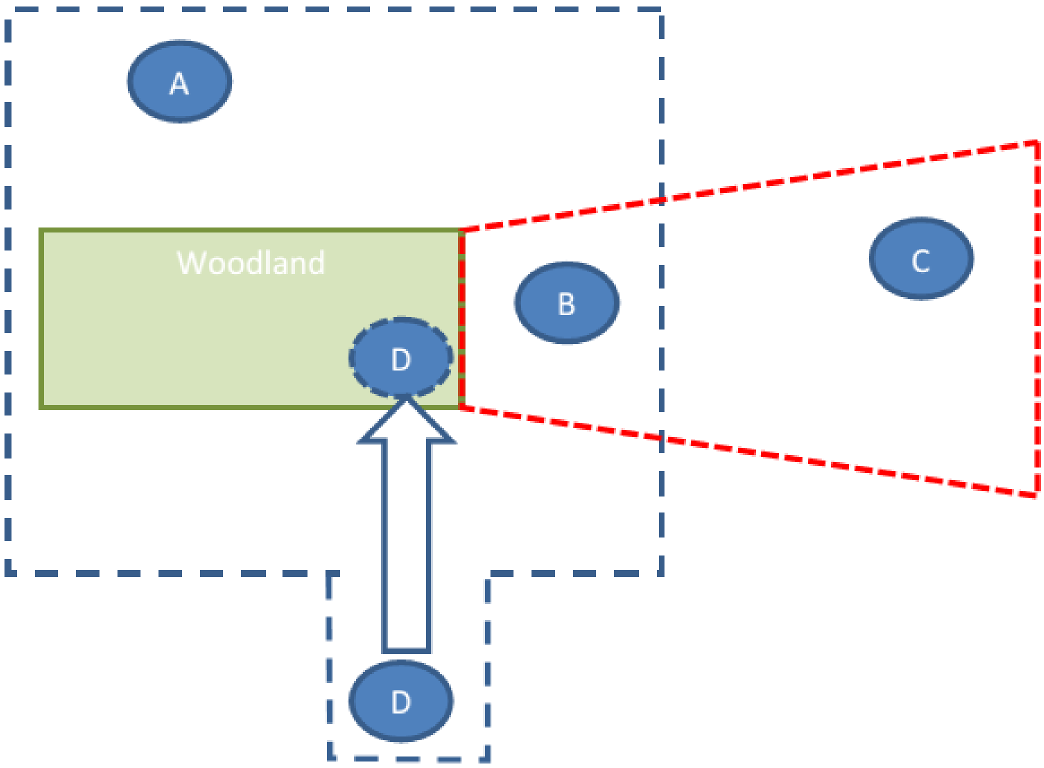

Figure 2-7: Conceptual model of the spatial extent of CES from (Ambrose-Oji and Pagella 2012 p. 40)

While the green area represents the point of provision, the hashed lines stand for alternative ES footprints related to the same ecosystem, the areas of reception then, the receptors (ellipses) are also physically movable. Therefore, receptor D enjoys the CES

[19] Ecology and Society publishes papers without page numbers

and somehow extend the blue dashed line when at the end of its trip he comes back to its starting point. Receptor D may even be really far from the environmental space, but it may have a visual connection to it. This representation still ignores the potential of information as a carrier that can have effect on people that never visited the provision point. This can be considered another peculiar spatial characteristic of CES as defined by Costanza (2008). Therefore, we can consider that videos, pictures, and even stories and legends in which a woodland (in the case of Ambrose-Oji and Pagella) play a role contributes to the creation of a cultural ecosystem footprint.

One of the main conclusions of Ambrose-Oji and Pagella (2012), as well as of UK-NEA FO (2014), is that the most appropriate methodologies with which to derive the spatial extent of CES are participatory mapping techniques. Participation might also include a broad range of stakeholders (Fish et al. 2016d), avoiding focusing on physical proximity as a way to measure the cultural impact of ecosystems. Ambrose-Oji and Pagella (2012) also underlined the inadequacy of land use change or census units to map CES. One of the main issue is the uncertainty of the extents for CES since neither natural nor census or administrative boundaries suffice in defining social effects and interactions who follows social and community boundaries enabled by accessibility. Therefore, mainstream methodologies to derive CES are not revealing the full impact of an area when defining the spatial extent of them.

> "data generated through typical interview techniques which do not integrate maps, photos or fieldwalks into the process do not deliver spatially explicit data, making it hard to relate non-material values to a specific place or feature of the physical landscape" (Bieling and Plieninger 2012 p. 651).

In statistical geographic analysis similar problems are referred as Modifiable Area Unit Problem (MAUP) (Openshaw and Taylor 1979). In the following section, we will show the studies that already attempted to define the environmental spaces.

2.6 Empirical studies mapping CES

In current practice, the investigation of CES has implied the integration of landscape ecology, landscape aesthetics, social sciences, and environmental or natural sciences (Schaich et al. 2010, Bagstad et al. 2013b, Hernández-Morcillo et al. 2013, Plieninger et al. 2013). Wider spectrum disciplinary collaborations are suggested to incorporate more elements of ecosystem services (Loft et al. 2016). The review of Milcu et al. (2013) shows how the focus on CES was weak and how not only conceptual but also methodological improvements were required. Kienast and Helfenstain (2016) developed a 6-point decision framework to evaluate the different components involved in ES models. They admit that CES constitute a different category. The limited reliance of CES on the physical components of space make them less prone to be modelled. Pettorelli et al. (2017) acknowledging the peculiar anthropocentric and cultural nature of CES excluded them from their assessment of ecosystem functions that can be detected using remote sensing. Nevertheless, in some studies statistical correlation, although weak, has been found with some physical elements mainly concerning recreation areas close to urban areas. In the following, we will show some of the main studies on CES mapping using the classification of methodologies according to Potschin and Haines-Young (2013, 2016b). We will merge their first two categories that imply large scale, wide area, and mostly desk based automated studies - the habitat or land use methodology and the system or model based methodology - together. This is because those methodologies are not well suited for CES. We will dedicate a specific section to the place based approaches with a concluding section dedicated to the studies with an innovative approach.

2.6.1Habitat and models approach

This is the approach used in the mentioned studies that used land cover, habitats, and modelled evaluations to estimate CES. Kienast and Helfenstain (2016) devoted a specific section to the modelling of CES. Mentioning all the unique characteristic of CES they acknowledge the difficulty in modelling and predicting CES. To this category pertain studies mentioned above focusing on the UK from Medcalf et al. (2012) and from Dales et al. (2014). Both attempts gave unsatisfactory results. Medcalf et al (2012) deemed those unsuccessful attempts to the lack of understanding of CES. They stress how data representing cultural services are often inaccessible mainly due to an insufficient understanding of them (Medcalf et al. 2012). As Fish (2011 p. 674) stressed is at least problematic to consider a "linear and deterministic relationship between ecosystems and culture" since "services and benefits do not simply arise from ecosystems, but are co-constructed through the interaction between people and their environments." (Fish et al. 2016a p. 330) or according to Chan and Satterfield (2016) such linear relationship is considered both laughable and offensive by researchers in culture. Therefore Dales et al. (2014) used a different approach more similar to the one used for the creation of scenarios in the UK NEA. They produced two maps of England, one based on the habitats and a second following the first version of the UK NEA (2011a) where not only land use but 14 different datasets have been combined. But when such maps were used in consultation with residents for a specific case study it becomes clear how they were not fit for mapping CES. Dales et al. (2014) embracing for CES a wider range of geospatial data, thus including designations, mention in the concluding notes of this work the inaccuracy at small scale geospatial data as the weak point of their approach. Alongside other recommendations there is the recognition of the limited usability of both the land cover and the habitat based approach for CES. It hints that alongside the UK NEA FO approach, the use of social media and innovative technologies can provide an

understanding of the subjective[20] experience of nature associated with CES. Bateman et al. (2013b) enhanced this study georeferring the Monitor of Engagement with the Natural Environment (MENE) survey (Natural England 2010). and adding network analysis and census data. They used a 2km^2 grid over all the UK. For CES they only considered recreational aspects. Bagstad et al. (2013a) proposed an alternative model of the spatial components and the relationships among them for ES. Using artificial intelligence, they distinguished between provisional and preventive ES, and tried to include competition among usage areas. Because of the specific features that characterise CES, some of these innovations have limited influence for mapping CES. Willemen et al. (2012) analysed cultural services using a cultural heritage study from the Dutch government. They developed a region wide study applying their metric then to a 100m grid. While the heritage landscapes are considered suppliers in the model they are negatively impacted by any land use change and they are considered as stationary in terms of demand. Moreover, any land use change between 1900 and 2000[21] is considered detrimental. Cultural heritage, computed this way, resulted as a burden for all management options. This model only considered the cultural intended as aesthetical appreciation of the historical landscape. According to this model we may think that even "rewilding" policies can be considered detrimental for cultural heritage and CES.
Sinclair and Pagella (2014) reviewed ES mapping techniques, comparing 50 published studies in which the importance of land use change in agricultural land was reported. Among these studies, few explicitly dealt with CES, mainly covering the stock of natural resources in which the services originated. Sinclair and Pagella (2014) found a very limited number of studies analysing flows and stakeholders. They found that recreation and tourism are the most often analysed CES. The mapping effort not only relied mainly on analysis units that were hardly suitable for the analysis of ES, but also lacked an assessment of uncertainty when dealing with the boundary of cultural entities. The MAUP affects several of this kind of studies since spatial data is never created purposely to analyse CES, therefore the aerial units reflect nature centric or socio-centric purposes. Plieninger et al. (2013) used a land use based spatial tessellation in their analysis and get to the conclusion that finer grained spatial units are required.

2.6.2Place based deliberative technique studies

The second group of studies is characterised by the fact that is focused on specific spatial units (region or landscape) and always involves locals and sometimes also tourists using deliberative methodologies often using participatory mapping techniques or PPGIS. This has been recognised as the most effective way to proper detect the spatial component

[20] Subjective per the authors of the report. We do not think CES are the result of additive subjectivities.
[21] In the range of 300 meters around cultural heritage

of CES. In Europe especially, where almost all landscapes are cultural landscapes, those methodologies strongly rely on cultural landscapes research (Schaich et al. 2010). Brown and Fagerholm (2015) reviewed the characteristics of this kind of projects when PPGIS approaches are used. While the review aims at the evaluation of all ES it is clearly stated how provisional and especially cultural services are the main beneficiary of this approach since they are grounded on personal subjective[22] experiences. Therefore, such studies are the ones that more involve the end users of CES. Several ES have been mapped this way either in a deliberative and non-deliberative way. Participants were more confident in placing points over maps to flag areas of interest. Moreover, participants taking part in surveys can stress the existence of CES that are usually either overlooked when biophysical or economic CES assessment are developed or in areas not considered worth cultural or ecological conservation status (Plieninger et al. 2013)[23]. The spatial analysis techniques used the distribution of points to measure the randomness of their distribution, to identify hotspots and to detect correlations between aspects of CES or between them and some characteristic of the environmental spaces[24].

Sometimes the analysis counted the occurrence of points inside specific spatial tessellations. One of the main limitations of this method may stay in the cohort of people interviewed. When surveys or invitations to engage in the local process are delivered to households most of the respondents fall inside a very specific section of society composed by whites, retired, highly-educated people, this bias is often enhanced by the snowballing effect. Several studies have been developed according to this approach. One of the most widely referred is Plieninger et al. (2013). Developing on this and including the new ontological framework as developed in (Fish et al. 2016b) and briefly reported here in 2.4.2. Fish et al. (2016a) developed an empirical study. Working in an area with well know identity and character they used a procedural mixed method approach using a structured questionnaire survey, mapping, group discussion, textual analysis, and arts-based methods. These to engage with local residents, excluding visitors or people who consume media about the area. They developed a procedure to engage in several ways and using distinct kinds of engagement techniques. In this process the creation or population of maps happened in three points of the process. Their mapping exercise consisted initially in the creation of a heatmap starting from a 50m grid and using a radius of 500m to analyse the density of green dots used. In this work they used green/red dots to represent services/disservices. In this work the first

[22] Subjective than not merely standardised, they can be somehow contrasting. See how Solitude and Sharing/Socializing ranks high and with very similar evaluations in the analysis of surveys run by Fish et al. (2016a p. 336). It all depends on the individual perception of solitude, solitude of a restricted group or solitude of the individual?
Another interesting piece of subjectivity (same work, same page) is the fact that someone list cows as a sight that makes feel connected with nature or even wildlife.

[23] In this study one of the most cited CES was social cohesion that hardly has explicit physical components

[24] Plieninger et al (2013) analysed the correlation with land use

stages elicited an understanding of community perceptions for the whole landscape it developed using PPGIS. The result of the mapping has been used as a tool to inform a group of participants (who individually answered to surveys and placed dots on maps) towards the deliberative phase of the process as already suggested by Kenter (2016a). Fish et al. (2016a) used another technique to engage people, mainly children, in the process to understand the local environment. With the support of an artist, they used GPS tracks to create experimental drawing techniques of people's movements in space. Mapping was involved at another stage when all participants were invited to sum up all kinds of multisensory experiences they had in the landscape (also some artist-led ones here not reported). Schoolchildren were involved in the development of this third level map and in the engagement with arts activities of the previous stage allowing the study to overcome the usual age engagement bias. One of the main characteristic of the place based studies is the tendency to deliver through the engagement of local people. But this often results in the overrepresentation of white, male, retired, (Plieninger et al. 2013, Fish et al. 2016a) and with mid-higher education people (Brown and Fagerholm 2015). One of the main conclusions of Fish et al. (2016a) is that using a mixed approach method a preliminary heatmap can help focusing the research and analysis with qualitative methods towards the areas that emerge as more significant. But the qualitative questionnaire making, combining participants in discussion groups helps making sense of hotspots and coexistence of services and disservices. This mixed procedural method is greatly enhanced by the fact that it combines several ways to grasp the experiences of space by different age groups and different physical and cultural capabilities. Nonetheless it still falls short of engaging out from the usual demographic pool of environmental participatory approaches. In the following section we show how some of the limitations of CES focused research have been mitigated embracing novel approaches.

2.6.3Actions as proxies for preferences Implicitly expressed preferences, interpretive approaches

Participatory mapping and interdisciplinary approaches are at present the two main ways to detect the spatial extent of CES. In relation to participatory mapping, receptors are involved in a process with researchers. An interdisciplinary approach often implies the combination and creation of several sources of information: surveys and extensive analysis of spatial and ontological relations between answers. Those studies relying on the bespoke created surveys are usually confined to well determined areas and are therefore limited in size and contingent. As it has been suggested in the introductory chapter, there are new large and varied ways to produce information related to space and people's preferences. Those sources are usually covering very large spatial areas and can be always accessible. Therefore, there is another way to detect engagement with the environment. The methodology is based on the assumption that actions undertaken by people can be considered as consequences of their preferences because “the true social value of non-marketed ecosystem services depends on the ways that services are used by different stakeholders” (Carpenter et al. 2009 p. 1308) see also Bagstad et al. (2013a) and the Research Box (2009). We are then advocating an innovative approach to revealed preference method to infer value from people's choices. In the new framework we adopted (Church et al. 2014, Fish and Church 2014, Fish et al. 2016b), the mutual interaction between environmental spaces and cultural practices gives rise to CES. Cultural practices are the activities that relate people to the natural world. Accordingly, we assume that cultural attachment to places motivates people’s actions; we can follow the digital traces left by these actions. Cultural practices are enabled by environmental spaces, but they are also shaping them. We have material traces then that can be analysed.

Another approach to detect people attachment for places is through the detection of elements that are created to enjoy the landscape features. Bieling and Plieninger (2012) analysed CES through their material manifestations in landscapes. This work needed extensive fieldwork to record the physical traces of activities for non-material purposes. Among the manifestations, the presence of hiking trails, benches, subsistence gardens (allotments), and memorial plaques showed people’s pleasure, interest in the surrounding environment, and its enjoyment of the symbolic shared value. As a development of this approach the EU funded HERCULES (HERitage in CULtural landscapES) project[25] using an interesting mixture of methods. It uses data from social media (Kuemmerle et al. 2014, Tieskens et al. 2014, Oteros-Rozas et al. 2017) and connects landscape research to cultural and social features in the landscape, it detects then, from one side the proxies of human preferences and the physical manifestations from another. The approach is very valid and grasps the value of citizen contribution but lacks real engagement with people in a more active role. People are just contributing spatial information (Shaw et al. 2017) and HERCULES missed the possibility to use the social component of UGC as a way to elicit connectedness with places. They conclude that interesting works have been started to involve citizens in citizen science like approaches for cultural landscapes. But we see how they lack the engagement that must be achieved through multi-layered deliberative approach, and they miss the full range of information and engagement potential that CGI and citizen science combined can reach for CES research.

[25] http://labs.kh.hercules-landscapes.eu/

Similarly, we assume that people go to specific places, and then, if they like them, share this enjoyment through the Internet, for example, by uploading pictures or tagging them online with friends. The traceable actions evidence not only the willingness to visit, but also the willingness and inspiration to share. Preferences, choices, and actions are proxies for CES. A value "if it is expressed in human action and choice, it is in principle measurable" (Boyd and Banzhaf 2007 p. 624). Some of these actions are the result of a sequence of individual decisions; for instance, the action of going to a place, taking pictures, and sharing them is a sequence of three decisions. Taking advantage of Internet tools makes some actions and decisions easy to undertake. Therefore, for this approach we can mention Raudsepp-Hearne et al. (2010 p. 5246) "We chose proxies that were relevant to the use of ecosystem services rather than their supply or stock, because we were interested in measuring the spatial patterns of current benefits associated with each ecosystem service". One of the most investigated benefits from ecosystems are the ones related to the aesthetic appreciation of them. But aesthetic judgment implies a broader range of benefits. Cooper et al. (2016). stressed how individual aesthetical judgment is based on social and cultural shared developments:

> "The aesthetic judgements of individuals of particular natural occurrences, e.g. a specific local landscape, will be the product of the historical development of general aesthetic sensibilities in society, scientific knowledge, local tradition and the actual features of the landscape, their health and disposition." (Cooper et al. 2016 pt. 221)

Moreover, they stress how this unintentional agreement can lead to more thoughtful judgments once developed with deliberative analysis. This is why the popularity of places for CES research has been attempted through the analysis of photosharing projects. In the reminder of this section we show how a similar approach has already been used in ES-targeted research. In the following chapters we will show how other studies, framed in different disciplinary frameworks performed analysis of people's spatial and cultural preferences using a similar assumption about actions and preferences. According to what is been assumed here Casalegno et al. (2013) attempted the detection of CES at a landscape level through the density of contributors to a photo-sharing website called Panoramio[26]. They analysed the relevance of geolocated pictures shared on one online project to evaluate the aesthetic value of landscapes in Cornwall, UK. Their main aim was to analyse the trade-offs with other ES. Their metric was based on the number of photographers for each 1km² area. Wood et al. (2013) used pictures from Flickr, another photo-sharing website[27], to evaluate visitation in 836 predetermined cultural and touristic sites worldwide. In another use of Flickr, Richards and Friess (2015) and Walden-Schreiner et al. (2018) focused on a small-scale analysis in parks. They targeted the content of shared pictures and developing the spatial analysis further to include the

[26] It will be described in the materials chapter. Panoramio has now been discontinued
[27] Flickr (http://flickr.com) will be described in the materials chapter.

analysis of environments and focal points using machine learning to understand the factors supporting people's preferences. Tenerelli et al. (2016), part of the EU funded project OpenNESS[28] (Operationalisation of Natural Capital and Ecosystem Services), used the number of pictures shared on Panoramio and Flickr to apply a geographical weighted regression to derive CES. Martínez Pastur et al. (2016), used the photos from one project only (Panoramio) and then run several tests over the sets of photos classified according to expert judgment. This mixed approach relied on one UGC project and its supporting community. From a technical point of view some studies focused on the interpretation of the position associated with photos. Yoshimura and Hiura (2017) used geolocated Flickr pictures, viewshed analysis and land aesthetics studies to explore the spatial extent of environmental spaces comparing supply and demand of natural capital but restricting their investigation to protected areas. In a similar manner also Tenerelli et al. (2017) on an alpine landscape used Flickr pictures and to them applied viewshed analysis and then manually analysing the content of pictures used optimised landscape metrics to characterise visual CES. In a similar manner Oteros-Rozas et al. (2017) analysed manually and in several stages a limited selection of pictures from Flickr and Panoramio and then run several tests studying the connection between features in the landscape, landscape character, Flickr and Panoramio and land use. Oteros-Rozas et al. (2017) derived metrics with the collaboration of experts in the ES field but also of local experts, to improve their results in 5 cultural landscapes in Europe.

Van Zanten et al. (2016) used three photosharing platforms[29] to assess aesthetic appreciation of landscapes and studied also how those results relate to landscape features[30]. Derung and Purves (2016) derived aesthetic evaluation on places from tags associated to geolocated photos. They derived a taxonomy of places from user generated content (folksonomy) and suggest their methodology for the detection of CES. Another study with the involvement of Purves (Chesnokova et al. 2017) analysed tags associated to pictures of a crowdsourced project focused on landscapes. They compared the adjectives used in tags to the aesthetic evaluation of the same pictures achieved through a crowdsourced visual survey. Another feature that helps embracing this approach is that the landscape scale studies demonstrated how CES comes often in bundles (Raudsepp-Hearne et al. 2010, Plieninger et al. 2013) therefore as we can detect some of the cultural practices that witness attachment for places we may assume that some others, undocumented are linked to them and bring to the deliberative process. The gradual acceptance in the ES field of photosharing as a source of information mainly for recreation purposes is witnessed by the inclusion of Flickr pictures as a proxy for

[28] http://www.openness-project.eu/
[29] Panoramio, Flickr and Instagram
[30] But the number of pictures they extracted from Flickr especially is incredibly low. They have probably got something wrong with their coding.

recreation in *InVEST* (Integrated Valuation of Ecosystem Services and Tradeoffs) one of the ES tools mentioned in 2.2.1. The InVEST Recreation model has been developed following Wood et al. (2013) and it uses counts of total photo-user-days for each grid cell or polygon[31] (Sieber and Pons 2015). While InVEST has been used for several studies on ES not all of them have used the recreation component of it. An explicit link between landscape research and photosharing is created with the *Rate my View* project[32] (Shaw et al. 2017) an explicitly designed smartphone app to collect views, information on views and perceptions of people and basic information on contributors. While this is a very interesting approach it is confined to a very specific app that has a specific use and may be less popular than the worldwide generic apps and projects we are analysing in this work.

2.7 Conclusion

This chapter helped in describing and understanding the overall field of the Ecosystem Services approach, sketching some limitations and debates. The focus on Cultural Ecosystem Services and their spatial component helped understand the new framework that has been adopted together with the analysis on the best practices to elicit environmental spaces. At the end we also have shown how a preference, choice based approach towards the detection of cultural practices has been already adopted to detect CES. While this is an important opening towards new methodologies we will show throughout the next chapter that more interdisciplinary opportunities have still to be uncovered ad used for CES research.

[31] *"One photo-user-day at a location is one unique photographer who took at least one photo on a specific day. For each cell, the model sums the number of photo-user-days for all days."* http://data.naturalcapitalproject.org/nightly-build/invest-users-guide/html/recreation.html

[32] http://www.ratemyview.co.uk/

3 CGI and digital cultural practices – digital engagement for places

In this chapter we analyse the area of research that can contribute to the detection of the spatial component of Cultural Ecosystem Services. In the previous chapter we adopted the ontological framework developed through the UK NEA FO (Church et al. 2014, Fish and Church 2014, Fish et al. 2016b). We will especially target the detection of spatially referred cultural practices that can be detected analysing digital spatial information. As outlined in the introduction, we intend to take advantage of the latest methodological and technological developments that since the beginning of this millennia defined the rise of data available for research as well as data centric science. In this chapter we first introduce the new digital interactive platforms and practices in section 3.1. Our interest is mainly geared towards the section that deals with volunteered and crowdsourced geographic information as in section 3.2. The main components of crowdsourced geographic information are outlined in section 3.3. In section 3.4 we show the idiosyncratic elements of CGI, while in section 3.5 we analyse the studies that derived qualitative information over places using CGI.

3.1 User generated content, citizen science and crowdsourcing

User Generated Content (UGC) or crowdsourced information are the more general terms used to define the data that we will use. The term crowdsourcing, proposed by Howe (2006) refers to the practice used by some organisations to engage people external to the organisation in executing micro tasks. In organizing this endeavour, the IT technological advances play a very important role. Therefore, some definitions imply that UGC deals only with natively digital data hence readily available for automated data consumption. Those definitions are tailored around technological affordances while our work looks beyond technology to support social-ecological research. Therefore, our approach, while taking advantage of Web 2.0 technology, is not focused only around its affordances. We will be more consistent with the natural, social, and geographical fields which investigate long lasting relationships between nature and settled population. We are therefore also including data that originally was recorded in analogue, non-digital format and using some automated system or digitisation it was subsequently converted into digital information. This is the approach towards biological records that has been in recent years strongly enhanced by the rise of citizen science.

Citizen science is the usual name for initiatives which pursue the active development of science engaging laypeople. Here we have the convergence of several heterogenous initiatives. Public Participation in Scientific Research (PPSR) (Bonney et al. 2009a), Community Based Participatory Research (CBPR) (Silka 2010), volunteer monitoring, community science, and participatory action research are also cited in Bonney et al. (2009a). Cohn (2008) mentions Bonney in several passages, in one of them Bonney states that citizen science projects are a partnership between volunteers and scientists to answer real-world questions. To refer to all those different approaches we will assume as an "umbrella term" citizen science (Bonney et al. 2009b). There is not a clear-cut definition of terms that has been ubiquitously accepted. A recent review cast a light in a quite intricate terminological situation (Eitzel et al. 2017a, 2017b). The same social and technological advances that lead to UGC have favoured the widespread development of citizen science but contrarily to the projects and methods considered by Shaw et al. (2017), Wiggins (2012) mentions technology-supported citizen science to distinguish it from more traditional activities of amateur naturalists. Social interaction and participation is strongly supported using ICT technologies (Hand 2010, Nov et al. 2011). New technologies, more able to interact with the environment, empower local communities (Paulos et al. 2009) allowing them to contribute to the shaping of their places (Paulos 2012). Citizen science projects are not intended uniquely as crowdsourcing endeavours with the hierarchical subordination of volunteers to scientists as described by Silvertown (2009) who limits the role of the citizen scientists to collectors and processors of data as part of a scientific enquiry. Citizen science is not itself a discipline but is a set of tools methodologies and practices that sketches new kinds of relationships between citizens and science. The ten principles of citizen science (ECSA 2015) make really evident how citizen science is part of the overall UGC phenomenon while having a quite considerable history of public engagement that predates digital technologies. In the broader family of UGC, the niche of citizen science is carved through the scientific rigour of methodologies and protocols.

The most popular type of citizen science is devoted to the collection of environmental observations or recordings. A considerable amount of ecology focused citizen science projects is following a paradigm that is based on the collection and classification of environmental data for biological records (Sutherland 1996). It is also called contributory citizen science (Bonney et al. 2009a, Shirk et al. 2012). Some citizen science projects are even focused on a single taxon (a group of populations of organisms that are considered a unit) or local environment. Ecological citizen science tends to create knowledge, detecting, collecting, measuring, and classifying data over individual species, environments, or natural phenomena. All the observed elements are located on the earth's surface and the location of the observation is an unmissable element. Moreover, to make observations somehow verifiable the observer is always recorded, sometimes even a determiner is added to the original record since amateur naturalists are really motivated to contribute in a rigorous way. Therefore, contributory citizen science has a natural, geographic, and a social component. A biological record conveys, space, time, nature, and an observer. The observations equate to an atomic element of a cultural practice.

Citizen scientists with interest in natural sciences use UGC to: create new biological recordings on the field, as a tool to aggregate communities of interest, and to digitise past collections of biological recordings to extend backwords the timeline of recordings. This have been done in several ways, sometimes using crowdsourcing like in the *OldWeatherProject* (https://www.oldweather.org/) to extend backwards climate recordings or like the work of UK's Local Environment Records Centres (LERC) who collect, collate, and manage environmental information from citizen scientists and professionals alike. Another feature that characterise citizen science as a very valid approach towards ES analysis and understanding is that it is not only a data creation activity. Citizen science implies participation, an involvement, and a knowledge creation process that is situated, localised, often social, and restricted in a period. Citizen science efforts and traditions are long held in natural history and nature lovers. This way it overcomes the limited period of engagement of ES designed short term engagement actions as stressed by Kenter et al. (2015) cited in Fish et .al (2016d). While an environmental observation is an atomic cultural practice, also other activities show a relation between spatial, natural, cultural entities and people. In the following sections we focus on digital information designed and standardised around the technological affordances of spatial information in the web, the geoweb. The potential of UGC is impressive, as it continues to expand daily, and it can be used for several purposes. Here, we are focusing on the detection of the spatial component of CES, therefore we will focus on UGC with a spatial information component.

3.2 Crowdsourced and Volunteered Geographic Information

UGC with spatial component and the enabling technology of the geoweb received an increasing interest in recent years. It is part of the phenomena of neogeography (though the term has not been in use any longer) (Turner 2006, Haklay et al. 2008, Wilson and Graham 2013) and there has been a wide variety of characterisations uses and definitions. A full glossary in Capineri et al. (2016) was developed starting from a similar work in See et al. (2016). The complexity and variety of terminologies suggests the flexibility of the underlying combination of technologies and practices and the impact that UGC with spatial content can have on multiple fields. Therefore, the spatial element can be seen as the connecting element between several disciplines. Capineri et al. (2016) gathered one of the most recent collection of UGC with spatial content focused papers. In this case the overall field has been called Crowdsourced Geographic Information. In the following section we will mention some ambiguous use of terms and set our terminology.

3.2.1 Terminological and ontological clarifications

The most commonly used term to denote UGC with spatial component is Volunteered Geographic Information (VGI) following Goodchild (2007a) that defined it as "the widespread engagement of large number of private citizens, often with little in the way of formal qualifications, in the creation of geographic information." (Goodchild 2007a p. 212). The striking resemblance with citizen science is also underlined by Goodchild (2007b) who mentioned citizen science as a precursor of VGI. The wide range of projects that used geographic information from private citizens now includes also information that has been created without deliberate actions or deliberate intention to create (digital) geographic information. Therefore, beyond geography, several disciplines have focused their attention to crowdsourced geographic information and the acronym VGI has been used with different meanings. This section shows some of those conflicting uses of the term, underlining the different perspectives under which they have been proposed. We will also introduce other classification criteria and the terminology that will be used throughout the thesis. To understand the complexity, we can look at the way Harvey (2013) and Fischer (2012) differentiate spatial UGC. They defined VGI and differentiated it from other categories using two different discriminating factors. Harvey (2013) used contributor control over the production of geographic information to distinguish VGI from Contributed Geographic Information, Fisher (2012) used intentionality to create geospatial information for mapping and cartographic purposes as the discriminating factor between VGI and iVGI (inVolunteered Geographic Information). VGI as defined by Fisher is therefore a subset of VGI as defined by Harvey.

To avoid misinterpretations, we will not use VGI to identify UGC with spatial components, but we will use Crowdsourced Geographic Information (CGI). The explicitness of the spatial scope of the projects (Antoniou et al. 2010) will be used to characterise CGI projects. While Antoniou et al. (2010) were referring only to collections of photographs, they proposed a distinction that is particularly useful for every CGI project. Spatially explicit and implicit projects are therefore listed. Spatially explicit projects "urge their contributors to interact directly with spatial features"…"while at the same time encourage that photos, and thus the content, be spatially distributed". (Antoniou et al. 2010 p. 100). Those projects are therefore more spatially oriented while other projects, whose main aim is to connect contributors through a shared social component, lacks spatial focus and the spatial information can be an ancillary attribute of the data collection process. Projects pertaining to the latter category are called spatially implicit. This distinction is important since the precision of geospatial information is quite different between the two broad categories of data collections. Campagna (2016) stresses how spatially explicit and implicit CGI needs to be treated differently since while they can provide different aspects of the engagement for places the spatially implicit CGI could lead to the creation of geospatial information not fit for mapping and cartographic production. The above-mentioned limitations are relevant since the internal project dynamics, the motivations of communities have been proved to influence the behaviour of contributors (Haklay et al. 2010). We are dealing with socio-technical systems (Suh et al. 2009, Ballatore 2014, Haklay 2016b).

3.2.2 Collaboration patterns

We propose here a new classification criterion for CGI projects, but it can apply to all UGC projects. They are characterised by the creation of knowledge through two collaboration distinct patterns. We call them *incremental* and *additional*.

Incremental projects are explicitly collaborative and drive contributors to create and improve one element of the data collection for one identifiable entity. For instance, where the main content of the collection is made by geographic features used for mapping and cartography, a building can be represented by a unique geographic representation as polygon. In this case while one contributor can create the polygon, other contributors can improve it either in geometrical terms or adding more semantic information attached to it. Semantic information can be in the form of free text or tags and it can include a description and several other attributes such as height, designer, structure, roofing materials, use, floors. In a similar way, consulting a crowdsourced encyclopaedia, users do not expect to have several articles referring to a specific historical/artistic building[33]. Users expect to have an article where contributors add information that is in this case mostly semantic information in the form of free text. At a certain point some contributors can also add a geographic reference and some tags to enhance discoverability in digital information systems. The content of incremental projects can include different kind of media such as: text, individual geographic features but also environmental observations or spatial tessellations. After an initial creation of the element the content is then enriched with information from the same or other contributors. We can have an enrichment that can be the enlargement of the minimal content, diversification, correction, reduction. The different enrichments are called revisions and create a new version of the information element. In incremental projects information also includes the history of data processing. This allow to follow the creation and refinement of entities through multiple contributions. The term *lineage* is used to describe the collection of revisions and versions of data. This is a relevant component for digital geographic information (Lanter 1990). As an example, what starts as a textual element can be enriched with pictures and then can be enriched

[33] At least in one language only

with a geographic reference. Therefore, in incremental projects the care of contributors for the digital representation of an element of knowledge is witnessed by the number of people that collaborated to the construction of the information about it. The incremental pattern is typical of geographic databases created for mapping and cartographic purposes. Therefore, when the main focus of the collection is on geographic elements, we have a spatially explicit incremental project. In spatially explicit incremental projects every geographic entity is usually associated with one or a series of tags (Haklay and Weber 2008) that add semantic information to the geometric attributes. When contributors are not constrained/encouraged to work on a unique element we have an additional collaboration pattern (or no collaboration at all).

Websites that collect photos can contain a limitless number of photos of the same entity and contributors are not encouraged to limit their contributions. For the remainder of this work, following Capineri et al. (2016) we will use Crowdsourced Geographic Information (CGI) to refer to all UGC with spatial content projects. We will use Antoniou et al. (2010) dichotomy to distinguish between *spatially implicit* and *explicit* CGI since this distinction has a strong influence on the geographic (mainly geometric) quality of CGI data. We will use the term *collaborative mapping* (CM) to identify all projects where contributors create digital geographic information actively and willingly to create geospatial data for mapping and cartographic purposes. Therefore, in the remainder of this work CM will include digital geographic information produced by contributors that in this context are often called *mappers*.

To characterise projects, we will also distinguish between *incremental* and *additional* contribution patterns as they have been defined above. The lineage of incremental projects will be also mentioned. In Figure 3-1,below, a graphical representation of the relation between the terms and concepts that introduced in this section will be used throughout the thesis. The combination of the two criteria suggested creates four categories.

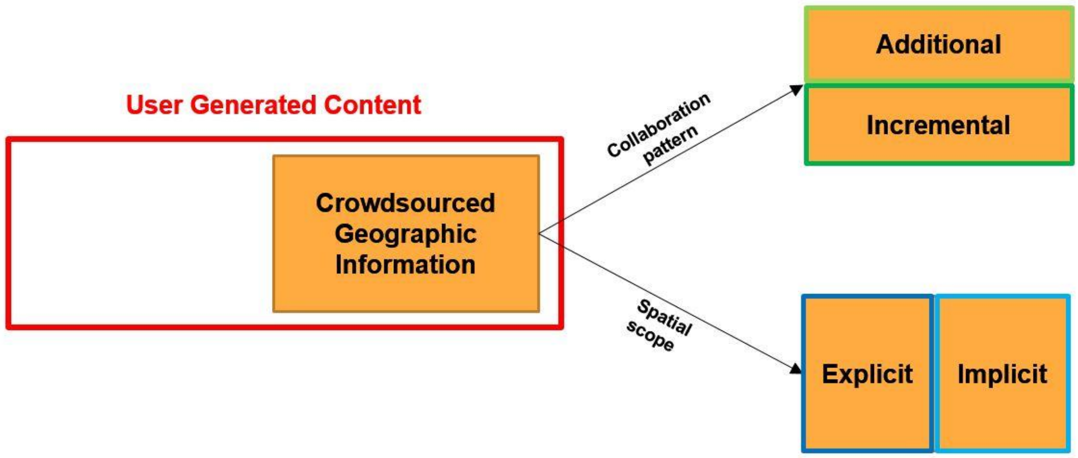

Figure 3-1: User Generated Content, Crowdsourced Geographic Information, and its analytical categories

In the following sections, we focus on the details of the components of CGI.

3.3 Characteristics and components of CGI

CGI combines, space, technology, social practices, and power relations (Capineri 2016). This is perfectly fitting our research since environmental spaces are the localised product of cultural practices which involves the creation of power structures amongst stakeholders. CGI is mainly composed by digital information associated with space and information about the making of the digital information. According to Capineri (2016) from a geographer's perspective the components of CGI are: geographic references (where), the content from which we derive information and knowledge (what), and the additional attributes pertaining to digital information, users, or contributors (who). This always includes the time of production of the information and sometimes some demographic information regarding the users or contributors. In a similar manner Campagna (2016) using a more mathematical notation defines the data model as composed by the geographic coordinates, time, contributor information, multimedia content, and a variable representing the popularity of a specific element. The above listed components and classifications are designed to describe the projects themselves. They miss to include in the projects several elements that contributed to the UGC phenomenon. Therefore, we propose here to add another component to the ones listed above. We call it the community support services, they can be tools or activities, they can be either set-up by the organisation or created by the community. In the following, we will describe some of the characteristics and issues related to the CGI components as proposed by Capineri (2016) adding community support services.

Focusing on the geographic reference, the ability to collect geospatial information can be done following three main approaches: referring the information to geographic coordinates, linking the information to a list of predefined place names of a gazetteer, referring to geographical locations using free text. The geographic references that do not rely on coordinates can have several orders of ambiguities (Amitay et al. 2004, Pasley et al. 2008, Capineri 2016) since geographic names can be used in different locations, they can also be used as names in non-geographic context, some places may have more than one name associated to them, we can also have homonymous geographical place names. Another ambiguity is created when collection of information referred to as place names are digitised and associated to the geographic centre of an aerial feature (centroid). Information that was originally associated to a large area is digitised as belonging to a specific point. Other than ambiguity error also plays its role, in not spatially explicit CGI projects inaccurate, vague descriptions and sometimes incorrect location coordinates can be found. Digital technologies for analysis and visualisation purposes favours when geographic information is recorded through coordinates. In this case, the geographic element is either represented through a point; a line; or a polygon. Points are either used individually as points of interest (POI) or used as building blocks for more complex shapes, then they are called *vertices* or *nodes*[34]. Nodes may be without semantic information since this is associated only to the line or polygon to whom they belong. The geographic reference is the main discriminator between spatially explicit and implicit CGI. Spatially explicit CGI are characterised by the fact that the geographic reference is part of the minimum information collected. In this case coordination and standardisation attempts are enforced from the organisation to ensure geographic data both in geometric and in semantic terms conforms to standardised models. For spatially implicit UGC the vast majority of information collected within the project may not have geographic reference at all.

Focusing on the main content of the collection component. Here the digital information conveys a very wide variety of information. We can have digital media (photographs videos), environmental recordings (often with associated photos), microblogging elements (Twitter, Facebook), commercial transactions, recommendations, and ratings (TripAdvisor, Foursquare, Yelp), commuting data or encyclopaedic articles (Wikipedia). The mentioned Fisher's (2012) VGI is constituted by projects where the geographic information and its classification is the main content. We call these projects collaborative mapping projects. In this case the main content overlaps only partially with the geographic reference since textual tagging schemas are always added to interpret the geographic features correctly. Textual and tagging schemas are called the semantic component of geographic information. The depth of the main content can range from simple free text, to more advanced multi-tiered textual based tagging schemas. Those taxonomies can be co-designed and coordinated by contributors' communities. Such taxonomies are called folksonomies (Vander Wal 2007) merging folk and taxonomies. Stricter taxonomies can be designed by the organisation (Google mapmaker, Foursquare, TripAdvisor) or can be derived from pre-existing scientific classifications of biota (citizen science).

Considering the additional information. This component is composed by ancillary information for the main purpose of the project. It can be of the same nature of the main content category. This component of CGI contains qualitative information that can be subjected to analysis from several disciplines. They can usually include also some information about the contributors, at least a unique identifier is requested to aggregate actions of the same individual. This information is often the only information included in the data about the contributor. Studies of contributors to CGI often relied on additional surveys or in the combination of spatial analysis of contributions and census information.

[34] A connected sequence of nodes constitutes a line, a connected sequence of nodes where the first and the last node are the same is considered an area or polygon

Projects that are designed to record a wide range of additional information provide a very rich heritage that has been widely analysed and several biases has been detected in the participation and engagement of contributors. The biases found on CGI information will be analysed more in details in a following section.

The last CGI component, which we added to previous classifications, includes the community support services. It usually encompasses IT and non-IT services. The IT services include communication facilitation and bespoke software developments. The non-IT part is strongly enabled by the IT tools but includes more traditional social events, gatherings, pub meetings, conferences. Therefore, we are here including a large amount of information on community dynamics, activities, and the supporting communication. In IT terms, communication facilitation such as the use of blogs, mailing lists, wikis, and live chats are often spontaneously created by the communities of contributors using resources external to the main organisation. Subsequently they may be embedded in the overall organisation. The software development, mostly opensource and also crowdsourced, define a sub-community who enables the overall communities with software to facilitate the editing process in terms of speed, interaction, precision and, standardisation of the outcome. Therefore, from a technological perspective there are applications for smartphones, editing software, internet bot designed with the only purpose of enriching the projects. Those elements are also an integral part of the projects and of their success. The communication facilitation supports also the creation and management of events to enhance the sense of community, facilitate participation, and ease productivity. To mention a couple, we can have gatherings to discuss or to share part of the activities of the projects that can be even branded differently in accordance with the main interest of the community of contributors. Therefore, we can have *mapping parties* and *bioblitzes*. While mapping parties are social localised gathering where volunteers work together to collect geospatial data for mapping and cartography, bioblitzes are gatherings of environmental citizen scientists to collect environmental recordings. Such social components have been used to analyse the internal working of communities (Hristova et al. 2013, Perkins 2014, Arsanjani et al. 2015a, Mashhadi et al. 2015). Those communication channels and socially constructed tools - while complementary to the main core of the projects - can be a very rich source of qualitative information on community practices and project development. Due to our spatial focus this thesis will not engage in the analysis of community support services.

3.4 Critiquing CGI

CGI being a subsection of UGC inherits mainly two of the weaknesses of data constructed in this way. Firstly, is questioned the reliability and quality of the information collected. Secondly, when looking at the social aspect of the overall UGC projects they tend to be developed by specific subsections of the populations, therefore CGI projects being affected by participation inequality mostly reflect some specific worldviews. Other kind of biases have been consistently found in several modalities of CGI including social media, citizen science and the collection of geospatial data for mapping and cartography. Those biases are often found in bundles and in the following we will distinguish them in three classes: social, spatial, and temporal biases.

3.4.1Quality of CGI data

The potential of CGI to produce reliable information is the main issue in literature. The main UGC project globally, Wikipedia, was at the centre of a heated exchange of articles between Nature and Encyclopaedia Britannica (Giles 2005, Britannica Encyclopaedia 2006, Liu et al. 2006) regarding the quality of entries. When a pool of experts conducted a peer reviewed analysis of excerpts of science-based material available online from Encyclopaedia Britannica and Wikipedia in 2005, the results shown that "the average science entry in Wikipedia contained around four inaccuracies; Britannica, about three." (Giles 2005 p. 900). This was an unexpectedly good result and witnessed not only the potential of Wikipedia but also one of the main methodology that will be used to control the quality of UGC and then CGI. The comparison with an external reliable source of information has been named extrinsic quality measurement.

The issue of quality of CGI involves also the spatial datasets that describe the physical world as in collaborative mapping projects. Therefore, geometric quality is the main concern usually followed by semantic accuracy. To assess quality and usability of data produced through collaborative mapping, several studies developed a wide variety of methodologies. Senaratne et al. (2016) compiled an interesting review of all quality assessment methods used. In essence, several studies witnessed how collaborative mapping can be locally more detailed and accurate compared to authoritative datasets. Haklay et al. (2010, 2010a) discussed the quality criteria for collaborative mapping identifying already the main trends. While the extrinsic measurement was dominant (Haklay 2010b) there was the first assessment of quality in relation of intrinsic quality measurements (number of contributors per km^2). Haklay et al. (2010) found a local correlation between geometric quality of collaborative mapping and the number of contributors. The framework or authoritative datasets have been used to measure the quality level of collaborative mapping also in other areas (Ludwig et al. 2010, Zielstra and Zipf 2010, Girres and Touya 2010). The growing literature on the use of extrinsic quality measures to assess the quality of geospatial data produced by volunteers drove to the refinement of intrinsic measurement of the quality of CGI data for cartography and mapping (Antoniou 2016, Senaratne et al. 2016). The social component and the degrees of freedom granted by some CGI projects lead to inconsistent tagging this is why several studies have been designed to mitigate inconsistencies using semantic technologies (Lemmens and Deng 2008, Ballatore et al. 2012, Gliozzo 2013). Thematic mapping has also been improved using CGI approaches. The International Institute for Applied Systems Analysis (IIASA) in Vienna (http://www.iiasa.ac.at/) has a group of researchers whose aim is to create citizen science projects to validate land use maps (Fritz et al. 2009, 2012). They published several works assessing the validity of their approach and the quality of the data obtained (See et al. 2013, 2016). Using quality of geodata as the final aim the above-mentioned studies end up stressing the role played by the community (Haklay et al. 2010, Napolitano and Mooney 2011, Mooney and Corcoran 2012a, Arsanjani et al. 2013). While most of the studies on CGI focus on the quality of geospatial data for mapping cartography and transportation networks, comparing them with authoritative geographic information, some studies developed methodologies to verify the data quality in terms of individual points data as Mashhadi et al. (2015).

Also citizen science has its quality issues. They have been hindering for long time its ability to be accepted as a reliable source of data for biological recordings. Bonter and Copper (2012) for project *Feederwatch* and Gardiner et al. (2012) comparing the results of validated and not validated citizen science projects focusing on ladybugs/ladybeetles found out that volunteers tend to overestimate rare and unusual species. Similarly, Boakes et al. (2016) analysing traits of observed species in Greater London reported that volunteers tend to record species that are easier to identify[35], alongside common and familiar species. Trust between contributors plays an important role to produce qualitatively good geospatial information (Bishr and Mantelas 2008, Mooney et al. 2010b, Keßler et al. 2011, Keßler and de Groot 2013). CGI is a social constructed endeavour; therefore, social considerations goes beyond the mere assessment of data quality. Haklay (2016b) described the various ways in which CGI is prone to bias in participation and their implications. The sum of them can also be mentioned as a sign of an underlying *Digital divide* (Arsanjani et al. 2015a) that unveils itself mainly in spatial, social and gender biases. There are several consistent biases that have been underlined in various studies (Nielsen 2006, Neis and Zipf 2012, Mashhadi et al. 2013, Graham and De Sabbata 2015, Quattrone et al. 2015, Sen et al. 2015). What is usually seen as a bias in the systematic collection of data using planned surveys such as in citizen science, in our case is seen as a pattern that bears some significance. In the following we characterise more some of the biases and social dynamics that characterise CGI as well as more in general UGC.

[35] Maybe in doubt they do not submit so more evident traits are more likely to support an identification.

3.4.2CGI Participation inequality: By few, for many

To describe the community of users and producers of CGI Capineri (2016) uses the term *produsers.* While this term is seen widely applicable to all web 2.0 projects, specific analysis of the patterns of contribution reveal that the mass production of the content is not as widespread as expected. As already signalled in the first chapter while the focus is on the crowd and very often the numbers of participants are used as indicators of the success of a project, several studies in UGC and CGI (Nielsen 2006, Lin 2015, Haklay 2016b) demonstrated that the vast majority of data (usually 90%) is provided by a narrow minority of contributors (usually 10% or even 1%). Haklay (2016b) warns that while those highly engaged contributors are collecting a lot of data they do not constitute the average contributor, they are statistical outliers. The highly committed statistical outliers are greatly influencing the resulting data collection in several ways. But this does not guarantee that their contribute is qualitatively better since low-volume contributors may offer deep and direct place specific knowledge while high-volume contributors are offering more fleeting observations for the areas that they cannot experience personally (Elwood et al. 2013).

Haklay (2016b) estimated that the main contributors to CGI spend around 37 hours a week contributing. Describing the potential impact of those collaboration patterns he also recommends studying the behaviour of such highly active contributors to analyse their influence on the shaping of the community. Accordingly, Glasze and Perkins (2015) call such ultra-dedicated contributors "a small and elite group" that steer the projects effectively determining their impetus and directions. Neis et al. (2012) to analyse the engagement of contributors to a CM project (OSM), analysed the dynamics of data creation as of 2011 detailing participation temporal and spatial biases. They found out the usual long tail distribution replicates at several scales. Firstly 62% of registered contributors provides no contribute. Neis et al. (2012) created groups of contributors according to their commitment. *Juniors Mappers* as the ones that have provided at least 10 nodes. Those are around 14% of registered contributors. While the *Senior Mappers* are the 5% of contributors that provided individually more than 1,000 nodes. The rest of contributors are the *Nonrecurring Mappers*. The long tail is replicated even when the analysis moved to focus to working sessions. Here 89% of work is performed by the 5% of registered contributors (the Senior Mappers) while 9% of sessions have been produced by 14% of registered contributors and 2% of sessions has been produced by 19% of registered contributors.

Purves and Mackaness (2016) highlighted the strong influence played by certain behaviours on analytical result of text analysis. The automated contributions of large amounts of data (bulk uploads) and the strong influence exerted by particularly assertive contributors can distort the analytical results. To stress the role of the non-IT community support services of CGI, Mashhaidi et al. (2015) also measured the impact of mapping parties on the activity of contributors. According to the mentioned paper mapping parties are successful in engaging and retaining contributors. They conclude that newcomers and low contributors are more productive and can be more engaged while heavy contributors are not that affected by mapping parties. Also in citizen science there is the same contribution pattern. Boakes et al. (2016) analysed citizen scientists' behaviour in Greater London comparing different projects both natively digital and digitisation of older records they found out the usual long tailed participation bias. Boakes et al. (2016) developed engagement profiles based on frequency and duration of engagement and found out that most of the volunteers were active for just one day and provided few observations each. The amount of people that just register and never contribute or that contribute with one item and them never come back is also considerable. This is the retention problem. It has been detected in every form of UGC including contributory citizen science. Wiggins (2012) describing the Great Sunflower Project reported that retention plays an even more important role than recruitment. Being aware of those different contributing patterns can help tailor the research methodologies and estimate their impact on the use of CGI data. In the following we will show the main biases usually found in CGI data.

3.4.3Social and gender biases

When combining CGI with other space referred information sources, several studies analysed contribution at several levels of spatial granularity. To infer social biases the analysis of data is not strictly needed. Specific features related to the engagement timings help to characterise the population sample, whose spontaneous aggregation is not completely accidental. Following on Haklay's findings (2016b) about the time spent participating in the creation of CGI some specific considerations arise. Different projects need different resources, sometimes not only time. Therefore, well paid jobs that allows plenty of spare time, but also financial resources are necessary for the highly committed participants. See et al. (2016) in their review stressed how the registration of contributors is most of the time lacking in data that goes beyond a username and an email address. Hence the studies that tried to derive the profile of CGI contributors analysed their spatial activities and footprints, studying their interactions, and mixing them with demographics of settled population (Neis and Zipf 2012, Arsanjani et al. 2013, Keßler and de Groot 2013, Quattrone et al. 2014a, Glasze and Perkins 2015, Mashhadi et al. 2015, Stein et al. 2015). The connection between CGI and census data enabled researches to derive social profiles of participants. But this approach is inconsistent with the fact that dedicated contributors to CGI are few individuals even inside an online community. This approach can be misleading given the very low numbers of contributors to CGI and given their spatial spread[36]. Coherently with this Steiger et al. (2016) strongly warn against the belief that analysing social media generated data will give an image of a real settled population. They also stressed how the heterogeneity of social media CGI data makes very likely that the information will address for several real-world phenomena. Neis and Zipf (2012) derived the distribution of contributors to OpenStreetMap (OSM) the most popular collaborative mapping project at national and local levels. Several times socio demographic variables have been connected to CGI practices finding significant correlations (Graham et al. 2013, Hristova et al. 2013, Arsanjani and Bakillah 2014, Quattrone et al. 2014a, 2015, Yasseri et al. 2014, Glasze and Perkins 2015, Mashhadi et al. 2015). More specifically Mashhaidi et al. (2015) investigated social biases in OSM and related the localised activity with social and population statistics at ward level in Greater London. They found out that population density and closeness to urban centralities have a strong correlation with completeness while a weaker link is with poverty. This stresses the digital divide concept that adapted to CGI is also labelled digital deprivation in Mashhaidi et al. (2015). Arsanjani and Bakillah (2014) using a logistic regression model found very similar results in a federal state of Germany linking 15 social economic and physical characteristics with areas of high contribution. Such areas have been identified aggregating the information of number of points and edits over a grid of 1 km^2 and applying then spatial statistics to derive spatial clusters. As for any social endeavours there is place for conflict along national and cultural divides epitomised by edit wars alongside pure vandalism (Ballatore 2014, Bittner 2017).

Cultural conflicts and disagreements are very useful for studies that tend to include alternative perspectives such as the ones hold by different stakeholders. But the process of cartographic production and mapping is in itself a social constructed process. As an example, Glasze and Perkins (2015) reports on the under representation of Mosques in German authoritative geographic information versus the more open and updated representation of the territory by volunteer mappers. Social biases are also detected in citizen science Phillips et al. (2012) underline one characteristic that all citizen science projects have in common “the fact that the majority of people who participate in citizen science in informal settings are self-selected” (p.85). The gender bias is also stressed by some studies. The usual assumed gender bias favouring males over females on UGC has also been detected in CGI (Schmidt and Klettner 2013, Schmidt et al. 2013, Steinmann et al. 2013, Stephens 2013) this results in the unbalanced detailing of gender specific information. Steinmann et al. (2013) while using a small sample made an interesting comparison between the gender biases of several UGC and CGI projects that while confirmed the general trend towards masculine majorities included also some

[36] We may even think that the intensive participation in online activities is a consequence or a contributing factor to local spatial isolation between the contributor and the surrounding settled community.

cases with feminine majorities. As Glasze and Perkins (2015) noted the creation of digital geographic information is a socio-technical practice, hence it reproduces societal structures keeping in the process also inequalities and biases as well as preferences of the map data producers. Lambio and Lakes (2017) demonstrated how social bias is rooted in the character of Flickr that represents a west centred worldview.

3.4.4Spatial biases

Hecht and Stevens (2014) found out that urban inhabitants tend to use more social media. The analysis of CGI helped defining new real life or 'natural cities' (Jiang 2016) making city dynamics more transparent and redefining the meaning of urban limits. The high dependence on very active contributors creates spatial and temporal patterns in the available data. Socio-spatial biases are derived when location of social posts (either photos or microblogging or ratings on dedicated websites) and time of contribution are combined to identify tourists and their most photographed locations (Girardin et al. 2008, Kisilevich et al. 2010b, 2010a, Wood et al. 2013, Floris and Campagna 2014). This kind of analysis can enhance greatly tourism tracking systems (Shoval and Isaacson 2007). Some studies stressed the difference between preferences of locals and tourists comparing specific targeted social media (Capineri 2016). Therefore, CGI is nicely fitting the need to have a tool that allows to identify spatial unities going beyond conventional borders (Capineri 2016, Jiang 2016). Hecht et al. (Hecht et al. 2007, Hecht and Gergle 2010, Hecht and Stephens 2014, Sen et al. 2015) detected and analysed the locality of patterns of several sources of CGI. Campagna (2016) addresses this characteristic stressing how the local grounded preferences for specific CGI makes important to analyse several of them alongside authoritative datasets to better improve the representativeness of CGI data. This aspect is certainly important when considering the care that should be placed in stakeholder selection to elicit CES (Fish et al. 2016d). While CGI is a very rich source of information it may lead to several statistical errors.

Neis and Zipf (2012) used several manipulations on OSM lineage to derive relevant spatial biases at national and local level. To determine spatial biases in terms of the country of origin of contributors, since this is information not directly available through OSM, it has been speculatively deducted by the contributor behaviour. They created a system to assign the location of a contributor to a specific area. As a threshold for the tentative assignment of locality they used a 1 km length of triangular shapes. They found out how deeply unequally distributed is the home location of the majority of contributors. At a global scale as for December 2011 three quarters of the contributors came from Europe and there was no relation with population or land surface. More patterns were computed that demonstrate how the most active or senior contributors have also a wider area of activity, such area decreases with the less committed contributors. But the percentage of very active contributors that have a quite localised focus is still significant since almost 37% of the most active contributors have a calculated range of activity that is less than 10 $km^{2.}$ This is important since locality and local knowledge are really relevant for map production but also for our research. At the same time the same group of very committed contributors has a large percentage of members mapping in an aerial even larger than 50km^2 (25%) while the less dedicated contributors in their greatest majority map in a neighbourhood with 96% of the less committed mappers covering areas of less than 5km^2. This is why CGI is very relevant at local level where together with dedicated wide-ranging mappers there are the local ones, but it also exerts the assessment that very active members are influencing and connecting the work of all the other contributors with their spatially wide-ranging contribution. In incremental projects, very active contributors while shaping the community can prevent other volunteers to join the project giving the impression that the collection is completed (Schmidt et al. 2013). Arsanjani et al. (2015a) describe how the production of geographic information from volunteers contributing to OSM is biased at national level in a global scale. The distribution shown in Figure 3-2, below, resembles quite closely the distribution of editors in Wikipedia as witnessed by Graham et al. (2014). But it also shows lessening biases towards Europe.

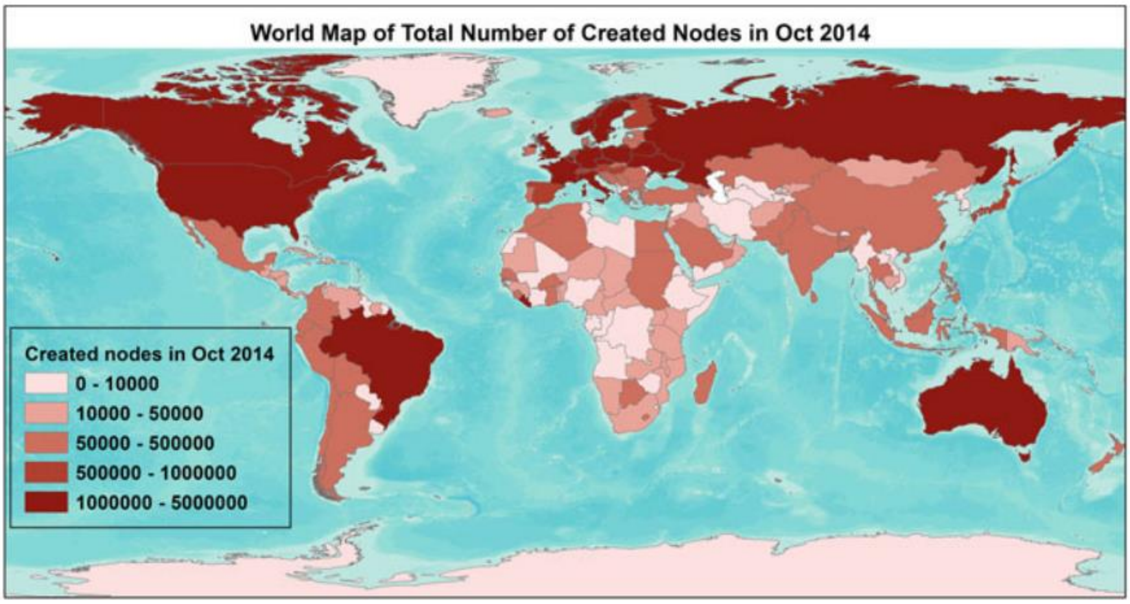

Figure 3-2:A world map of the total created nodes in October 2014(Arsanjani et al. 2015a p. 7)

More strikingly when the number of active contributors is averaged against the resident population the strong biases emerge even more as in the following Figure 3-3, where we have the relation between mappers and resident population of each country.

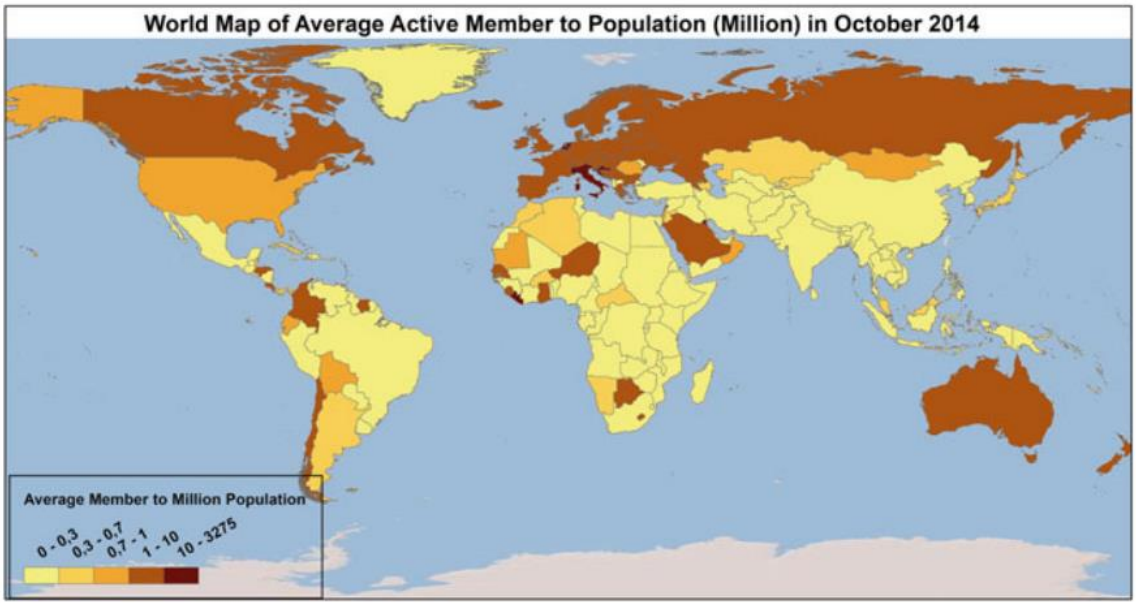

Figure 3-3:A world map of average number of active members to population (million) in October (Arsanjani et al. 2015a p. 8)

Very often spatial patterns are also social ethnic and cultural patterns. Bittner (2017) compared the use of two different CM projects in Jerusalem and found out that they are inversely successful around the Jewish/Arab spatial and cultural divide. At some extent CGI reflects the cultural biases of the society within it progresses. Also citizen science is hugely affected by spatial biases, often found in connection with temporal biases. They can be found in environmental recording efforts (Tulloch and Szabo 2012, Tulloch et al. 2012). Spatial bias tending to favour more wide lakes and easily accessible areas are also contributing to biases in citizen science collected datasets (Surmacki 2005) because volunteers like to detect as more species as possible and they prefer to gather information in proximity of specific features like water courses. Similar spatial biases emerged also when environmental records have been examined in urban settings (Boakes et al. 2016).

3.4.5Temporal biases

Campagna (2016) suggests how the timely production of CGI can lead to a better planning but even more a better timely and promptly responding management of the cities (smart cities). He also proposes the implementation of more explicit and tailor designed CGI systems intended to engage citizens in the management of urban spaces. Barve (2014) compared the timely provision of photos collected on Flickr and the environmental records provided to the Global Biodiversity Information Facility (GBIF)[37]. He found out that they are both affected by temporal biases. While Flickr had a hype in the first years of the century and then slowly declined, it shows less variability along the

[37] http://www.gbif.org/. GBIF includes environmental observations from citizen scientists, researchers, and automated monitoring programmes

year compared to GBIF. While GBIF records are more constant along the years but are affected by seasonal biases. More specifically focusing on citizen science the “weekend bias” (Sparks et al. 2008, Courter et al. 2012) has been often detected. It is because the average volunteer has more spare time to make observations in the weekend.

Neis and Zipf (2012) analysed several dimensions of temporal biases, in terms of days of the week and hours of the day of general contributions to a CM project. Biases have been detected in terms of weekdays where the activity is more acute on Sunday and then has a minimum around Thursday and Friday to increase back towards the end of the week. Hourly biases are starker with mostly evenings and in general daily hours (10-21) more used to contribute. Moreover, they analysed the temporal patterns of volunteers according to their classification amongst *Senior* (highly committed), *Junior*, and *Nonrecurring* mappers. As expected the life span of contribution is longer for senior contributors that so often coordinate and influence newcomers. Sporadic mappers tend to map in shorter time ranges. The contingent nature of CGI has also to be considered. Some projects are enduring and lasting, but several initiatives expire, change their terms of access, or are suppressed. This issue is called project continuation and sustainability in Haklay et al. (2014). Amongst the platform used by Antoniou et al. (2010) half of them are no more in service, as well as the only platform used in Martínez Pastur et al. (2016) and in Casalegno et al. (2013). See et al. (2016) reviewed a list of projects analysed five years before by the spatial group of the University of California in Santa Barbara[38] and found out that half of them were no more in operation. This volatility may affect the way CGI should be considered thoroughly. Given the above-mentioned limitations we can now approach the analysis of CGI projects keeping in mind the warning for the use of CGI Haklay (2016b) stated that:

> "When using and analysing crowdsourced information, consider the implications of participation inequality on the data and take them into account in the analysis."

In the following sections we focus on the qualitative value that can be achieved analysing CGI projects.

3.5 Qualitative value in CGI as CES components

The relation between digital information space and people has been analysed and seen under different perspectives. We are then taking advantage of the several depths that geographic knowledge reached using web 2.0 and big data approaches, technologies, and practices to elicit qualitative information rather than only quantitative one. In this section we show how several components of CES, and some specific practices, goods, and benefits have been widely investigated using CGI. We will refer them to the

[38] Elwood, S.; Goodchild, M.; Sui, D.Z. Vgi-Net. Available online: http://vgi.spatial.ucsb.edu/. Accessed by See et al. on 6 December 2013 it is no more accessible online as of 17/06/2017

ontological framework developed by Fish et al. (2016b) and visually expressed by Fig (2-5). In the following subsections we show how those non-material benefits have been elicited. We have seen that cultural practices enable cultural benefits such as identities that are usually locally grounded and contribute to the sense of place. CGI can complement or help to improve traditional ways to create geographic information but its real value is in giving voice to localities "the most important value of VGI may lie in what it can tell about local activities in various geographic locations that go unnoticed by the world's media, and about life at a local level."(Goodchild 2007a p. 220). CGI constitutes then an unprecedented opportunity at a so large scale but also at so fine-grained temporal and spatial scales. Campagna (2016) from a planning perspective stresses how CGI is still mainly overlooked for studies aiming at planning while it can "capture citizens' interests, intentions, perceptions, and needs" (Campagna 2016 p. 46).

CGI can then support policy making but its localness targets local, social, and territorial structures as well as local identities. Elwood et al. (2013) stresses how the social aspects included in CGI can lead to the development of a more place centric perspective on GIS suggesting the development of *platial* geographic information system. They argue that since CGI is created by humans it more resembles human discourse. As such it refers more to places than to coordinates and precisely delimited areas as in traditional GIS. Geographers are particularly sensible for this kind of analysis. Consequently, computer science educated researchers focus on data quality while geographers focus more likely on the social and place related added value that CGI can deliver. Accordingly, Capineri (2016) stresses how CGI gives an unprecedented possibility to locate in space and time practices that are a consequence of sense of place and belonging. The production of geospatial information is also equated to a form of appropriation of space. Spatial analysis can combine several CGI elements to elicit "sticky places" because CGI knowledge production is able to detect multiple identities and perceptions as well as inclusion and exclusion. There are already several studies that taking advantage of CGI target sense of place and belonging (Purves et al. 2011, Adams and McKenzie 2013, Haywood 2014, Purves and Derungs 2015, Jenkins et al. 2016).

Inspiration from landscape and scenic locations have also been widely analysed mainly using photo sharing platforms (Comber et al. 2015, Dunkel 2015, Seresinhe et al. 2017). As Dunkel (2015 p. 177) stated for photos "the action of taking a photo is recognized as not only being triggered by the immediate environment but by all aspects of cognition: personal preferences, memories, opinions, and more." The analysis of citizen scientists' motivations helps us to understand that this cultural practice enables several immaterial benefits. Contributory citizen science makes use of the appreciation of nature and tends to spread knowledge, awareness and possibly action. Citizen science, can contribute to create spatial/environmental knowledge that can not only drive towards individual environmental stewardship to support conservation and biodiversity but also contribute to create a better picture of places people appreciate for their natural characteristics. Volunteering in environmental recording is motivated by non-material, cultural gaining, and enjoyment of nature as in Hobbs and White (2012) "motivations for participation identified by both organisations and participants themselves centred on learning about their local area or species using their garden, enjoyment of the activities and making a contribution to conservation." Individual and local knowledge has been networked to achieve broader scientific coverage and understanding. The cultural benefits that can be intended as experiences, identities, and capabilities. The social and altruistic components play a very important role in the implementation and success of citizen science projects (Jackson et al. 2015), but also learning and emotional attachment play a role in it (Geoghegan et al. 2016). Citizen science has seldom played a role in social-ecological systems or ecosystem services research. Shaw et al. (2017) explore the connection between citizen science and landscape research. After an interesting overview on citizen science methods and classifications they firstly restrict their interest to projects that employ web-based infrastructures. In doing so they then show what we simply call CGI projects since they collect geographic information about material features of the landscape or alongside with aesthetic judgments and some information about the contributors.

3.5.1Range of analytical techniques on CGI

This wide range of potential targets of qualitative analysis has been reached with a wide range of analytical methodologies. Following Campagna (2016), who focuses on a social media data centric approach, we have five main categories of analyses: Spatial analysis, temporal analysis, spatial statistics of contributor preferences, multimedia content analysis, contributor behavioural analysis. To enrich this list, we can mention Purves and Mackaness (2016) that referring to geolocated photo collections list several ways to extract qualitative information from them. Given that all the analysis reported are not analysing explicitly the photo media used they can be applied to almost all other CGI. Some methodologies tend to characterise places as with the extraction of vernacular place names, the elicitation and extraction of place's semantics, the extraction of locational information from text, the analysis of people's movements and the elicitation of the occurrence of an event. Purves and Mackaness (2016) stressed also the importance of spatial scale and the MAUP issue willing to aggregate point patterns before running spatial analysis on CGI data. Moreover, using spatial clustering techniques CGI has been able to experience places capturing social and cultural practices and processes (Capineri 2016).

In the following sections we will offer an overview of methods and results achieved using digital spatial information generated by people while in or about the open spaces. We aggregated the analysis of those studies in three groups. The first focus on photo collections with associated position. It is the most commonly used family of CGI to elicit cultural benefits and our methodology mainly is based on such studies. The second section focuses on the cases when citizen science has been used as a proxy for CES. The third section is devoted to the only incremental sources we are going to use, this group contains the prominent UGC and CM projects.

3.5.2Photo collections to characterise places

CGI photographic collections are characterised by a digital photo as content that has been referred using a geographic reference system. Photographic collections are in general loosely coordinated and the contribution pattern is in the vast majority additional. Those resources are not always spatially explicit, some of them are mainly photo collections and the possibility to refer photos to geographic places is an additional option, hence those are spatially implicit. Projects whose aim is to collect photos about landscapes are spatially explicit. Crowdsourced photographic collections are diverse and heterogeneous because the availability of *Application Programming Interfaces* (APIs) enables the easy downloading of metadata about photos, as well as the photos themselves. More interestingly, APIs allow the downloading of registration nicknames of contributors, comments, and tags that, together with the time and place associated with the photos, have allowed some studies to focus on the spatial behaviour and preferences of contributors. Therefore, photo-sharing websites are a suitable source of information for mapping people's preferences among landscapes and natural environments.

Several studies have used photo-sharing websites with different purposes. Nov et al. (2010) analysed the dynamics of photo-sharing communities. They underlined the fact that a sense of belonging to a community, supported by the social-networking functionalities of every application, boosted participants' performances in terms of contribution. Antoniou et al. (2010) aimed to evaluate the role of photo-sharing websites as a source of geographical information. An interesting discovery of their study was: "Spatially speaking, the users are not interested in the small, relatively unpopular, niches of space but focus on the mainstream places" (Antoniou et al. 2010 p. 108). Following this definition, we assume that the most popular areas are those where more people take photos, subsequently sharing them online. Contributors behave selectively and tend to share information regarding specific locations. We assume that this is an expression of their preferences among places. Antoniou et al.'s (2010) evaluation of photos against a hypothetical isotropic (spatially even) distribution demonstrated that the real behaviour of contributors was characterised by a strongly unequal distribution in which the density and clusters of photos and contributors revealed people's preferences for places. Girardin et al. (2008) also detected similar behaviours. They analysed the patterns of geolocated photos in close proximity to tourist attractions in the province of Florence, Italy. They also underlined the existence of hotspots, and then detected the digital footprints of tourists as they travelled in Italy. Similarly, Kisilevich et al. (2010a) conducted their analysis assuming the qualitative derivation of spatial preferences from the quantitative spatial density of photos shared online. They examined the attractiveness of places by comparing two photo-sharing websites and an online recommendation system for tourism. Kisilevich et al. (2010b) applied the analysis of spatial patterns from photo-sharing websites for only urban environments. It is also an example of the integration of diverse sources of information. Similarly, Kachkaev and Wood (2013, 2014) applied their methodology to an urban focused project. They created a leisure pedestrian navigation system eliciting people's preferences from a blend of four online photo-sharing services[39]. Their study, alongside those of Li et al. (2013) and Antoniou et al. (2010), deduced that similar Web 2.0 applications implied different target populations and hence different behaviours. Urban centrality was the target of another qualitative evaluation based on the density of photos shared online (Hollenstein and Purves 2010). The reliability of photo collection positioning recording has been also questioned. Senaratne et al. (2013) evaluated the correctness of geolocation of photos. The main issue is related to the fact that a point can represent either the position of the camera/contributor or the position of the photographed object or landscape of interest. Purves and Mackaness (2016) argue that by time the majority of photos are not geotagged by the contributor at home, but they are natively geotagged by smartphones or cameras, therefore they will record the position of the instrument. Therefore, using the positions associated with photos, a specific radius of analysis should be set to weigh also this characteristic.

Seresinhe et al. (2015, 2017) made two studies analysing both the content and the position of crowdsourced photos and combining the positions and ratings of photos with a vast cohort of datasets from census to environmental to the combination of photosharing and VGI data. Therefore, those studies have a relevant potential since they try to harness CGI to evaluate sceniqueness and its relationship with human wellbeing. They firstly (Seresinhe et al. 2015) analysed the relation between colours in the photos and sceniqueness, then they analysed the relation between sceniqueness and wellbeing aggregating and comparing census, environmental data, and surveys on people wellbeing. The second study (2017) implied more CGI sources exploring other ways to

[39] Panoramio, Flickr, Picasa and Geograph.

get value also from tags through semantic query expansion. They also implemented a regression analysis.

3.5.3Citizen science upside down- the mapping of recorders' spatial preferences – the CGI cultural practice

Citizen science collected environmental observations are used to create very rich biological recordings. This makes the geographic reference part of the main content but not the main focus of the collectors. Citizen science biological recordings are spatially explicit. They have a wide range of possibilities in terms of coordination and the contribution pattern is in the vast majority additional. There is a very limited amount of studies that focused on the use of citizen science for CES research. Powney and Isaac (2015) describing the uses of biological records ignore this potential. In Canada Raudsepp-Hearne et al. (2010) used citizen science to get people interest for places:

> "observations of rare and endangered species by amateur nature enthusiasts were used as a proxy for nature appreciation rather than general biodiversity indices, because we were interested in where biodiversity was being appreciated." (Raudsepp-Hearne et al. 2010 p. 5246).

Similarly, Queiroz et al. (2015b) in Sweden did not use generic environmental observations but the ones provided by citizen scientists and collected in the Swedish *Artportalen*. Their approach is evidently very similar to ours:

> "The Swedish Artportalen is an open database where any individual can report a species observation. This means that many observations are reported by amateurs, i.e. not following a standard or systematic sampling scheme. Therefore, the number of observations reported within a certain geographical area is not necessarily representative of the actual diversity of species of the area, rather reflecting the level of interest that people living or using the area have on biodiversity."(Queiroz et al. 2015a p. 10).

They also used the number of observations of rare species to measure the appreciation for biodiversity. But focusing on the number of observations rather than the number of volunteers makes their methodology vulnerable to detect over-active volunteers and their preferred places rather than places of shared interest. Environmental observations by amateurs have a long-lasting tradition and they often focus on urban areas and most notably on private gardens. While the importance of private gardens has been underlined in several reports as well as several studies aiming at CES (Camps-Calvet et al. 2015, Calvet-Mir et al. 2016, Fish et al. 2016a) the swathe of environmental observations there collected has not been used to witness for it. Boakes et al. (2016) explored volunteer behaviour and found out patterns and biases that witnesses spatial temporal and taxonomic preferences as well as biases induced by social organisational patterns. Societies and social/environmental organisations strongly influenced the spatial concentration of volunteers regardless of intrinsic natural values that can be found in open areas. Moreover, familiarity more than rarity, so a sense of belonging was more influencing citizen scientist's behaviour. Barve (2014) made an interesting blending of CGI, photosharing and citizen science. His main purpose was to improve citizen science surveys. He used photos shared on Flickr to derive species distribution. He crowdsourced the species identification adding just the tags used in the photos posted on Flickr. He crowdsourced CGI, he analysed the different temporal patterns of contribution from Flickr and GBIF and identified some elements that influence the ability of volunteers to correctly identify species from photos. He also found out that on Flickr people tend to share more unusual sightings and rare species compared to GBIF. Another potential use of citizen science for CES research is to create specific projects to gather people's perceptions (Fish and Saratsi 2015, Costanza et al. 2017, Priess and Kopperoinen 2017). Since citizen science " projects have a huge potential to increase our knowledge base and contribute to improving decisions and management" (Priess and Kopperoinen 2017 p. 224), moreover "citizen science, and decentralized data collection (involving citizens) would allow models to reflect real-time changes in stocks and flow" (Costanza et al. 2017 p. 14). But in the overall vision proposed citizen science is seen as a tool to introduce participatory and deliberative approaches and trigger bespoke data collections protocols rather than taking advantage of data already collected with other purposes.

3.5.4Incremental spatial knowledge

In this section, we analyse projects whose communities' contribution pattern is incremental. Those projects can be either spatially explicit or implicit. In general terms the incremental contributor pattern witnesses the construction and refinement of knowledge. When the target of communities is geographic knowledge and representation we have spatially explicit projects. Those are the CGI projects for mapping and cartographic purposes which we call Collaborative Mapping projects. When the target of communities is the creation of generic knowledge collections we have a spatially implicit project. The vast majority of elements pertaining to spatially implicit projects are not related to a specific location on earth's surface. For geographic implicit projects the geographic reference is ancillary and is usually a point. In the following, when dealing with spatially implicit projects, we will analyse only the studies aimed at defining the popularity of entities that are spatially related. Incremental projects have been analysed to elicit connectedness with local features and places, but we have found no evidence of the use of the collaboration patterns as proxies for CES.

Several studies have been performed analysing spatial and temporal biases. Spatial biases are already a sign of contributors' familiarity with areas. A peculiar characteristic of this group of CGI projects is that, being in their nature incremental, they promote collaboration. Collaboration is a key to our evaluation since it implies shared interest. To elicit attachment for places through the production of collaborative mapping mainly two methodologies have been implemented. The first methodology simply assumes that quality of CGI information is a consequence of numbers of entities. The second is more focused on collaboration and takes advantage of the lineage of data. In both methodologies the most commonly used approach is to analyse the distribution of minimal elements. They are POI and nodes, when we deal with CM, while they are words when we deal with spatially implicit projects. The number of nodes used to define an area is the most commonly used quality measure applied to collaborative mapping. Therefore, an area with more detailed information is an area that arose more interest. Studies aiming at the analysis of the validity of CGI to extract land use information examined which kind of land use is more accurately mapped. Arsanjani et al. (2015b) found out that there are local variations but in general some categories of land use can be used to create reliable land use maps[40]. Amongst the other categories wetlands, forests, and water bodies are well mapped. Similarly, focusing on natural features Mooney et al. (2010a) used a formula to compute the density of nodes used to define an area. They found out that not every natural feature attracts the same level of attention, some features were over represented while others under-represented.

Golf courses and urban greens, in particular, received more detailed mapping. Estima and Painho (2015) willing to evaluate CGI as a reliable source of land use data focused on the analysis of POI using a minimum spatial unit of 25ha. They just compared collaborative mapping data and authoritative data regardless of collaborations of contributors. Estima and Painho (2015) found that only some specific land types can be reliably extracted from OSM. But their methodology already excluded local focused studies with a so large minimal spatial unit. In Wikipedia, the largest UGC resource which is incremental and spatially implicit, quality has been associated with number of words (Blumenstock 2008). But the most commonly intrinsic used metrics for the quality of articles are the ones stressing collaboration. Therefore, the studies that are focused on the lineage of data. To analyse the collaboration between contributors a common approach mediated from social network analysis (SNA) tends to unravel the co-editing patterns from the lineage. This method originated to analyse wikipages (Wilkinson and Huberman 2007) has been used consistently also in CGI research (Keßler et al. 2011, Mooney and Corcoran 2012a, 2014, Mooney et al. 2013, Stein et al. 2015). Those methods are focused on unveiling the depth of editing collaborations, some of them are based on the concept of *interlocking* (Stein and Blaschke 2010). The interlocking co-editing interaction takes into consideration the sequence of edits to the same elements

[40] One of the most used approach to maps ES consists in the use of land use maps. The use of social constructed land use maps can be also an interesting approach to evaluate ES.

by different or recurring contributors. This approach has been also used to analyse CM projects. Neis and Zipf (2012) made a very contributor centric study that we mentioned in several sections of this chapter. Trying to create maps to represent the aerial of every individual contributor, they found out that a bug in OSM editing software created version changes on object that were not involved in the specific editing activity. This may have influenced all studies that analysed interlocking and co-editing patterns. This insight emerged during this specific study although they were not interested in collaboration patterns. Arsanjani et al. (2015a) designed a contributor index to study the temporal and collaborative aspects of the development of OSM alongside an attempt to predict its development using cellular automata Markov model. The study is conducted over a regular grid (100 m). Interestingly for our study they also try to characterise the contribution with four variables *quantity*, *interactivity*, *semantic*, and *attractivity*. While quantity and semantic variables simply count nodes and average tags per cell unit, interactivity and attractivity are better fitting our research purpose. Interactivity counts the number of edits a node has been subjected to, attractivity counts how many contributors worked in the same area. Arsanjani et al. (2015a) found out that artificial surfaces are mapped earlier and in more details than agricultural and natural areas. As for most studies on collaborative mapping the focus was on data quality for mapping and cartographic purposes, hence other social and natural aspects have been overlooked. For instance, this study considered all the metropolitan area of a German city, therefore the mutual values of open areas have been overshadowed by the inclusion of urban areas that for mapping purposes includes much more details per square metre that a rural or natural area.

Interlocking is strong also where two only editors are contributing with several edits on the same object. This equates more to a dialogue rather than a social activity. We are therefore interested in the interactivity or attractivity according to Arsanjani et al. (2015a) which more efficiently than interlocking (Stein and Blaschke 2010) can express interest for an area. Loai Ali (2016) analysing collaborative mapping aiming at the semantic extraction of elements noted that usually natural features in the 10 most populated German cities (he analysed just the tags related to the supposed presence of grass like forest, garden, grass, meadow, and park) individually are just included in the database and never refined and only half of them is adjusted by another contributor. Loai Ali (2016) stresses how the heterogeneity of contributors highly affects the quality of the final resulting data collection. To extract qualitative information on society in CGI from mapping and cartography a relevant work is the one from Mashhadi et al. (2015). They went through several data preparation procedures, removing bulk uploads and considering only POI rather than including all road networks. In focusing only on POI, they compared the POI collected on OSM with two other CGI datasets considered as baseline for information. Our analytical approach, while mainly derived from CES research, resembles this approach in measuring popularity of real world features through collaborative interactions in the creation of the CGI element. Some studies focused on the internal dynamics of the projects but in our case, we are interested in the spatially grounded interactions that can witness the existence of environmental spaces. We assume that where there is more care for space, where there are more people trying to increase the quantity and quality of geographic information and knowledge, there we are witnessing a digital cultural practice. As for the photo collections incremental CGI has been used to determine some cultural benefits enabled by CES such as sense of places and inspiration (sceniqueness). Combining a crowdsourced encyclopaedic collection and social media (Wikipedia and Twitter) Jenkins et al. (2016) applied several statistical models to extract the sense of places from locational and textual information. They have not used a measure of popularity as the number of contributors but used the number of geolocated tweets and the number of accesses to Wikipedia pages in April 2014 for London, New York, Los Angeles, and Singapore. While extremely interesting it mixes how crowdsourced information is created and how it is consumed. Moreover, it focuses on one month only data to determine the sense of place using also geolocated tweets as given by the streaming API. But that according to the authors' admission the API gives are just a sample of 1 percent of the geolocated tweets that are just 2% of all tweets.

Seresinhe et al. (2017) used collaborative mapping only as a source of polygons to distinguish indoor and outdoor photos. But checking this approach against a computation intensive method proposed and documented in literature, it resulted less reliable to distinguish indoor vs outdoor taken photos. They use CGI extensively. They used *scenic or not* (http://scenicornot.datasciencelab.co.uk/) project that crowdsources aesthetic judgments on outdoor photos collected through another crowdsourcing project. It is a UGC over a CGI. But they neither did use the number of nodes nor the lineage of the collaborative mapping to deduct popularity of places. They used CGI just like was just an alternative to authoritative geographic information to derive the shape of buildings. But there are no studies that assess the quality of collaborative mapping as a reliable source of information on buildings. The already cited Bittner (2017) compared two CM projects to analyse the relation between digital and physical communities finding out that the two projects were reflection of the two opposing worldviews on the same contested area. This analysis used statistical census areas and was mainly focused on cultural differences rather than focusing on the relation between the populations and the features of the territory, but it stressed how the identity of communities, their cultural social and religious biases are reflected in collaborative mapping purposes.

3.6 Summary

In this chapter we have shown how CGI is a very heterogenous field. This open the possibility to several analyses not only of the produced data but also of the modalities of cooperation and competition. We have also seen how the social component is an unmissable element of the creation and success of such projects. We have also seen how ecological elements and attachment for place and digital enabled cultural practices can be detected through the analysis of the wide range of projects. We have also identified which cultural benefits have been detected using CGI finding out that a wide range of them has been detected using different kinds of CGI. While PPGIS retains its leading role in the detection of CES other approaches, while less effective in detecting CES are handier, cheaper, faster and at larger scale for the desktop creation of explorative studies. In the following chapters we will demonstrate it.

4 Methodology

4.1 Introduction

The literature review revealed how interdisciplinarity is strongly advocated in the ES field and even more when dealing with CES. This openness allows to have fruitful interactions with other disciplines like the ones focused on landscapes. This allowed to adopt sound and relevant methodologies in terms of interactions with local communities and analysis of local environments. But not all employed methodologies rely on long lasting and sound traditions. The endorsement of a cautionary approach is strongly advised when innovation and data intensive methods are adopted. Even the calls for a science without theory does not imply the ignorance of data lineage, the purposes for which data has been created, and the methodologies of data creation and validation. Those elements still set the relevance and the validity of data for the research at hand. A couple of such gaps strongly influence the geographical and analytical aspect of several studies aimed at ES. The previous chapters highlighted the main research trends related to the detection and mapping of CES and the opportunities offered using crowdsourced geographic information. The focus of this chapter is on the identification of gaps in section 4.2, we formulated two categories of them. One is mainly based on the substandard implementation of transdisciplinarity, a second one is based on the missed or underdeveloped opportunities. Following the gaps in section 4.3 we describe our research objectives alongside a detailed description of the expected results, in section 4.4, as well as the methods to achieve them in section 4.5. Sections 4.6 to 4.9 describe several stages of the designed method referring them to the relevant literature. Section 4.10 explains the cartographic design of some maps. A final summary (section 4.11) will recall the main points raised in this chapter.

4.2 Research gaps

In the following sections we explain the limitations that we have identified in the current research methods to elicit CES. Most of the time the gaps derive from the use of available data. Therefore, either data is used ignoring the methodologies and procedures to create and appropriately use them as shown in section 4.2.1, or there is an excessive focus on limited if not one only data sources (section 4.2.2). The studies on large scale models and assessments for other ES and their reliance on available data such as land cover, land use, and habitat led to the largely unsuccessful attempt to derive CES likewise (section 4.2.3). Even the most successful methodology to detect CES has its limitations (section 4.2.4) that can be mitigated once we embrace a wider range of data and attitudes, but also in this case attention have to be paid to the way data has been generated (4.2.5).

4.2.1Aerial unit and census data bias

The overall perspective on ES and conservation poses no limits to the scope and spatial extent of ES investigations, but most studies fall short of embracing this spatial opening (Palomo et al. 2014). Therefore, the aerial of spatial analysis when dealing with conservation and ES is debated. When dealing with the mapping of cultural hence ES with considerable non-material components even more difficulties are embedded. Studies aiming at CES tried to characterise their spatial extent using data on provision areas as well as data on receptors. This leads to the identification of designated areas and parks as spatial units of CES provision. But this excludes areas where the engagement for nature is less planned and excluded from everyday life. Dales et al. (2014) interestingly stressed how designations do not give actual numbers of visitors and how designations are quite wider compared with areas that actually provide benefits. Hence, there was no way to balance designated areas internally and to compare their comparative "performances" in terms of CES provision.

Focusing on receptors, studies on CES relied on the use of census data and associated spatial units. But the way they are delimited introduce a bias in the spatial analysis of open space areas (Pagella and Sinclair 2014) being census units intended for resident “sleeping” population and not for moving and nature connected people. In fact, standard census spatial statistical units are characterised by two main features. They are designed and tailored on resident populations statistics. We can consider the case of the UK census, but census spatial units are always created in a similar manner. We have a multi-tiered system that at every level aggregates units from the smaller level. The basic spatial units are postcode areas that cover all the territory. Data at postcode level are not available, the lower level with spatial statistical data available (Output Areas or OA) is constituted by aggregation of postcodes. But most of socio-ecological studies are not conducted at OA level, they are conducted at the higher level that is constituted by aggregation of OA. The Lower Layer Super Output Areas (LSOA) are geographic areas created aggregating from four to six OA. The threshold used for the creation of LSOAs is based on population and is between 1,000 and 1,500 inhabitants. Number of households and socio-economical homogeneity are also considered (Cockings et al. 2011). Census spatial units tend to be smaller on more densely populated urban areas. Therefore, this tessellation not only ignores every environmental aspect but also creates spatial units that in rural and natural areas can be so wide as to reach a surface of hundreds km^2.

In fact, for England and Wales we have the largest LSOA is 683 km^2 while the average LSOA is 4,35 km^2 thus including also the small urban LSOA that can be even less than 0.01km^2. Studies are conducted at LSOA level because at this level more statistical classifications (such as rural-urban classifications) are derived from raw data. The use of census units creates then a bias in spatial resolution. Therefore, studies that use LSOA to investigate CES or other studies, when working on open spaces, conflate very large portions of space using spatial units designed on human settlements where human settlements are quite rare. Hence census units as large as LSOA are failing to help detecting CES[41]. The lack of awareness of the way census units are created, bounded, and how the urban-rural classification has been derived for them in the UK implied also an erroneous use of LSOA data for regression models. Therefore not only the direct use of census units and census data but also all data which used the census spatial tessellation such as the rural-urban classification (Department for Environment Food and Rural Affairs 2013) are not suited to investigate the spatial extent of CES. Ignoring the inherent presence of census data in LSOA delimitations make this variable counted several times when applying regression models or investigating correlations that consider other census or census-derived datasets. An example of all the above cited improper uses of census data is in (Seresinhe et al. 2015).

4.2.2Recreation and tourism reliance

At present, the detection of CES has received marginal attention (Ambrose-Oji and Pagella 2012) and, when it has been performed, relies only on specific assumptions and mainly targeted recreation and tourism flows (Maes et al. 2012, Milcu et al. 2013, Pagella and Sinclair 2014). But this does not derive from a research priority but from the fact that recreation and tourism are easier to account for (Fish et al. 2016b). Those datasets are used on almost every large-scale study or model. More precisely we already reported in section 2.6 the three main clusters of studies using the classification of methodologies according to Potschin and Haines-Young (2013, 2016b) that attempted to map CES. In 2.6 we aggregated the large-scale habitats and models approach so also in the following section 4.2.3 we analyse them together.

4.2.3Consolidated methodologies based focused on large area model based studies (habitats, land use and systemic approaches)

The model/system based approach includes the studies conducted at national and even global scale where the extensive use of GIS involving network analysis, land use maps and large-scale surveys and census data, targeted mainly tourism flows and recreation (UK National Ecosystem Assessment 2011a, 2014, Medcalf et al. 2012, Willemen et al. 2012, Hernández-Morcillo et al. 2013, Plieninger et al. 2013, Bateman et al. 2013b, Dales et al. 2014, Maes et al. 2016). Those studies sometimes developed tools to analyse ES

[41] Large units may be useful for entire landscapes evaluations but their bounds might be derived from visual criteria rather than demographic then.

trade-offs including the other kind of ES whose measurement and detection is more related to biophysical elements. A first weak point of some of such studies is the cited use of census spatial units. In Bateman et al. (2013b) the creation of a trip generation function relied on the spatial resolution of LSOA. In this specific case for less densely populated areas LSOA can include areas quite larger than the minimum spatial resolution of two km^2 used. Willemen et al. (2012) modelling ES[42] to analyse trade-offs scrutinised a Dutch landscape at a regional level. Time series of land use maps were investigated to elicit change that is considered detrimental to cultural heritage designed areas in the vicinity. Cultural heritage has been included in the model through a governmental study[43]. This study, based on land use change targets explicitly cultural heritage without explicit involvement of people or of any form of survey. It will hardly fit in a social-ecological view of CES since it more resembles the studies based on land use change that we describe below.

Dales et al. (2014), after trialling a land use based map of the UK, used a 1km grid combining 14 spatial datasets mimicking more closely the UK NEA approach. Commenting on their work themselves acknowledged the limitations behind both approaches, first of all they were "not convinced that using habitats as proxies for CES is the best way of presenting the evidence base" (Dales et al. 2014 p. 33). They assume that using a 1km grid will not work on spatially finer grained elements. To curb some of the above-mentioned limitation some studies trialled the georeferentiation of the Monitor of Engagement with the Natural Environment (MENE) survey (Natural England 2011) to assess people's preferences (Bateman et al. 2013a, Day and Smith 2016). But the MENE survey even being a very useful and insightful resource has some limitations and several adjustments have to be performed on it before it was used to develop the model (Sen et al. 2014). In terms of intrinsic limitations of MENE we can list six of them.

Firstly, it targets only open space recreation while other elements of CES that can be obtained through other surveys have been ignored. Natural England developed a list of complementary surveys to get people engagement for natural and cultural spaces[44], a very similar list can be found in the specific section of the UK NEA devoted to shared plural and cultural values (Kenter et al. 2014b p. 63). Secondly, it targets only one-day activities, then ignores overnight stays. Thirdly, it includes only responses from England residents that are older than 16, then ignores visits from abroad, other UK nations and younger population. Fourthly, geolocation is based on a gazetteer and for larger open

[42] They prefer to use landscape services to better explain how the service is provided by the entire landscape.

[43] The referred study is in Dutch so we do not know if any stakeholder consultation has been performed during its development.

[44] Natural England. (n.d.). Note on surveys undertaken in the outdoor recreation, leisure, sport, heritage, and related areas. Retrieved November 9, 2016, from http://webarchive.nationalarchives.gov.uk/20140711133551/http://www.naturalengland.org.uk/Images/MENE-survey-comparison_tcm6-38001.pdf

space areas it can be quite imprecise. Fifthly, the survey is run on a predetermined calendar missing a day to day granularity in temporal terms. Sixthly, the results are available one year after the survey has been conducted. Lastly all this kind of surveys are generally very costly (Kienast and Holfenstein 2016). Another innovation introduced in Bateman et al. (2013a) was the use of network analysis to compute visitation numbers and distances. While it surely added a very valuable information for smaller scale and near urban areas it falls out of scale when considering larger spatial units less remotely serviced by roads. One study in the Republic of Ireland (Upton et al. 2015) used a combination of CM and a survey on forest visitation and management to model recreational forest management practices. But they used CGI only as a cheap source of geodata for natural areas. This study uses CGI data ignoring its social component but also ignoring to check if the data downloaded can be realistically used as a replacement for authoritative data. The quality of CM has been widely checked but most of the studies focused on the quality of road networks, few studies focused on natural elements (mostly green areas in urban settings) and no study focused on trails and nature walks that has been widely used in Upton et al. (2015). While this family of project provide some interesting results, they meet two overall main drawbacks as mentioned in Kienast and Holfenstein (2016). The type of output and the stakeholders' involvement are not clearly formulated along with the knowledge about the services themselves, but mainly what influences the mapping purposes is the lack of knowledge of spatiotemporal resolution of data. Such studies rely heavily on the availability and quality of the spatial data, especially land use/land cover data that have been proved to be ineffective for CES studies. While the availability of land use data can be considered a secondary concern for the UK, it cannot be scaled up at the global scale. The availability of very accurate land cover/land use maps is not widespread. Moreover, and the ones derived through remote sensing (Pettorelli et al. 2017) needs intensive validation, as suggested in Fritz et al. (2009, 2012). Secondly it assumes that supply and demand are covering entirely the defined spatial units.

The latter is obviously incorrect for mapping purposes since several reasons can create a limitation to the delivery of the ES from some specific areas, such maps show ecosystem functions rather than ecosystem services (Pettorelli et al. 2017). While this is a limitation for CES mapping Potschin and Haines-Young (2016b) stress how from the engagement and policy construction point of view this can be seen as an opportunity showing stakeholders the potential ES not enjoyed. An important result from a place based study that applies to the results achieved through this methodology is the weak relation that links designated areas with local people perceptions. Fish et al. (2016a) found out that local respondents were less moved towards the natural environment notwithstanding the study area had overlapping and diverse nature designations.

4.2.4 Place-based deliberative approach studies

The second group of studies is characterised by the fact that is focused on specific spatial units (region or a landscape) it may involve locals and tourists using deliberative methodologies often using participatory mapping techniques or PPGIS (Public Participation Geographic Information Systems). This has been recognised as the most effective way to proper detect the spatial component of CES. This kind of studies while very effective in getting local preferences have the following limitations (Brown and Fagerholm 2015): It is very difficult to scale up at a national level (context dependent). Such studies include recruitment and local engagement therefore they require a long and costly process that often ends up targeting self-selected audiences made up by the same age, educational, and social group usually made up by: white, highly educated, retired, mostly males (Brown and Fagerholm 2015, Fish et al. 2016a). Place based studies are also very often targeted only at residents (Plieninger et al. 2013, Fish et al. 2016a) missing the opportunity to stress the contribute that CES could provide to enhance tourism as a sustainable way to make nature and culture really sustainable[45]. Therefore, the study and selection of involved stakeholders must be done very carefully (Fish et al. 2016d, Kenter 2016a).

4.2.5 Novel methodologies that use crowdsourced information (desk-based interpretive methods)

The third group is made up of studies that take advantage of a more transdisciplinary approach. In this group we include also studies that are not directly aimed at CES but that targets some of their benefits (landscape beauty/sceniqueness, sense of place, belonging, inspiration, knowledge, judgment). Kenter (2016a) mentions the role that desk-based interpretive methods can play for the assessment of values for various ES. They are deemed to be rapid and relatively low cost.

Media analysis for instance is acknowledged as capable of delivering an image about conflicting communal values and beliefs of different social groups. A quick snapshot that, in relatively simple contexts, will suffice as an assessment for ES. CES in particular are deemed to be the most suitable for interpretive methodologies. "Cultural services are generally more idiosyncratic (i.e. they are expressed differently in different places), which advocates for a more interpretive approach" (Kenter 2016a p. 291). The interpretive studies in this group use techniques and datasets derived from people’s actions, those actions of care or interest for the environment are traced through their digital footprints. While this way these studies overcome several of the gaps listed for the other two groups of studies they embrace an approach without being fully aware of the limitations but also

[45] In Plieninger et al. (2013) the authors suggest capitalising on CES to provide local people clear benefits.

with hidden potential associated with crowdsourced information. Therefore, still lack in several points as follows:

CGI data is used just as data and not as the result of a project that implied the engagement of volunteers with all its components. They ignore the different identities of online communities ignoring internal dynamics (different social media component) and motivational biases (different purpose-built projects), therefore ignoring the differentiation among apparently similar data sources themselves (Casalegno et al. 2013, Tenerelli et al. 2016, 2017).

They lack socio-demographic awareness when dealing with crowdsourced geographic information (Capineri 2016, Haklay 2016b). While they assume that their results work for the overall population, specific studies demonstrated that they are affected by social, temporal, gender, and spatial biases as often stressed in several studies. The following studies that are not targeting explicitly ecosystem services stress the presence and the need to reckon with socio-demographic profiling of CGI (Mooney and Corcoran 2012b, Hristova et al. 2013, Li et al. 2013, Quattrone et al. 2014b, Comber et al. 2015, 2016, Mashhadi et al. 2015, Campagna 2016, Longley and Adnan 2016)

In engagement terms they also fall short on assessing and evaluating the implications of participation inequality (Nielsen 2006, Yang et al. 2015, Haklay 2016b).

They use a limited amount of resources focusing on just one kind of tool a time. Using just online photos (Purves et al. 2011, Casalegno et al. 2013, Martínez Pastur et al. 2016, Tenerelli et al. 2016, 2017, van Zanten et al. 2016, Jeawak et al. 2017, Oteros-Rozas et al. 2017); or just citizen science population surveys (Raudsepp-Hearne et al. 2010, Queiroz et al. 2015b), while (Mooney et al. 2010a, Neis and Zipf 2012, Arsanjani et al. 2015a, Bittner 2017) used CM only.

They use a wide range of analytical methodologies lacking to define a common ground that can be inclusive of a wide variety of datasets. In the following some of the variations when dealing with spatial analytical methodologies.

Quadrat counts with differently sized quadrants (from 2km to 27m wide); point pattern analysis (Tenerelli et al. 2017), temporal patterns analysis; Quantity of pictures, or spatial edits vs quantity of people submitting observations, pictures or contributing to the online maps; Analysis of tags, analysis of the content of pictures; correlation between land use, land features.

Some available sources of information that have the potential to give some aspect of the engagement for the natural environment have still not been used. Amongst the unexploited sources for social-ecological research we can mention: the abundance of geolocated social media data that are already used in cultural geography or spatial planning to target benefits or choices not related to the environment; The large amount of information about participation and interaction in the construction of CGI. That is often used just to have free geodata; The numerous collection of GPS tracks used already in tourism studies; The ever-increasing data produced by geosocial games like geocaching (Cord et al. 2015) and the growing success of games like Pokémon go (Juhász and Hochmair 2017) and Ingress. Not lastly the possibility to extract place information from text as in geo digital humanities or cultural Geography (Murrieta-Flores et al. 2013, Gregory et al. 2014, Derungs and Purves 2016). Environmental volunteering constitutes a cultural practice and contributes to healthier lives (Seymour 2016) but data documenting such practice has not been used directly for CES studies yet (Seymour and Haklay 2017).

The use of CGI data has been combined with traditional techniques in a way that a bottleneck requiring manual expert evaluation on every single crowdsourced element is required (Richards and Friess 2015, Martínez Pastur et al. 2016, Oteros-Rozas et al. 2017, Tenerelli et al. 2017). This limits the possibility of a large-scale adoption of the methodologies proposed[46] (Jeawak et al. 2017). This ends up creating a project that while using CGI covers an area whose extent is not larger that the areas covered with place-based deliberative approaches. Those are obviously more expensive but also more effective engaging populations in the deliberative process. Moreover, the expert selection inserts a bias in the pool of material used for the research[47].

A couple of studies sum up several of the above cited limitations of this group of works. Seresinhe et al. (2015, 2017) while combining several kinds of CGI and using several technologies made a work that was too reliant on technological affordances, so the social aspects looks underdeveloped. In particular their regression model relied too heavily on the resident population that is included as a census data but that also plays a role in the delimitation of LSOA. Moreover, they used LSOA units to derive sceniqueness of places, but LSOA are quite large in non-urban environment therefore their calculations may imply the aggregation of several *places* falling in the same area or the creation of false positive results. Unsurprisingly, the main outcome of the second study is that scenic areas are non-urban. Unfortunately, the technologically complex and articulated approach lacked awareness of basic UGC biases and the limitations of the use of census units for environmental studies. There was no assessment and no evaluation of underlying biases in the cohort of crowdsourced pictures used. In the second study, also the use of the rural urban classification of LSOA[48] falls short of getting environmental valid results. The conclusions of the two papers show how the methodological approach is hugely flawed

[46] In a case by case manner it can be reduced crowdsourcing the evaluations (Kachkaev and Wood 2013, 2014, Barve 2014)

[47] The mentioned case studies end up selecting a subset of pictures to check them manually using this way not more than a thousand of them.

[48]"The classification should be therefore regarded as a statistical classification based on population and dwelling density, not one based on landscape or the nature of a place." From https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/539241/Guide_to_applying_the_rural_urban_classification_to_data.pdf

by an overtly optimistic attitude towards technological affordances missing to assess the validity and appropriate use of CGI, census data, landscape analysis and social media uses to elicit landscape values.

Interdisciplinarity can drive to misleading assumptions on terminology. Lin et al. (2017a) while delivering a very interesting case study over the combination of some of the tools to evaluate ES as described in chapter two, used a misleading terminology. They claimed that they used crowdsourcing while using Facebook to advertise an online survey (Lin et al. 2017b).

4.3 Research issues and objectives

In the introductory chapter we already anticipated the main research questions in section 1.6. In this section we first answer to the ones that have already been addressed in section 4.3.1 and then we propose more detailed research questions for the analytical work in section 4.3.2.

4.3.1Answering to the research issues

In the following in Italics the research questions arose previously, they were associated in three groups the first was on the understanding of the intrinsic characteristics of the individual domains and the questions and answers are as follows:

What is the potential, nature, and limitations of UGC? For the individual domains we wanted to have a better characterisation of UGC with its peculiar characteristics and limitations. Those are then inherited by CGI. We found in section (3.4) that social, cultural, and spatial biases and participation inequality issues are features that characterise both limitations and values for our CES oriented research.

What are examples of uses of CGI to elicit social-territorial bonds? CGI has already a relevant heterogeneous dedicated corpus of research focusing on places. The most consistent bodies focused on the photos shared online, analysed both in content and position. We have also seen how other types of CGI have received a very limited interest as potential proxies for CES. Even biodiversity recordings run by volunteers, a cultural practice in itself, have been used very rarely to characterise places. Several studies analysed communities behind CGI, but those studies very rarely focused on social-territorial relations.

How CES research is conducted and what are the best practices? We have found a growing corpus on ES and CES research. In general terms three main approaches has been used for ES research while just one is the mainstream CES research approach. It is based on public participation deliberative approaches. Best practices and lessons have been learnt but some limitations still hold, and we will try to amend them with this research.

The second group of questions is about the examples of already performed cross contaminations between CES research and CGI. The research questions and answers for such methodologies are as follows:

How CGI can fit in ES research? The use of CGI in ES and CES research is compatible with the main frameworks adopted for ES and CES. The CES conceptual and analytical frameworks that have been analysed can include the use of CGI. We found several studies using CGI for CES research. But we also found several studies that, without references to the ES approach, use CGI targeting non-material benefits from human wellbeing that have been included in CES. More explicitly citizen science for biodiversity recordings is intrinsically a cultural practice located in space. Desktop interpretive studies are valued as relevant for CES research.

What are the early uses of CGI for ES and CES? The overwhelming majority of studies using CGI for ES and CES research use online platforms that share pictures when aiming at the analysis of people's perceptions both analysing locations in geographic terms and characterising landscapes through the analysis of the content of pictures. CM is very often used in ecology focused studies, but it is mostly employed as a cheap source of geographic data disregarding its socially constructed component.

The third group of questions is about the design of the methodology for our work, and the questions are as follows: *Which types of CGI have potential for CES research?* There are several options available. Learning from the experiences gathered in the previous chapters we can say that the kind of CGI is not crucial. All information that can be associated with mass participation and that has at least a unique identifier for every individual or group of contributors is suitable. CES focused research, concentrated on geolocated photo collections in several occasions. This allowed us to design our methodology to improve from the best emerging practices on photo collections and then we broaden the research method with some adjustments to other kinds of CGI.

Is there the possibility to design a methodology that allow us to compare the different CGI sources? There is not a large pool of studies that compared the different CGI projects but statistical and geostatistical methods that focused on one only CGI source or that compared different ES will be adopted. Going beyond the existing studies involving CES and CGI we will pick and adopt statistical and geostatistical methods that compare and analyse the spatial behaviour of CGI communities.

The fourth group of questions is about the characterisations and findings. Those questions are addressed mostly in the following analytical phase. We here evolved those research issues and questions in more thoughtful set of research questions that considering the results of the previous chapters will guide us in the design of the methodology and in the development of the analytical phase.

Does individual CGI projects inside the families works similarly? This can be helpful because we know that spatial and social biases are embedded in the spatial and temporal adoption of projects. Therefore, comparing the projects in areas where their adoption coexist can help understanding not only their differences but also their potential mutual replacement for areas where one project is spatially absent or in case one project is discontinued.

What are the aspects of places that can be elicited using CGI? As we have seen in the previous chapter, several aspects of CES have already been targeted using CGI. Improving the analytical methodology, broadening the base of CGI types involved we will also stress the aspects of CES that can be elicited from unprecedently used CGI sources for CES research.

Is it possible to implement a fully automated large-scale study to elicit CES? As we have seen studies targeting effectively CES are always at local scale while large scale projects have not been able to contribute effectively to CES research. Some results obtained using CGI relied on manual stages that hindered the possibility to scale up the analysis.

What are the limitations that have to be considered evaluating our work? We already know that CGI data has some specific features, we will try to evaluate how them and other factors influence the functionality and usability of the final outcomes.

What are the potential developments? This work is not seen as a development towards a predefined scope. From it several studies, tools and processes can be developed. Once the results have been evaluated we will be able to envisage some developments.

4.3.2Research objectives for analytical phase

Given the current development stage of our research after having reviewed all relevant literature we are now able to propose a limited but very targeted set of three research questions that sum up our main contribute to the advancement of CES research.

Demonstrate that a wide range of crowdsourced geographic information can be used to detect people's attachment for places.

The overall objective of the thesis is to demonstrate how crowdsourced projects can contribute to the detection of environmental spaces. This comes from the need to expand research methods that have been explored in sometimes erratic or incomplete ways trying to embrace innovative approaches. Achieving this research objective will enable us to evaluate a novel, quicker, and cost-effective way to support research in cultural ecosystem services and inform policy. To achieve this objective, the research project developed on the seeds of innovation that are emerging and providing interesting insights. A limited amount of data and methods used in this work have been used for similar purposes beforehand. Sometimes research projects were not aiming at unravelling social-ecological relations. This project is trying to complement such studies not only adopting the best fitting analysis methodologies but also using insights on the dynamics of crowdsourced information and communities coming from several projects, environments, and organisational settings.

Understand and identify which components of Cultural Ecosystem Services are more likely targeted with the use of a wider range of crowdsourced geographic information

The first seminal works in the field demonstrated that user generated information is a very rich source of information to elicit people's preferences. Relying on information that has been collected for other purposes, the analysis can emphasise some elements of the experience of the territory while neglect others. We stressed which component of CES has been targeted with the studies that already used CGI. Widening the quantity of CGI sources and using a unique methodology we are going to identify which other elements will be more likely targeted. It is therefore very valuable to understand both the specific components of CES that are targeted as well as those who are neglected by the analysis in this work. This can lead to improve the understanding not only of our work but also of the seminal works that preceded us to inform policy in a more correct way and to help design new research and policy agendas. Therefore, this work focused on a limited number of crowdsourced geographic information sources and scrutinised them against the features in the landscape. Moreover, the case studies have been chosen not to emphasise the computational challenges of a large-scale adoption but to focus on the fine-grained characterisation and identification of the analysis results.

Develop analysis methodology and research approach appropriate for the operationalisation of CES investigation using CGI

The overall family of crowdsourced geographic information includes several kinds of projects and technological platforms. Only some of them have been used in the few seminal works to explore social-ecological systems. Very rarely more than one of them have been used in the same study. Almost never the relation between them has been investigated. The link between the project and the information biases in the data, and their relation to the project's design and volunteer's motivations has been not explored. Therefore, while those studies used a heterogeneous range of methodologies they were unable to correctly elicit social-ecological systems due to the lack of awareness of the mechanisms that govern the behaviour of different groups of people in the open spaces. Studies on UGC, CGI and citizen science demonstrated that the use of untested user originated information as a barely cheap alternative to authoritative information is not correct.

Even more important is the fact that user generated information is dependent on the nuance of aggregation of people that generated the information. Therefore, applications that looks very similar from a technological point of view can drive to different results. Therefore, to extract valuable insights and operationalise our approach, the definition of a methodology that stresses the biases embedded in user originated information is mandatory. It will enable a more correct use and analysis of this kind of information, it will reduce the influence of the biases of individual projects, it will enable further studies to build on top of our initial exploration that focuses on user contributed geographic information. The common minimal elements given by the availability of human activities with associated positional information and the awareness that we are dealing with user generated information will support the definition of a suitable analysis methodology. The methodology developed for the present study therefore, considers not only the elements of spatial uncertainty associated to the detection of cultural elements but also the internal dynamics of social groups that collect information with spatial references. The materials used are: explicitly recorded geographic information readily available for data analysis and consumption, information regarding the collaborative construction of this information including enhancements and number of contributors. Studies demonstrated that this material is able to achieve the proposed research objective. The comparative spatial analysis of the collaborative information will enable this study to stress biases between projects. Our willingness to avoid the analysis of the individual elements with position associated (photographs, observations and articles or geographic elements) is explicitly designed to allow a large-scale implementation of our methodology. This work will conduct to further recommendation to operationalise our approach.

4.4 Main expected results

The definition and test of the methodology will allow us to reach two main results. Firstly, we will achieve a characterisation for each project. Secondly, we will achieve a characterisation for the families of projects that share a similar scope or a similar technological platform. Without the need to analyse individual contributor’s motivation an overall characterisation of every individual project will be performed. This will also allow to compare projects independently from their individual technological platforms. When dealing with a family of projects, the identity and characteristics of every project will be measured against their ability to stress the existence and the relative magnitude of a CES. Given the exploratory analysis characteristic of the current study, a thorough evaluation of the analysis outcome will be performed to characterise the commonalities shared among projects pertaining to the same family of crowdsourced geographic information. As for the identities of individual projects, a group characterisation will be sketched.

4.5 Methodology overview

The thesis follows a sequence of actions. Their sequence is shown in Figure 4-1.

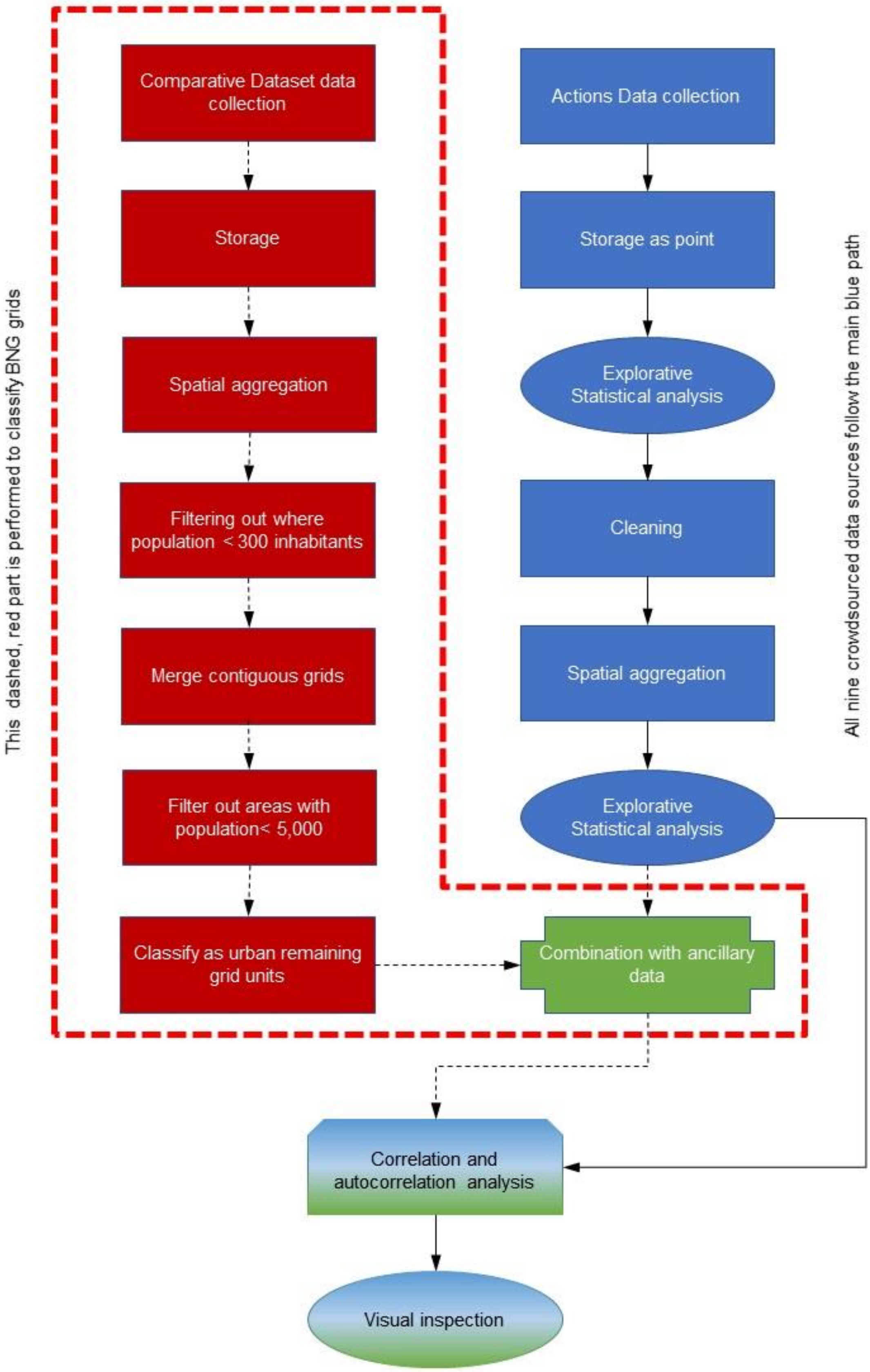

Figure 4-1: Flow diagram of the developed methodology. Dotted lines and arrows identify processes that are not followed in all case studies

The main elements of the current work are ecological and social elements and their interaction in space determining the perceptions of environmental spaces. This then need to be undertaken in a methodology that considers residency places and places people care for, either performing place based activities or contributing to the creation of information related to places. The comparison between the places where people reside and the density of people's digital actions is our lead towards the places people care more. The number of people that contribute to crowdsourced projects gives the density of people's actions. Crowdsourced data offer a very wide variety of data that has been analysed in several ways. In this work, since its purpose is to compare different projects, there has been the need to select the common elements that can be found across all projects and that can provide a measure of people's engagement. We therefore suggest that a basic measure of this engagement is the number of unique contributors per square km, as the proxy that has been used for the attachment of people for space. This is in line with what made by Casalegno et al. (2013) and Tenerelli et al. (2016) Purves et al. (2011). Some works as in Tenerelli et al. (2016) used the number of pictures rather than the number of unique users.

On another side Osterman et al. (2015) used different methodologies that are more specific to the CGI type they used; that drove to more accurate analysis and conclusions. While the relevance and quality of those studies is remarkable their analysis technologies are tailored to some specific characteristic of the CGI source at hand. Given the wide range of projects here analysed those methodologies cannot be applied to some of our projects therefore not contributing to our explorative and comparative analysis. To evaluate the interest in places for some CGI projects we used to evaluate the movement of people from residency urban places to outdoor areas. Therefore, there are two main kinds of data used in this work, one focusing on digital footprints of people and another focusing on people's residency places. The assessment of residency places has been used as a baseline to compare the spatial distribution of the resident population versus the spatial distribution of the digital footprints of cultural practices. The creation of the baseline distribution of resident population required a unique data extraction and modelling process.

Contrarily, to extract and use the information from the crowdsourced projects several distinct procedures has been implemented. This was needed to extract the common components from every individual project and prepare them for analysis by bringing them to a standardised format. The analytical phase will than make use of the standard format that allows to compare the different projects. The discrepancies that forced to standardise the data are due to the way platforms are designed, the services projects offer and the openness of the data they hold. Another aspect of the standardisation is the result of the focus on positional information. Here, there is a need to bring all data set to a common reference system. We have then two main blocks of data, census data and CGI data. Those must be made comparable using a common spatial arrangement. The left hand red column in Figure 4-1 is devoted to the collection of demographic datasets, their aggregation over the chosen tessellation and the consequential classification of the spatial units. Not all our analytical work will require the use of demographic data, this is why we encircled this section with a dashed red line and all this part is characterised by dashed arrows used to describe the workflow. The right hand blue column in Figure 4-1 is devoted to the preparation and collection of CGI datasets. Here there are all the phases for their analysis, the data collection, the preliminary statistical analysis, the removal of false positives or data preparation, the aggregation over the chosen tessellation and the final analysis of those achieved spatial data. The blue column characterises processes that will be performed on all CGI datasets, this is why the arrows connecting the various phases are depicted with solid lines. At the bottom of Figure 4-1 we have the two final analytical tasks that will be performed on all CGI datasets. The importance of the spatial tessellation and the census data is highlighted by the first two sections that follow. A brief explanation of the way space is subdivided and named in the British National Grid is the first section that follows (4.6) this section explains the choice of the analysis grid and the need to align geographically all crowdsourced and demographic data. The need to overcome the limitations caused by the use of census units for ecological studies, as explained above, induced us to develop a procedure to allocate population to our analytical spatial tessellation and to classify the spatial units to distinguish the urban and the non-urban ones. Section (4.6.1) describes the demographic data extraction and manipulation.

Figure 4-1 shows how several passages are devoted to the collection and preparation of crowdsourced information. In terms of crowdsourced data in the following there will be the explanation of the main phases of the project and their implementation related to the different projects used. The phases are:

- Collection (4.7)
- Storage and preparation (4.8)
- Analytical Section (4.9)

The processes specific for every individual project will be explained in the chapters devoted to the individual projects. The analytical phase taking advantage of the gradual standardisation process combines all datasets and calculates their correlation. The last computation made for analytical purposes is the spatial statistics accounting for the autocorrelation of the distribution of contributors. To answer to the research questions posed in the previous paragraphs a final visual inspection of results is needed. The analytical phase is the one where the different datasets are subjected to the same processes therefore this phase will be explained just once in this chapter. The visual inspection phase compares the result of the analysis with the real wold natural and cultural features trying to understand the real impact of the approach undertaken.

4.6 Geoinformation common framework

The British National Grid (BNG) is the reference grid used by Ordnance Survey (OS) on all its products. It is also called Ordnance Survey National Grid. Therefore, it is the reference system used on all official maps in the UK. Its design makes it different from the most recent geodetic based reference systems. Therefore, a grid based on the BNG is not overlapping with a grid based on other systems like WGS84 that is now globally the most common reference system. All geographic data in this work has been analysed according to this grid and not based on a grid aligned to WGS84 since some of the data used (as shown in chapter 7) has been provided only with a reference to several levels of the spatial subdivision of this grid. While all the data that has been used in this work with reference to WGS84 is made by points then they can more easily be projected in a BNG grid. The British National Grid upper level is characterised by squares of 500km identified by a letter. Every grid is recursively subdivided in smaller grids whose size is 1/10, whose naming starts from the upper left corner and then continues horizontally and then at the end of the line it continues at the following left side of the line. Following a system to subdivide and name every grid recursively using letters to identify digits with letters and numbers to identify numbers there is the following hierarchy:

1 digit “A” identifies a grid of 500 km

2 digits ”AB” identifies a grid of 100 km also called "myriad"

4 digits – 2 figures ”AB01” identifies a grid of 10 km also called "hectad"

5 digits with final letter "AB01X" identifies a grid of 2 km also called "tetrad"

6 digits – 4 figures ”AB0213” identifies a grid of 1km also called "monad" or "grid square" in Geograph

8 digits - 6 figures ”AB024135” identifies a grid of 100 m also called "centisquare"

10 digits - 8 figures ”AB02461357” identifies a grid of 10 m

12 digits -10 figures ”AB0246813579” identifies a grid of 1 m

Therefore a 12 digit grid is inside a 10 digits grid. The containing 10 digits grid can be derived deleting the numbers that in the example above occupy the position of the digits 8 and 9. This characteristic will be used to manipulate the environmental observation datasets that contain observation with different precision levels. In order to collect, analyse, and visualise in map format the result of the analysis, in this work the following data has been used:

GB National Grid Squares [SHAPE geospatial data], Scale 1:250000, Tiles: GB, Updated: 2012, Ordnance Survey (GB), Using: EDINA Digimap Ordnance Survey Service, <http://edina.ac.uk/digimap>, Downloaded: Tue Jun 25 17:31:07 GMT 2013

GB National Outlines [SHAPE geospatial data], Scale 1:250000, Tiles: GB, Updated: 2005, Ordnance Survey (GB), Using: EDINA Digimap Ordnance Survey Service, <http://edina.ac.uk/digimap>, Downloaded: Tue Jun 25 17:31:07 GMT 2013

The different CGI datasets come with different ways of conveying geographical information. Data obtained through APIs included geographic information as points referred to the WGS84, which is used for GPS coordinates. Ecological data dumps are more problematic for two elements, the way they are georeferenced and the spatial precision. In terms of georeferentiation, ecological data are always referred to the British National Grid. Some of the ecological data also include longitude and latitude data but they are often referring to a corner or the centroid (centre of the geometric representation) of the corresponding British National Grid unit.
Referring to the precision issue, ecological data dumps often include in the same file data referred to different precision levels. As an example, the same dataset includes recordings that are referring to a grid of 100 metres grid unit, while some others are referring to one km grid unit or to 10 km grid unit. For the development of the following analysis this implied that all observations collected with higher precision to the 1km grid level had to be aggregated at the 1km grid level. To run the spatial analysis and to create a uniform methodological background a 1km grid based on the British National Grid (BNG) has been chosen as the reference grid for the following reasons.

- It represents a good approximation of locality as in Purves et al. (2011) and in Neis and Zipf. (2012).
- It is used as the basis for most of the reports on territory and landscape in the UK as the NEA scenarios (UK National Ecosystem Assessment 2014), The Countryside Survey (Carey et al. 2008, 2009)
- Most of the environmental observations in the United Kingdom, where all the case studies are located, are recorded using this or a higher level of precision. Moreover, environmental observations in the UK are always primarily if not exclusively referred to the 1km BNG.
- A 1km spatial grid, even when not based on the BNG, is used in several cases in literature and in different disciplinary domains that are combined in this study. As an example, in Geographic Information Science Antoniou et al. (2010) and Li et al. (2013)

that used a similar grid; in Ecology Casalegno et. al (2013); in cultural landscape studies van Zanden et al. (2016), Arsanjani and Bakillah (2014) in CM research.

- At continental scale also, most studies analysing ES and CES use the 1km grid (but based on WGS84) as the EU wide HERCULES project on cultural landscapes (Kuemmerle et al. 2014).
- The derived demographic data used is aligned to the BNG and its manipulation is optimal using a compatible grid

Therefore, not only the crowdsourced data but also the demographic data will be manipulated to be available as an information layer referred to the 1km BNG.

4.6.1 Demographic data and urban classification of grids

The use of datasets where human actions are recorded for other purposes implies the possibility that irrelevant or misguiding information is harvested alongside relevant information that links humans with ecosystems. This misguiding information is more likely concentrated in urban environments where the CGI projects have a strong social/visual component and nature is not an obvious target of actions. This required the adoption of a methodology to select non-urban environments. Demographic data is the main comparative dataset used, the flow diagram of its derivation and use in the overall project is described in the left hand side red column of Figure 4-1.The population map of the area based on 1km² grid units has been derived from the online tool developed for the Population 24/7 project (Martin et al. 2015). Data was initially hosted by UK Data Service in http://www.osgl.soton.ac.uk/data/pop247/, where it has been obtained on 24 March 2014[49].The project allowed to download the statistics of population aggregated for demographic classes and time ranges. For our project there is the need to distinguish between the places where people reside and places where people go voluntarily.
Therefore, the data collection focused on night time residency of all age groups. This required to download different datasets. The dataset came in raster format, the pixels are aligned with the BNG and they cover areas of 200X200 meters. The raster format files were summed up point by point to have a raster map of all population in the area. After their procurement, to verify if the process gave a correct estimation of the resident population, the resulting total population has been calculated and compared with the total resident population as can be calculated summing up all Output Areas intersecting the study area. After a positive evaluation of the procurement and aggregation of age ranges the resulting raster map of population has been intersected with the 1km BNG shapefile of the study area. This gave as a result the 1km BNG shapefile with a calculation of the resident population in every grid unit.

[49] Data is no longer available in UK data service.

The 1km BNG population shapefile is needed to follow the procedure used by OECD (The Organisation for Economic Co-operation and Development)[50] to classify urban and non-urban areas. The methodology to derive urban clusters, as described by Dijkstra and Poelman (2014) has been applied. The methodology was developed to compare urban areas at a trans national level. It avoids using demographic or administrative delimitations since they are usually not consistently defined in distinct countries. In our case it is helpful since for environmental spaces administrative and census units are not fit for the purpose of mapping. The first two steps of the methodology are as follows: Firstly, select all 1km grid units where there is a resident population exceeding 300 people. Then aggregate contiguous 1km grids in a unique area and calculate its total resident population. The so derived larger areas that overcome a total population of 5,000 are classified as urban. The resulting urban areas greatly overlap those obtained by the GIS of the European Commission (https://circabc.europa.eu/sd/a/a932d937-82fe-48b6-9c14-1d549ac494f3/URB_CLST_2006.zip) using a different $1km^2$ gridded population referred to WGS84 (Gallego 2010). The resulting non-urban grid units are the focus of this study.

4.7 Crowdsourced data collection

The collection of data followed three main modalities that are related to the way data is accessible. Collection of data dumps is the most common way with five projects that have been accessed this way. The collection through APIs was the second most common way of data collection[51]. Some data was not available through either dump or APIs but was readable on the web. Data dump collection occurred in all the three categories of communities analysed. Amongst them there are dumps freely available and accessible online and dumps that has been accessed through an agreement with the institution holding and collecting data. The format of data provided includes, Tab Separated Values (TSV) files, Comma Separated Values (CSV) files, a specific implementation of eXtensible Markup Language (XML) and shapefiles (the most common GIS industry standard). Data dumps obtained by collaborating institutions (GiGL, iSpot, iRecord) are tailored to the specific case study areas. Data dumps obtained freely from the projects (Geograph and OSM) are often covering larger areas, therefore a clipping process is needed. Given the high specificity of this process it will be described with each specific data source project. Application Programming Interfaces (API) are interfaces designed to share parts of the content of a specific application. The ones accessed for the current work have some commonalities, but they differ in size and

[50] http://www.oecd.org

[51] Data from OpenStreetMap is accessible through API but it lacks the lineage information that is used for this work. We then used a couples of global data dumps which include all lineage for OSM data.

extent. There are APIs which allow the access to a limited amount of data. Sometimes there is a lack of documentation – in the extreme case in just one simple webpage (Panoramio). In more information-rich cases there is the possibility to access several services including testing versions also called sandboxes (Flickr, Wikipedia and Wikimapia) that allow to run part of code to have a preview of the results. In the latter the explanation of access methods and parameters takes several web pages often organised in wikis. The APIs work as follows. The user posts a request and the API send its reply in a bundled file using specific compact formats, that then need to be parsed. The most commonly used is JavaScript Object Notation (JSON), it is a lightweight data-interchange format specifically designed to deliver data in the web. When queried for geographic areas APIs can convey in one bundle file all geographic data points contained in the area, it can go from zero to hundreds of elements. Geographic data points constitute the minimum spatial information unit. The process used to collect information from APIs follows several steps:

- Decide extent of case study area
- Select grids covering the study area from the BNG
- Study API of the individual project to check
    - Information downloadable
    - How to geographically query it
- Create the database to store data
- Design code to query API and store in the database results
- The code goes by every grid unit of the BNG and
    1. Extracts the coordinates of the bounding box (BBOX)
    2. Queries the API including the BBOX coordinates
    3. Reads the result sent from the API
    4. Parses them
    5. Tests if the data point at hand is already in the database
    6. In case the data point is not in the database the code stores it in the database.

Point 5 above is necessary since APIs always request two non-adjacent vertices and then it derives a rectangular area which sides are delineated along WGS84, while the grid used to extract the BBOX is based on BNG therefore the sides of the grid unit are not aligned. This implies that there is always an overlap between two adjacent bounding boxes. This is further explained in Figure 4-2 where the bounding box chosen for the first case study has been chosen aligned to the BNG. In Figure 4-2 a translation of the vertices of the BNG bounding box to WGS84 coordinates was performed. When dealing with APIs whose geographic parameters must be inserted using WGS84 there is the need to use the WGS84 bounding box of a BNG grid unit as can be seen in Figure 4-2 below.

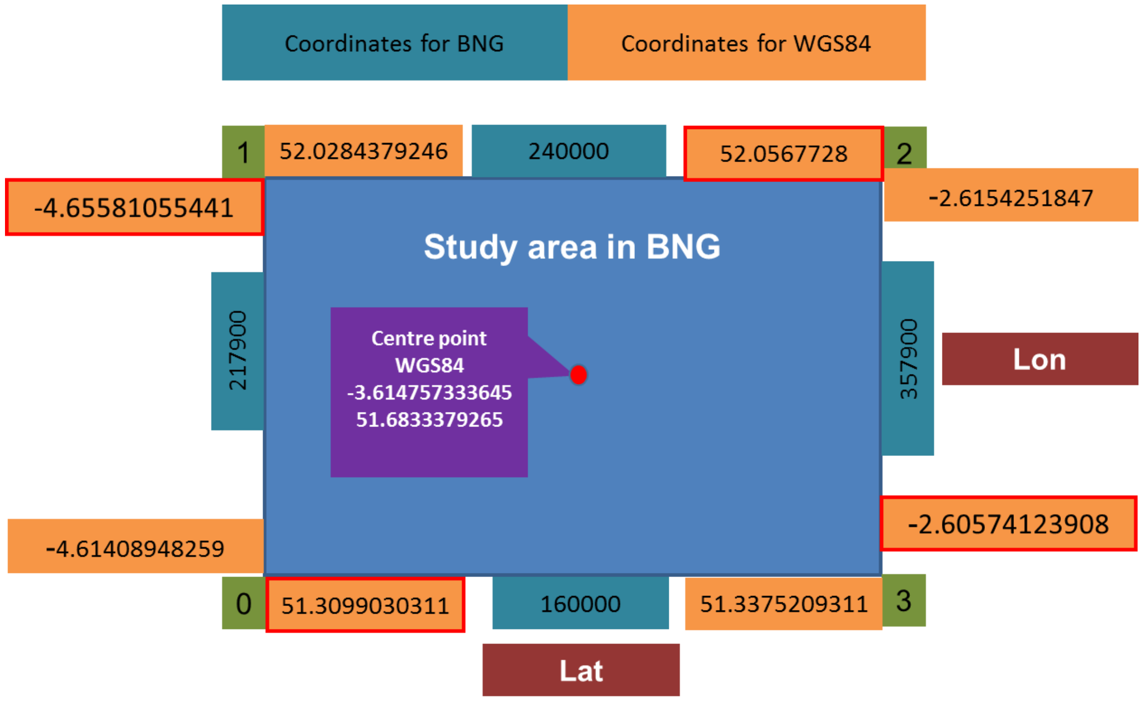

Figure 4-2: BNG grid units and WGS84 bounding boxes relations

The need to avoid not queried areas than defined the bounding box area as the one having the maximum and minimum longitude and latitude of each vertices of the WGS84 BBOX. The chosen coordinates are the red bounded WGS84 coordinates in Figure 4-2 above. Sometimes the information to analyse people's preferences, was accessible online but neither available through data dumps nor through APIs. To use such information a process of reading and storing in the database has been performed through the design of specific code (DOM parser). This process is often referred to as web scraping and is widely used in GIS to extract demographic information (Chow 2013) and in Ecology is also considered a viable web technology to extract information from web pages (Vitolo et al. 2015).

4.8 Information storage and preparation for analysis

We used a database to store, query, and elaborate the information collected for this work. Some of the used information sources include more than one million elements. Database based on SQL (Structured Query Language) are also the main supporting storage and analytical tool for GIS. The storage of data points information implied the creation of a database using PostgreSQL enabling the PostGIS extension. A specific table of the database included the grids covering the study areas. To assign every data point to a specific BNG grid the points must be linked to the grids. A GIS process called spatial join operates this way. When every point got its reference to the corresponding BNG grid unit then they can be queried and the resulting information aggregated at the 1km BNG level. We processed the demographic data similarly but starting from a raster file. The demographic table then underwent other specific spatial queries to assign the OECD classification for every 1km of the BNG as described in 4.6.1. When the tables representing the number of contributors to every project and the OECD classification include a column with the references to the 8 digits BNG code they can be connected by the database in a table or queried to perform the analysis. Not all data points can be considered as being originated by a spontaneous action from people when connected to the physical environment. Analysing all data points may lead to several false positives or type I errors. Therefore, some data preparation has been needed to remove all such automated edits before the individual data points are aggregated for every BNG grid unit. The designed methodology based on quadrat counts of number of contributors and the selection of urban areas only allow us to filter out already some of the elements that can mislead our analysis. For photosharing projects, we do not need to apply any other data preparation technique.

A characteristic that specifically applies to the citizen science biodiversity recordings data is the heterogeneity of geographic reference. Since the main purpose of biodiversity recordings is the recording of observations, their intended use is not mapping. Therefore, the main aim of observer is to identify in the most correct manner the observed "element of nature". In fact, in almost all datasets there is a field where there is reported an expert (*Determiner*) who double checks the observations and corrects the classifications. In terms of geographic reference there are several levels of complexity. Firstly, we found data recorded using two different reference systems, often even in the same project and information table. Secondly, we found data recorded using several precision levels in a unique file. This happens especially when referring to the British National Grid. Thirdly environmental observations are often not related neither to a grid nor to a couple of coordinates, but they are often referred to a "place" that can be an urban park, a woodland. Then digitizing this content for their exploitation through GIS those observations are geographically referred to the geometric centre of the shape of the area they refer to. So even if an observation has a geographic point associated to it this does not correspond to the exact location of the observation. In chapter 7 we will show the data preparation procedures and the resulting datasets for citizen science biological recordings.

Incremental projects are subjected to two main misleading factors. The first is that the main objective of those project is the production and refinement of knowledge, not necessarily knowledge about or in nature. The second misleading component is more subjected to the action of automated edits. In contrast with all previous projects, incremental projects are often improved through some automated processes. The need to filter out edits aiming at not natural elements can be done in two ways dealing with incremental projects. The first option takes advantage of the folksonomies that are produced by users and select only natural elements. The second option consists on the use of the same method we used dealing with photographic collections, we filtered out urban areas using the OECD classification methodology for 1km grids. Then we focused our attention to non-urban areas. Our chosen option is the latter because this way one can evaluate not only the interest for the specific natural object but also its relationship with not explicitly natural object to understand which of them acted as a catalyst of localised interest. To deal with the second misleading component, the need to filter out automated edits from our incremental projects is unprecedented.

While photosharing projects and environmental observations are not subjected to such problems, incremental projects are subjected to automated edits. *Bot*s are piece of code that edits data stored in the internet. Bots can perform several kinds of works: they are used for imports, grammar, and language check, in general they offer automated services that are performed over large amounts of data. Since their actions are automated and not supervised, their contributes are not considered as a valid proxy for human preference or at least attention for a specific place in the real world. We followed several approaches to remove automated edits. The most affected projects have a rich documentation on automated edits and most of them can be traced this way. To add another level of scrutiny we also added a string matching search over the usernames of contributors to search for names, or part of them, that can refer to bots or imports. This procedure will be explained more in details in chapter 8, devoted to the analysis of incremental projects.

4.9 Analysis

As mentioned before Steiger et al. (2016) warned about several limitations that can arise when dealing with CGI from social media which is particularly rich and heterogeneous but also strongly biased source. Our analytical phase tends to stress the grids or groups of them where there is a statistical significant density of contributors. This is assumed to be a proxy of interest for the area. This study is an exploratory analysis; therefore, a statistical analysis of partial results follows every stage. The analysis followed several steps. It first focused on the creation of the count of the number of contributors for each 1km grid unit of the BNG. A second optional stage used the demographic data as described in (4.6.1) to distinguish and classify between urban and non-urban areas. Then statistics were developed to compare the different projects of the same group using a scatterplot diagram and calculating a correlation index. The last analytical stage is the application of spatial statistics to the total amount of contributors to derive clusters and hotspots of grids. The analysis used a GIS to verify whether the projects analysed are suitable sources of information from which to infer the location of environmental spaces.

4.9.1 Quadrat counts

The methodology used several tools applied in exploratory spatial data analysis (ESDA) based on quadrat counts (De Smith et al. 2007). The main methodological references are several studies. Considering a first group of studies where photosharing websites has been used Kisilevich et al. (2010a) used one approach with two different types of baseline data, Flickr and Panoramio. Casalegno et al. (2013) used only Panoramio data and demonstrated the clustering of Panoramio's imagery in specific hotspots. Tenerelli et al. (2016) applied a quadrat count over the number of pictures shared on Flickr and Panoramio. In the follow-up study Tenerelli et al. (2017), just used Flickr and a spatial tessellation of 27 m, Hochmair (2010) used a combination of sources and aggregated information along main roads on 500 meters grids. Lastly the HERCULES project (Kuemmerle et al. 2014) aggregated all its information at 1km level including also Panoramio images. In contrast to Kisilevich et al. (2010a) study, which focused on temporal patterns, here they examined spatial patterns.

We have several reference studies for the application of this kind of analysis. In the field of cultural geography van Zanten et al. (2016) for a continental scale study used the number of "posts" for Flickr Panoramio and Instagram aggregated over a 5km grid. Purves and Derungs (2015) text mining geographic information from tags and textual description associated to images, used a smart way to tessellate space with adaptive grids that adapt to the amount of content found in the individual grid cell. In the latter case there was an aggregation of information for every grid unit. On a similar subject Jenkins et al. (2016) did not use spatial aggregations using Wikipedia and Twitter but their study is more focused on the meaning rather than on the location of places.

Seresinhe et al. (2017) seeking to determine sceniqueness of places used Flickr, Scenic or not[52] and OSM. This work also confirmed how the number of contributors more than the number of pictures is a proxy for spatial preferences. The combination of Scenic or not and Flickr has also been used by Jeawak et al. (2017) to evaluate sceniqueness analysing Flickr tags. They also extracted other environmental information such as land cover, species distribution, and climate information just analysing Flickr tags. One of the main conclusion of Jeawak et al. (2017) is that Flickr data can complement but not replace authoritative information. In the collaborative mapping side, we can mention two studies: Neis and Zipf (2012) determined the activity areas of an individual contributor using several aggregations and simplifications where a threshold of 500 m has been

[52] An online tool that asks people to evaluate the sceniqueness of pictures taken from the Geograph photo collection

used; Arsanjani et al. (2015a) used OSM lineage to count the number of contributors in grid units of 100m and called it *attractivity* of the area. Studies focused on collaborative mapping projects used ESDA for large scale studies as in Comber et al. (2015, 2016) where data points aggregation for crowdsourced land use has been made on a 50km grid. Roll et al. (2016) used Wikipedia articles accesses aggregating global spatial information in degrees (going from a 110 km grid at equator to 0 at poles). Haklay et al. (2010) used 1km grid cells to aggregate contributors to a collaborative mapping project. As mentioned before, participatory mapping is considered the most valid methodology to detect cultural ecosystem services. Using PPGIS (Brown and Fagerholm 2015) cited studies where polygonal or raster gridded analysis is the result of the calculation of a kernel density estimation from point data for the mapping of ecosystem services.

4.9.2Contributors vs actions

The decision to focus on the number of contributors rather than the number of spatial objects created is in accordance with several studies (Girardin et al. 2008, Kisilevich et al. 2010a, Kachkaev and Wood 2013, 2014, Casalegno et al. 2013) focusing on pictures. Those studies used the number of contributors sharing pictures online as a proxy for the relevance of place. Wood et al. (2013) used a photograph-user-day metric, which mixed spatial and temporal aspects but revolving around the number of users. Similarly Cord et al. (2015) using open air games to assess recreational values used the number of user accessing a specific location (quest) rather than the number of locations per spatial unit.

On the other hand, there are several mentioned studies that used the number of spatial objects or actions rather than the numbers of contributors as (Antoniou et al. 2010, Richards and Friess 2015, Tenerelli et al. 2016) or as environmental observations (Raudsepp-Hearne et al. 2010, Queiroz et al. 2015b) or as number or the spatial extent of geographic objects added as in Hochmair (2010). In between the two extremes we have (Haklay et al. 2010) that compared the extent of spatial objects with the number of users who contributed to their creation. The study of contributions to OSM is often performed to study the community or the reliability of data. In one case (Polous et al. 2013) the number of contributors and contribution to OpenStreetMap has been used to see if local knowledge about events can be extracted analysing the history of the development of OSM.

Our choice, while supported from a considerable amount of literature, came after one of the multiple preliminary analysis of data that we performed at every stage of development. It was evident that some individuals perform several actions in a specific area in relation to a specific interest that is not related to the natural or cultural environment. Therefore, an approach based on the density of actions can be misleading in drawing conclusions. The detection of such biases happened both in the case of photosharing and in the case of citizen science. For example, around Swansea Airport, there are many pictures from one contributor, who is passionate about plane-spotting. During the analysis of citizen science in London two of the most important hotspots are determined by two individuals that collected tens of thousands of moths' observations using light traps at home. The analysis tended to create two main statistical elements. The number of contributions per grid cell, the number of contributors in each grid cell. In addition, for the citizen science case one more parameter for evaluation have been derived: the volunteer productivity. The volunteer productivity per grid cell has been defined as total number of observations/total number of volunteers.

4.9.3Study environments

In terms of spatial coverage in this study there will be a different focus in the different case studies. In the case of photosharing, several studies have already examined urban areas, our case study on photosharing platforms as well as the one on incremental CGI focus on non-urban environments. This work extends part of the work of Casalegno et al. (2013), but including a wider spectrum of projects, as well as broadening the analysis to overcome the strong coupling that occurs between one application and its contributors. This is in accordance with the outcomes of studies focused mainly on urban areas (Antoniou et al. 2010, Kachkaev and Wood 2013, 2014). The distinction between the focus on urban or non-urban areas is also distinctive of studies on collaborative mapping projects. Studies on OSM demonstrated how data quality increases in urban areas (Haklay 2010b). Therefore, urban areas have the most contentious data with all highly edited elements in OSM (Mooney and Corcoran 2012a) coming from them. A specific study on natural features in OSM was only interested at evaluating data quality (Mooney et al. 2010a).

OSM has also been used in Ireland to assess recreational values of forested areas, but unfortunately it focused only on them (Upton et al. 2015) and OSM has been used as a cheap source of accurate data, without previous evaluations on it fitness for the purpose comparing OSM data with authoritative datasets. This thesis work considers the number of contributors sharing activities on the three platforms for each case study as the main proxy for cultural attachment to the particular territory.

To focus only on urban areas the urban classification of grid units as derived in (4.6.1) has been used. This allows to filter out the urban areas where the vast majority of photosharing and incremental activities were performed. This work focused on the areas classified as non-urban according to the definition used in the second stage of the typology produced by OECD (Dijkstra and Poelman 2014). The classification was based on population counts in each grid unit and in the sum of the population of neighbouring grid units. Biodiversity recordings are implicitly focused on nature, therefore for the citizen science case study we omitted to filter out urban areas.

When we analyse citizen science originated biodiversity recordings, for each dataset we tried to encapsulate the variety of reasons a grid cell might be a recording hotspot, e.g. the 'home patch' of one prolific volunteer, popular visitor spot, area target of a specific conservation organisation. Other evaluations are possible such as the influence of high taxonomic diversity, high species richness, but they are specific to the citizen science case study only and their development can be found in Boakes et al. (2016).

4.9.4 Contributor ratio and Correlations

To assess how behaviour differed by platform, a per grid unit comparison was performed. This does not consider proximity and topological relations among grid units. Since all nine projects considered have a wide range of participants, we needed to define a parameter to compare between data collections which range from few thousands to more than one million elements. The grid Contributor ratio (Cr) has been derived for every grid square. The Cr is the percentage of contributors for each grid unit compared with the total number of contributors to the specific platform in the study area. As an example, the total amount of contributors to a specific project in the case study area is 1,000 while in a specific BNG grid there are 20 contributors. The contributor ratio for the chosen grid for the chosen project is therefore Cr=20*100/1,000. Therefore, the contributor ratio Cr for the specific project for the specific grid unit is 2. The Cr has been used for most of the work to compare visually and computationally all projects. It has been used to visualise through maps and scatterplots spatial distributions and overall projects correlations at the individual grid unit stage.

In those maps, a unified range of Cr values classes allows to visualise the spatial distribution of volunteers for one project but at the same time allows visual comparison with the other projects. To compare in a grid by grid basis the three platforms two methods have been used. The design of scatterplot diagrams and the estimation of the Spearman's rank test. The scatterplot method provides a visual impression of the correlation grid by grid. Scatterplot diagrams give a very visible image of the correlation between two projects. But they can give misleading information because there is not additional information associated to overlapping points. The visual analysis of scatterplots, is very helpful to determine outliers.

For the calculation of Spearman's rank test, the use of Cr is not needed since this statistical method ranks the values for every project and compares the ranks for every grid units in both values but does not consider the relative magnitudes. The following stage of the analytical process stresses proximity and topological relations among grid units calculating spatial autocorrelation. It finds out if a spatial distribution is clustered or dispersed. We used the global Moran's I. Following, to detect the existence of spatial clusters, we used the local Moran's I. More precisely to detect spatially statistically significant grid units, such as spatial clusters and outliers, was used Moran's I local indicator of spatial association (LISA; Anselin 1995) as implemented in ArcGIS 10.2.2 spatial statistic tools. This analysis which implies topological analysis gives as a result both clusters of grid units with high and low numbers of contributors and grid units that differed markedly from surrounding grid units. These two steps were also used by Raudsepp-Hearne et al. (2010), Casalegno et al. (2013), and Plieninger et al. (2013). Frequent assessments helped to verify if the results of the analysis are coherent with the unprocessed data and to verify the correct implementation of the methodology. Several cycles of analysis and assessment strengthened the methodology and gave precious insights on the characterisation of the projects and families of them. Several visual inspections helped in shaping the methodology improving the data preparation process to avoid type I statistical errors.

4.10 Cartographic representations

The current work is an exploration of the potential of CGI to support CES research. The maps here developed are used as a tool to support the evaluation of data. Therefore, their intended use is not to engage the wider public but to explore how CGI resources interact amongst them and how they are related to the urban and rural areas analysed.

Particularly challenging has been the definition of a method to allow the visualisation of multiple choropleth maps referring to individual platforms for each area. To be functional to the purpose of this study such maps must represent in the best way possible the distribution of data for every individual project but also help us to compare the mutual performances of the CGI projects. Therefore, the bins used to represent contributor ratio and volunteer productivity have been designed considering not only the individual map, but the series of maps designed for the same area and for the same family of projects. Therefore, the bins have been chosen to allow readers to see the spatial distribution of every individual project but also to compare them considering also how many grid units fallen in the same bin.

All distribution of contributor ratio is characterised by zero or near zero inflated values and by a long tail. We decided to collect mean median and standard deviation of all maps we wanted to compare and then we designed our bins around such numbers. We found out that a good representation of our data emerged with bins doubling their range at each stage.

In some maps we used a divergent scale going from red to green. We used it for the fact that the red used for small and marginal numbers are not overshadowed by green. We considered this as a fair use of the optic characteristics of these colours for two reasons.

Firstly, the maps are not intended for public consumption, so the untrained eye will not associate red and green to positive and negative judgments. Secondly, our spatial units are the BNG grids. They are purely geometrical abstractions, they are not designed around administrative or census units or any other delimitation that may associate a judgment with a population or a place.

4.11 Summary

This chapter introduced the main gaps in literature, identified the research questions and highlighted the methods and the resources used in this work. The gaps stress how there is a loose and episodic consolidation of methodology to assess CES. The gaps stressed also the limitation of the most popular research methodologies used for the evaluation of CES. Cost efficiency and length of engagement with local communities have been identified as ever recurring. In the gaps section there also been an assessment of the current use of CGI to detect CES. Inconsistent approaches, limited and sporadic use and overconfidence on one project only have been identified as limiting factors. The methodological part demonstrated how this work starting from nine different CGI projects developed a tailored methodology that can be applied to all of them to acquire standardised settings that allows a trans platform comparison.

5 Data sources and study areas

5.1 Introduction

In this chapter we will show in section 5.2 the sample areas, and introduce the crowdsourced geographic information projects used in section 0. The sections 5.4 to 5.6 introduce the individual families of CGI data and their specific features. A summary, in section 5.7, will recall the main points raised in this chapter.

5.2 Study areas

The three case studies have been developed mainly in two areas of the United Kingdom, south Wales, and London. South Wales has been chosen since it represents an area with a wide range of environments with none of them overwhelmingly predominant. Urban, peri urban, blue, and green areas with a National Park are closely knitted in south Wales.

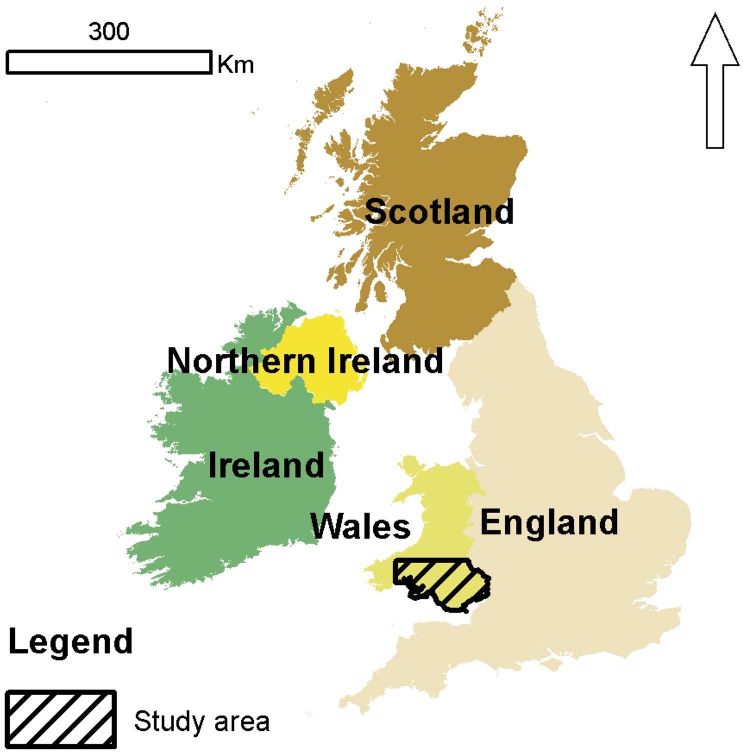

Figure 5-1: The first study area in the United Kingdom

The second study area is delimited by the administrative unit of Greater London. The large metropolitan area is a fascinating challenge for the detection of CES. While being mostly famous for its man-made artefacts, London hosts a large amount of blue, green, and natural areas. South Wales is the setting of two case studies. For the first case study we have chosen a study area of approximately 800,000 ha as shown in Figure 5-1. This area is characterised by diverse environments. The first study area covers, from north to south, the Brecon Beacons National Park in Powys to the coastline facing the Bristol Channel (Figure 5-2). West to east, it stretches from the delimitation of the

Pembrokeshire Coast National Park to the Wales-England border. Within the wider case-study area, there is a broad range of landscape habitats, according to Scott's (2002) classification. The southerly part of the study area is characterised by diffuse urbanisation, with the main urbanised areas around Cardiff, Newport, and Swansea. There is a linear urbanisation along roads that follow the natural valleys and watercourses, called *The Valleys*, toward the Brecon Beacons area from the coastline.

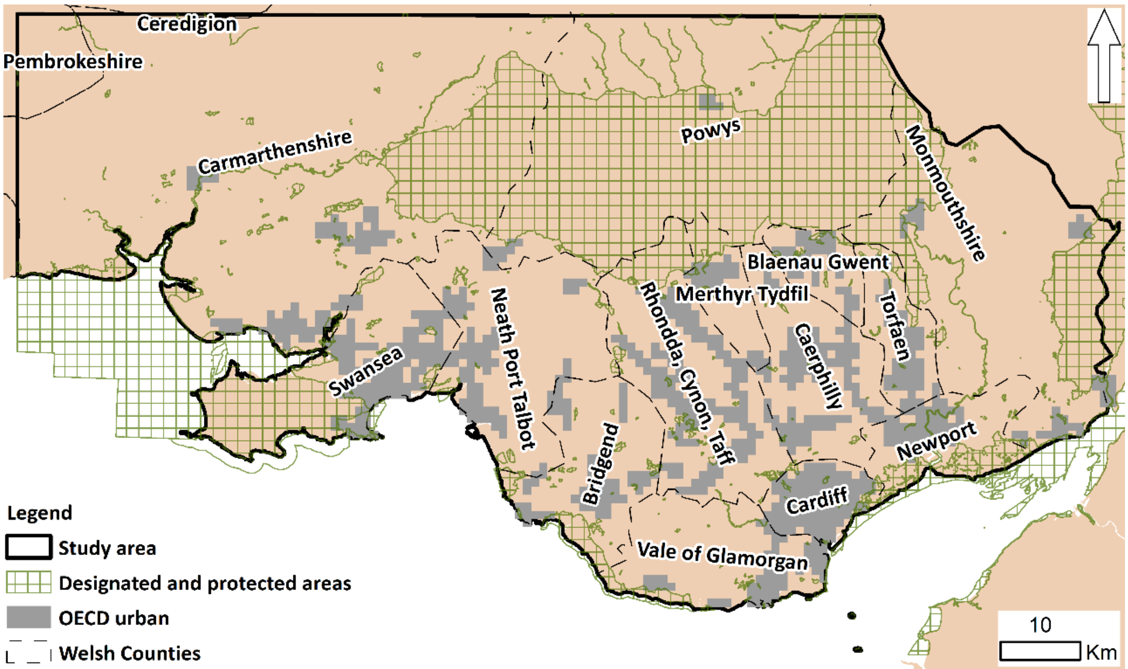

Figure 5-2: Urban areas and designations of the first study area in south Wales

According to the UK 2011 census, more than 2 million people live in this area, in 14 counties. There is a diverse mix of designations due to European directives and national and local legislation. These designations include a National Park, Sites of Special Scientific Interest (SSSIs), Special Areas of Conservation, Special Protection Areas (SPAs), Wetlands of International Importance, National Nature Reserves (NNRs), Marine Nature Reserves, Areas of Outstanding Natural Beauty (AONBs), Heritage Coasts, Biosphere Reserves, Biogenetic Reserves, and Local Nature Reserves. Descriptions of these designations are provided by the Joint Nature Conservation Committee (http://jncc.defra.gov.uk/page-1527).

The designated and protected areas are mainly concentrated in the northern part of the study area and along the coastline, where the intricacy of designations leads to overlaps, especially in the Gower Peninsula in the west, across the border between the counties of Swansea and Carmarthenshire. The delta of the River Loughor in the southwest and the surrounding areas have several overlapping designations. Hence, the study area offers a variety of environments for this study, with access to natural parks and natural reserves being relatively easy.

For the third case study we have chosen the easternmost half of the above described area (Figure 5-3). The presence of the full range of kinds of place, natural as well as cultural, but in a smaller area helped in the analysis of the third case study. Another criterion to resize the study area has been the richness of information, with more than one million data points at hand in this smaller area the third case study overcomes the first and equals the second case study in term of computation intensity.

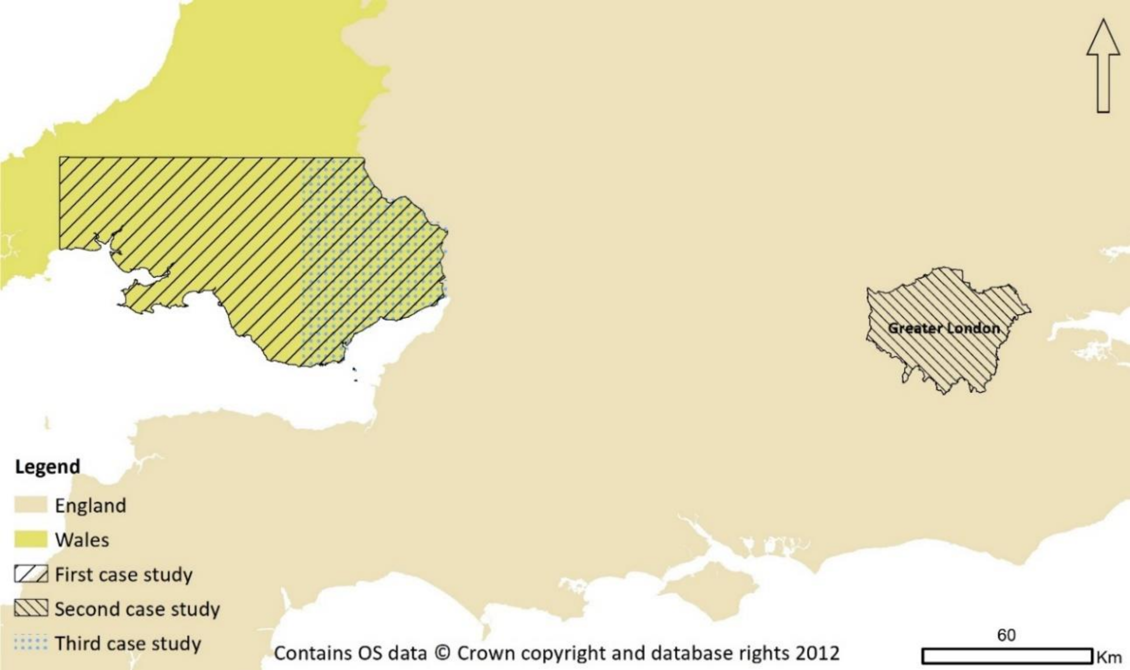

Figure 5-3: Three case study areas in southern United Kingdom

The second case study is set in Greater London, as shown in Figure 5-4, below.

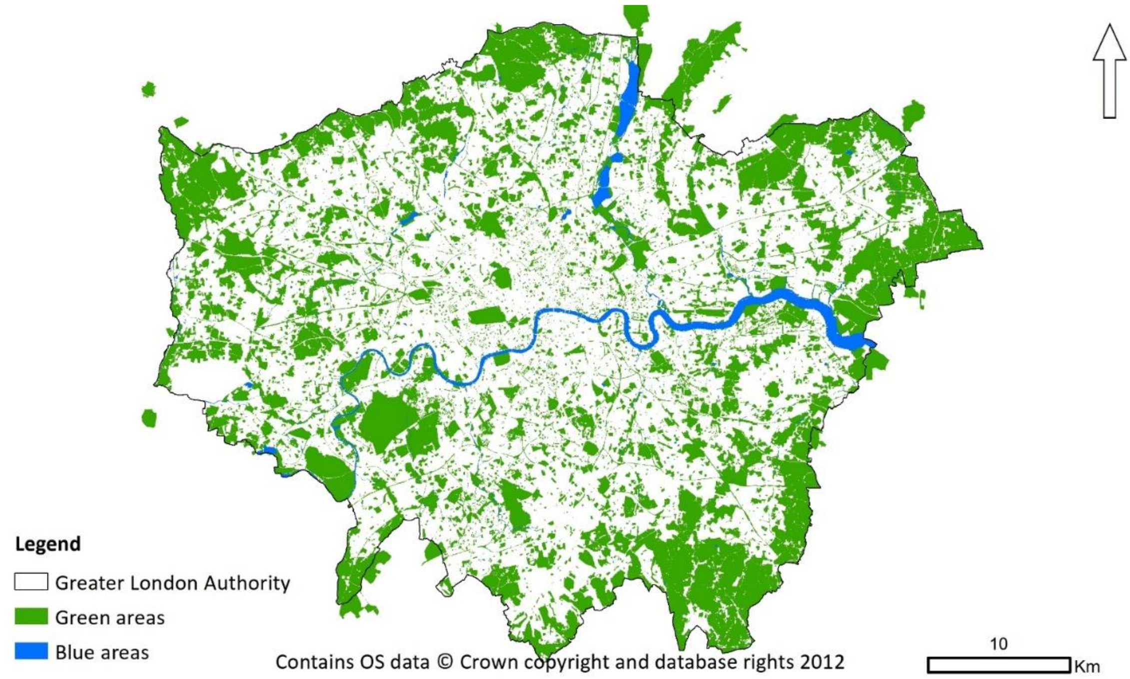

Figure 5-4: Green and Blue areas of Greater London

The administrative unit covers an area of 1,569 km^2 and has around 10 million inhabitants. 33% is vegetated greenspace with a further 14% estimated to be private vegetated garden green space; 2.5% is blue space such as rivers, canals, and reservoirs. Greater London contains 185 protected wildlife sites with 19% of its area designated as Sites of Importance for Nature Conservation (GiGL 2016). The analysis in Greater London focused on the use of citizen science data.

5.3 Datasets

In this section, we offer an introduction and description of the nine projects grouped in three families. The datasets, practices, and communities we have chosen gained different levels of interests in the scientific community. To give a picture of this in the following Table 5-1 we collected the mentions of some keywords used in the two prominent collections of papers focusing on CGI (Sui et al. 2013b, Capineri et al. 2016).

CGI keywords	**European Handbook of CGI (Capineri et al. 2016)**	**Crowdsourcing Geographic Knowledge (Sui et al. 2013b)**
Citizen science	130	121
Flickr	67	13
Geograph	12	0
GiGL	0	0
iNaturalist	3	0
Instagram	40	0
iRecord	0	0
iSpot	2	0
Mapmaker	2	5
OpenStreetMap	156	174
OSM	277	270
Panoramio	28	1
Pinterest	0	0
Twitter	223	39
Wikimapia	21	11
Wikipedia	32	145

Table 5-1: CGI projects mentions in main books

The locational information on pictures was collected in August 2015 using APIs and data dumps. Citizen science information has been collected in March 2014 while Incremental data has been collected in the summer of 2016.

5.4 Photosharing

The first group of projects used to elicit spatial preferences is made by Flickr, Panoramio, and Geograph. The three projects share similar platforms that consist in the facilitation of online photo sharing. Alongside the main platform those three projects have some community resources. Flickr and Geograph community forums are integrated in the main platform. Panoramio's platform was less structured but the community was quite active and maintained a quite vital forum on an external server and a mailing list. Flickr is the most developed with the possibility to upload pictures directly from mobile phones and is integrated in the Yahoo family of services. Panoramio is no more operational since November the 4th 2016. Google, decided to concentrate Panoramio and Map Maker on Google Maps, its main geospatial product. Google local guides is a new geospatial product based on crowdsourced reviews. The main CGI for Google is now concentrated in Local Guides. As seen in chapters two and three Flickr and Panoramio have already a consistent literature supporting their use for the investigation of landscapes. In particular they analyse tourist flows and behaviour or explicitly aim at the detection of cultural ecosystem services. As stressed above their analysis have often been decoupled or when used together their individual character has not been assessed (Tenerelli et al. 2016) in one case only (Hollenstein and Purves 2010) the characteristics of different photosharing projects has been attempted. Geograph provides a project with a very high degree of locality for the UK and has a very similar platform to Flickr and Panoramio while its management and its aims are different from the other two projects. In the following a description of the individual projects and at the end of this section Table 5-2 shows the common fields among the three projects.

Flickr[53] is an application created in 2004 with the purpose of photosharing and photo storage online. There is not coordination of the data collection, the collaboration pattern is therefore additional. The geographic component is not in its main purpose[54]. Therefore, Flickr can be considered spatially implicit. In the global ranking of websites Flickr occupies a position around the 250th. It is explicitly stated that Flickr is a photo management and sharing application. The geographic component is an added value to the sharing of pictures therefore some contributor can advertise a picture geolocating it in a place where it can attract curiosity regardless of the place where the picture has

[53] http://www.flickr.com/
[54] http://www.flickr.com/about/

been taken. The content of the picture can be absolutely not geographic or geolocated with a very low degree of accuracy[55]. Flickr allows comments, votes (likes only), and tags for every picture. Contributors must be registered. Flickr data are available through a public API[56]. The API is quite rich allowing to download several kinds of data not only about the pictures but also about the success of the pictures themselves in terms of likes, comments, and viewings. Since those specific elements are not available in the other platforms they have not been used in the present work. To access the data an API request includes also an app registration, for this work the app "Cultural Ecosystem Services detection" has been created. We developed specific PHP code to query the Flickr API. Special care must be devoted to the design and test of the query. A first data collection was affected by a bug in the data collection code therefore the number of Flickr images was largely underestimated (~5,000 vs ~445,000). This happened because there was a missing check where the number of pictures that can be harvested overcome the limit allowed for individual request. This work uses the Flickr API but is not endorsed or certified by Flickr.

Panoramio was an application explicitly intended for photosharing over landscapes. It was spatially explicit and additional. The service acquired by Google in 2007 was closed on the 4th November 2016. Some of the functions of Panoramio as a source of crowdsourced information were transferred to Google maps while all social media components alongside with all images that are not used have been removed from the web. Panoramio, around the date of its cessation was ranked around 3,500th in the Alexa global ranking of websites. Panoramio was the main provider of geolocated pictures overlaid on Google maps and Google Earth. The site allowed comments votes (favourite only) and tags. The fact that pictures inserted in Panoramio could be seen at a global scale overlaid on Google maps was a great incentive for participation. Contributors had to be registered. Panoramio data and pictures were available through a public API. The Panoramio API was handier in comparison to the Flickr API. It had a very limited number of parameters to set and query for. The query got a JSON formatted file.

The *Geograph Britain and Ireland*[57] project aims to collect geographically representative photographs and information for every square kilometre of Great Britain and Ireland. The scope in this case is explicitly geographic and the overall idea is to spread and cover all the territory of the United Kingdom. All the territory is viewed as in the tessellation of the BNG at different levels (1km and 100 meters) and contributors are encouraged to cover as more grid units as possible but also to deepen the knowledge of already photographed areas to add as more details as possible. Positions of pictures are then chosen not based

[55] There are entire photo albums regarding a touristic town geolocated in one point.
[56] http://www.flickr.com/services/api/
[57] https://www.geograph.org.uk/

only on aesthetic appreciation of the landscape but also on the need to cover all the territory. The project does not discourage multiple pictures of the same grid square. Therefore, Geograph does not fit perfectly in one of our collaboration patterns. It seems additional but it is also incremental aiming at augmenting information from photos for the atomic element of information which is the 100 m^2 grid unit or centisquare. Here there is a collaboration that is typical of incremental projects since contributors behave knowing what other contributors have done. It is incremental in an additional way for the centisquare element. Participants are incentivised with games and leaderboards. The website is ranked lower than 70,000 globally but retains a position around the 3,700th in the UK. Updated data are available as SQL or TSV dumps in a specific section of the project. Other formats are available, but data is not as frequently updated as the dumps mentioned above. The data collection process to import Geograph data used the UK wide data dump. Several APIs are claimed in the specific page[58] but at the time of this work the documentation was lastly updated year before, in May 2012. The safest way to gain possession of data has been evaluated as the acquisition of a data dump using a TSV formatted text file and using a database schema available in MySQL format. For the fulfilment of this work only one of the table of the database has been created. The TSV file contained all data covering the UK with information for 3,548,905 images. The table has been imported in a GIS and then the bounding box as identified in 4.7 has been used to geographically restrict the number of images to the final amount of 117,850 images.

A twin project has been developed recently to cover the territory of Germany. A project derived from Geograph is *Scenic or not*. It is a website that asks for a qualitative aesthetic evaluation over pictures that are taken from the Geograph project restricted only to Great Britain. The evaluation is achieved voting from 1 (not scenic) to 10 (very scenic). Pictures achieve a vote that is the average of the votes received. Information about pictures that have been voted more than three times are delivered through a TSV (Tab Separated Value) file. This resource once discontinued has been taken over by the Data science lab of the Warwick business school to analyse how sceniqueness impacts human well-being (Seresinhe et al. 2015, 2017). This resource alongside the other qualitative information associated with Flickr and the other projects of this work creates a good ground for further developments.

The main information downloadable from the three platforms has been compared with the following Table 5-2. There are the main components of CGI as described in chapter 3. The position in space given by the couple longitude and latitude, the possibility to identify the contributor that in some platforms has even a dedicated web page. There are

---

[58] http://www.geograph.org.uk/help/api accessed 03/01/2016

fields dedicated to the main content with the web address to access the pictures. Going further than the minimum elements we have also the timestamp and the possibility to associate a title to the pictures. Flickr and Geograph can provide more data. For the sake of the platform comparison the following table is restricted to the fields of Panoramio.

	Geograph	Flickr	Panoramio
Photo id	"GRIDIMAGE_"= 3251774	"photo id"= 7946078662	"photo_id"= 62701946
Contributor id	"USER_ID"= 1712	"owner"= 55898913@N04	"owner_id"= 1766081
Contributor nickname	"REALNAME"= Humphrey Bolton	"ownername"= interbeat	"owner_name"= Karen James
Contributor URL	https://www.geograph.org.uk/profile/1712		"owner_url"= http://www.panoramio.com/user/1766081
Picture URL	"https://www.geograph.org.uk/photo/3251774"	"url_o"=https://farm9.staticflickr.com/8181/7946078662_6861cfb343_o.jpg	"photo_url"= http://www.panoramio.com/photo/62701946
Photo title	"TITLE"= Rhossili Bay	"title"=Barmouth Bridge, Wales	"photo_title"= Brechfa Pool at Sunset
Timestamp	"IMAGETAKEN" =2012-12-02[59]	"datetaken"=2012-08-26 13:59:28	"upload_date"= 26 November 2011
latitude	"WGS84_LAT" =51.568803	"latitude"=52.719942	"latitude"= 52.030905
longitude	"WGS84_LONG"=-4.292971	"longitude"=-4.046819	"longitude"= -3.2862

Table 5-2: Fields comparison for photosharing platforms

[59] Geograph date and time are saved as a free text (Character varying) and no standardised and consistent format has been used

5.5 Citizen science biodiversity recordings

The second group or family of projects is made up by a selection of three citizen science biodiversity recording initiatives. The second group was chosen since the actions and patterns of volunteers reveal not only the location of biodiversity but also the spatial preferences of environmental focused volunteers. We are not going to use the OECD classification of 1km grid squares because the actions are explicitly targeted at elements of the environment. In this case volunteers often belong to associations or clubs. The analysis of their spatial behaviour can also lead to interesting insights on the relation between environmental associations and cultural ecosystem services in the metropolitan area of Greater London.

UK biodiversity is exceptionally well recorded, representing almost one-eighth of the records on the Global Biodiversity Information Facility (GBIF) (Powney and Isaac 2015). While all UK environmental data is available through the National Biodiversity Network (NBN) gateway, but surveys are not merged in a coherent way. At the beginning of this case study the download of all NBN data with a geographic query covering a 10km grid gave a data table with 570 columns but most of them were repeated (e.g. there was 16 times NBN_COMMENT). Therefore, a different approach has been followed aiming at more advanced ways of data management. This has been achieved through the collaboration with organisations of data collectors or local records centres.

This work compared three volunteer-contributed datasets from Greenspace information for Greater London CIC (GiGL), iSpot, and iRecord, encompassing a broad range of volunteer motivations and interests (Local Records Centres 2014, Pocock et al. 2015, Silvertown 2015). Those three projects while focusing on environmental recordings have quite different structures and functioning. iSpot and iRecord are the most similar. Both rely on IT and mobile services to collect environmental records from volunteers. Both provide social media services. iSpot is the one whose focus on the community is stronger, since it is designed and driven in a way to strengthen the community supporting it. To motivate a peer reviewed process iSpot also created a system based on layered roles and leaderboards. iRecord has tools to help people gather in real space. The possibility to organise activities or bioblitzes is therefore provided. GiGL has very limited recording tools but takes advantage of the long-lasting tradition of British environmental volunteering and citizen science. GiGL provides a very limited support to stimulate the physical or IT mediated gathering of volunteers. GiGL relies on hundreds of organisations and natural history societies therefore it has not the scope to promote community building. In the following sections we describe more in depth the three organisations that provided us data. They all provided us data dumps.

5.5.1GiGL

GiGL (http://www.gigl.org.uk/) Greenspace information for Greater London is London's environmental records centre. It collates wildlife observations via data sharing agreements with recorders and natural history organisations. GiGL hosts online forms to facilitate partners' data collection and a live form to collect species records that do not form part of any particular project (http://www.gigl.org.uk/submit-records/) although <0.1% of the GiGL data analysed here came from this latter source. GiGL's collated dataset was begun in 1996 and although it also includes historic records, 99% of data are from 1975 onwards. Records are contributed not just by volunteers but by professional bodies such as the Royal Parks. Only surveys for which GiGL was confident all records were collected in a voluntary capacity were included in this analysis. This may include volunteers who are 'professionals' in their day-job but are contributing their own personal data collected in a voluntary capacity on their day-off. The majority of GiGL's volunteers are affiliated with natural history groups (e.g. The London Bat Group, Barn Hill Conservation Group) who will often be targeting a taxon or local area. Every time they are georeferenced as the lower left corner of the corresponding precision level of BNG. Surveys are not coherent having four different precision levels. Moreover, there are often several observation points overlapping therefore there is not a correct visual representation on map. In GiGL there are often environmental records that are referred to the geometric centre point of a specific place (around 65,000 of them). Therefore, the geographic harmonisation and the quadrats count methodology helped to overcome such limitations. GiGL holds more than 3 million records coming both from professional surveys and from surveys performed by volunteers. To access only volunteered collected data GiGL filtered out almost two million observations and provided three files, (two points and one polygon) with a total of 1,102,746 observations.

GiGL data

GiGL data, the overwhelmingly richest collection of environmental observations we got, comes from 176 surveys. The integration work made by GiGL reduced the levels of complexity compared to the iRecord data. The overall dataset included 27 fields. Amongst them is the verification field with 5,388 determiners. The majority of the observations are verified (611,795 or 55.48% of all observations). GIS data came in one polygonal (with reference to the BNG) and two point files. The polygonal files included only precision 1,000 observations and relative polygonal areas. Point files included all observations with better precision levels. Amongst the other fields there was also the coordinates for the OSGB reference system (eastings and northings) but they were pointing at the lowest left corner of every grid. In geometric terms they were associated with the intersection of 4 grid units. Data has four different levels of precision as in Table 5-3 below the distribution of precision levels for GiGL data. It shows a good potential for more spatially fine-grained studies with almost 80% of observations recorded at 100m precision.

Precision	Observations	%
1m	90,182	8.18
10m	46,474	4.21
100m	878,113	79.63
1,000m	87,977	7.98
Total	1,102,746	100.00

Table 5-3:GiGL precision level and observations

We needed first to associate the points to the 1km BNG grid, then we merged all of them in one only 1km BNG grid file covering all Greater London and including all GiGL observations coming from volunteers. The further analysis has been conducted with a GIS.

GiGL surveys and contributors

The surveys included in the data provided from GiGL were not equally successful. The first 26 most successful (less than 15% of surveys) gathered almost 75% of all observations as well as the top 10 gathering more than 50%, as can be seen in Table 5-4, below, where LNHS stands for London Natural History Society

Pos.	Survey	Observations n.	Observations %	Cumulative %
1	LNHS Moth Records	170,438	15.46	15.46
2	Essex Field Club	126,605	11.48	26.94
3	LNHS Bird Records 2008	44,744	4.06	30.99
4	LNHS Bird Records 2009	42,266	3.83	34.83
5	LNHS Bird Records 2006	37,909	3.44	38.26
6	LNHS Bird Records 2005 part b	29,213	2.65	40.91
7	LNHS Butterfly Records 2011	29,000	2.63	43.54
8	LNHS Bird Records 2002	24,603	2.23	45.77
9	LNHS Bird Records 2001	23,811	2.16	47.93
10	LNHS Bird Records 2007	23,769	2.16	50.09

Table 5-4: Top ten most successful GiGL surveys

In terms of contributors in Table 5-5 we have the list of the top fifteen GiGL contributors and the surveys in which they contributed. In total we have 14,668 pairs survey/contributor. In bold the lines where the contributor is anonymous, not provided, or anonymised. The top 173 contributors/surveys couples (1,18% of the total) gathered more than 50% of observations.

Pos.	Survey	Volunteer	Observations	%
1	LNHS Moth Records	Rachel Terry	45,235	4.10
2	LNHS Moth Records	GV Geiger	31,043	2.82
3	Essex Field Club	P.R. Harvey	30,071	2.73
4	**Herts & Middlesex Butterfly Conservation**	**Volunteer - Butterfly Conservation**	**16,314**	**1.48**
5	**LBG Records 2010**	**London Bat Group Recorder**	**14,522**	**1.32**
6	LNHS Moth Records	David Howdon	12,986	1.18
7	**LWT Wildlife Garden Survey 1999**	**Member of Public**	**11,466**	**1.04**
8	Essex Field Club	C.W. Plant	11,227	1.02
9	**Barn Hill - various**	**Various - Barn Hill Conservation Group**	**10,553**	**0.96**
10	Essex Field Club	J.E.D. Milner	10,521	0.95
11	Essex Field Club	D.A. Smith	9,570	0.87
12	**Essex Field Club**	**RIS recorder**	**8,350**	**0.76**
13	**TRP Richmond Park Butterfly Group**	**Richmond Pk Butterfly Surveyor**	**7,954**	**0.72**
14	LNHS Moth Records	A. Culshaw	6,732	0.61
15	**House Sparrow Survey 2002**	**Member of Public**	**6,639**	**0.60**

Table 5-5:Top fifteen GiGL contributors and respective surveys.

Half of the observer-survey top couples are anonymous or collective. This must be considered in our evaluations. Some surveys use a collective name while others simply "Member of the public" when referring to volunteers. This creates virtual volunteers highly productive as can be seen in Table 5-6 where the second most productive volunteer is "Member of Public". In Table 5 6, we stressed in bold anonymous or collective accounts. The top ten recorders contributed with 231,499 observations, amongst them 79,795 or 34,48% come from collective accounts. Another characteristic of some of GiGL data collections is the presence of recordings with two or three people assumed as recorders[60].

Contributor	Observations
Rachel Terry	45,299
Member of Public	**38,950**
GV Geiger	31,043
P.R. Harvey	30,071
R.M. Callf	18,397
David Howdon	17,488
Volunteer - Butterfly Conservation	**16,314**
London Bat Group Recorder	**14,522**
C.W. Plant	12,011
Various - Barn Hill Conservation Group	**10,617**

Table 5-6: GiGL top 10 contributors, in bold anonymous or collective accounts

5.5.2iRecord

iRecord (www.brc.ac.uk/irecord) is an online data portal closely linked with a mobile app that allows observers to manage and share their wildlife records and thus is particularly useful for keener recorders and *panlisters* (those who record across all taxa). In the global ranking of websites iRecord occupies a position around the 620,045. In terms of website access, it is characterised by an overwhelming majority of accesses from the UK where it ranks around the 54,669th position and where 47,6% of its traffic is from the UK. It started collecting observations in mid-2012. Sightings are checked by automated rules and by experts. Contributors can submit wildlife sightings via a general form or forms tailored to particular projects or taxonomic groups. iRecord has links with many established national taxa recording schemes and regional co-ordinating Local Record Centres and is helping those communities move online. iRecord stores data from partner organisations who agree to share data, some of which may be collating data for specific goals or using specific methodologies, for example, the Garden Bioblitz (www.gardenbioblitz.org) an annual survey in which volunteers list the wildlife in their garden, or iRecord Ladybird (www.ladybird-survey.org) which has been widely publicised by the media. iRecord contains a subset of data from iSpot for verification purposes and so this was removed to avoid duplication. The geographic information associated with iRecord observations had two precision levels and was geo-referenced

[60] As an example, "A.R. Duckett" participating in several surveys and working in several occasions in association with "D. Johnson" who also partnered with someone else in several other occasions.

in a coherent way using references to the BNG, while using both OSGB and WGS84 references for longitude and latitude references. All the above are spread in 8 precision levels.

iRecord data

iRecord data came with several degrees of complexity since it involves 29 surveys. Some of the observations were flagged as already shared with iSpot or GiGL. The dataset included 35 fields. Amongst them is the verification field with 43 verifiers. Nonetheless most of the observations are not verified (10,570 or 81.94% of all observations). The geographic reference was a real challenge. We had two projections used. WGS 84 was used for 2,662 observations (25.7%) while OSGB was used for 7,696 observations (74.3%)[61]. In iRecord we also had 7 precision levels. In Table 5-7 the distribution of observations and the precision level associated to them. It is evident how ignoring all observations that have been recorded with a precision larger than 1,000 meters we are ignoring 129 observations that amounts to ~1% of the total observations. Precision 1,000 is the second largest group after precision 100.

Precision (m)	Count	%
100,000	7	0.05
10,000	110	0.85
2,000	12	0.09
1,000	3,441	26.68
100	3,559	27.59
50	2,662	20.64
10	2,693	20.88
1	415	3.22
Total	12,899	100

Table 5-7:iRecord precision levels of original dataset

Some observations with different projection and different precision levels belong to the same survey. In Table 5-8, below a limited sample of the multi-layered complexity of the geographic reference of the iRecord data we received. We colour coded the table to highlight the three degrees of complexity. The white rows represent surveys with unique projection and precision. The yellow rows represent surveys with just different precision in the same survey. The orange rows represent surveys with different references in the

[61] In essence all observations with precision level 50 have been also recorded as referred to WGS 84

same survey. The red rows represent surveys with different both projection and precision in the same survey.

Projection	Precision	Survey	Observation
OSGB	100	UK Ladybird Survey | BBC Breathing Place Ladybird Survey	31
OSGB	1000	Plantlife Wildflower Count | Wildflowers Count	1,448
OSGB	10	ORS | Othoptera Recording	3
OSGB	100	ORS | Othoptera Recording	12
OSGB	1000	ORS | Othoptera Recording	17
OSGB	100	NatureSpot | NatureSpot records	1
OSGB	1	Nature Locator | Plant Tracker	1
WGS84	50	Nature Locator | Plant Tracker	246
OSGB	1	National Plant Monitoring Scheme | Wildflower Survey	10
OSGB	1000	National Plant Monitoring Scheme | NPMS Paths	518
OSGB	1	National Plant Monitoring Scheme | Inventory Survey version 2	345
OSGB	1	Mammal Society | Mammal Atlas	3
WGS84	50	Mammal Society | Mammal Atlas	14
OSGB	100	Mammal Society | Mammal Atlas	236
OSGB	1000	Mammal Society | Mammal Atlas	4
WGS84	50	iRecord | Rinse	4

Table 5-8: iRecord geographic referencing variety colour coded. In yellow just different precision in the same survey, in orange different reference in the same survey, in red different precisions and geographic reference system in the same survey.

iRecord surveys and contributors

The surveys included in the data provided with iRecord data were not equally successful. The first most successful six surveys (less than 20% of them) gathered almost 75% of all observations and the first 11 surveys gathered more than 90% of all observations, as it can be seen in Table 5-9, below.

iRecord surveys	Observations	%	Cumulative %
iRecord | iRecord general data	1,655	15.98	15.98
Plantlife Wildflower Count | Wildflowers Count	1,477	14.26	30.24
iRecord | iRecord Ladybirds	1,440	13.90	44.14
iRecord | iRecord Butterflies	1,148	11.08	55.22
iRecord | Garden Bioblitz 2013	1,042	10.06	65.28
iRecord | iRecord Import	983	9.49	74.77
National Plant Monitoring Scheme | NPMS Paths	518	5.00	79.77
National Plant Monitoring Scheme | Inventory Survey version 2	345	3.33	83.1
Nature Locator | Plant Tracker	279	2.69	85.79
Mammal Society | Mammal Atlas	257	2.48	88.27
iRecord | Garden BioBlitz 2014	244	2.36	90.63

Table 5-9: Top 11 iRecord surveys, all observations are 10,358

In terms of contributors in Table 5-10 we have the list of the top ten iRecord contributors and the surveys in which they contributed. In bold the lines where the contributor is anonymous, not provided, or anonymised. The top 39 contributors/surveys gathered more than 50% of observations. We can also see that while iRecord is an app based biological recording initiative it has some records associated to couples of citizen scientists. This supports strongly the case for having access to the richest collections of data. We refer especially to nicknames or name/surname pairs associated to each observation.

Contributor	Survey	Observations
Harvey, Martin C.	iRecord | iRecord general data	486
shilland, ewan	iRecord | iRecord Butterflies	397
Richard Bullock	iRecord | iRecord Import	311
Anonymous, Anonymous	**iRecord | iRecord Ladybirds**	**292**
McAusland, Marion	iRecord | Garden Bioblitz 2013	277
Atkin, Paul	iRecord | iRecord Butterflies	267
	UK Ladybird Survey | BBC Breathing Place Ladybird Survey	**191**
Sylvia Hicks and Julian Hicks	National Plant Monitoring Scheme | Inventory Survey version 2	188
marcus jordan	iRecord | iRecord Import	186
Hastings Iqball, Ruth	Plantlife Wildflower Count | Wildflowers Count	171

Table 5-10: Top ten iRecord contributors and respective surveys

There are several anonymous or in any way undistinguishable or not individual contributors in the iRecord collection. Below in Table 5-11 some of them.

Contributor	Observations
Anonymous, Anonymous	296
	288
anon.	8
Environment Agency, EAinvasivesKSL	6
Bede Centre Nature Group	6
Unknown	3
Morden Hall Park Nature Group + Science Club	2+2
TOTAL	611

Table 5-11: iRecord anonymous or collective contributors

Table 5-11 contains just 611 observations, they amount to 5.9 % of iRecord records. They lower the precision of our contributor centred evaluations reducing the number of contributors that are counted for each grid units considering that most of them are concentrated in specific surveys.

5.5.3iSpot

iSpot (www.ispotnature.org) is a web application developed by the Open University to help anyone, from beginner to expert, with species identification. In the global ranking of websites iSpot occupies a position around the 326.165th. While firstly developed in the UK, a South Africa focused development is now providing 49% of the online visits while only 14.3% of visits are from the UK. Participants began contributing observations in mid-2009. Participants submit photographs to an online social network, purpose-designed to crowdsource the identification of organisms. Over 94% of observations submitted to iSpot receive a determination, verified at around 92% accuracy. Contributors do not need to be able to identify the species they have photographed although they are encouraged to suggest an identification. Other iSpot contributors will then confirm or correct the suggested identification. Here, in a similar manner to Geograph for photo collections, iSpot is a kind of exception amongst biological recording projects. Mainly it is not itself strictly a population survey. Encouraging collaboration over observations it is more characteristic of incremental than additional projects. For the current study iSpot provided a national data dump with 59,773 observations.

iSpot data and limitation

Before approaching the issue of geographic harmonisation, it is important to acknowledge one limitation that concerns specifically iSpot data. It makes a very interesting case for the of CGI as secondary data, with a different purpose from the one stated by the data custodian. iSpot rather than being a simple data collection project, a citizen science biodiversity recording project, is more focused on training volunteers to properly identify and classify nature. Therefore, the way the system is designed could lead to the fact that most of the contributors may never have a position associated to them and even more interestingly the incremental process is not built around the place of the observation but around the observation. Contributors can even simply participate in the identification process without contributing with sightings. The organisation, provided for our analysis the iSpot data that met two quality criteria for the identification and classification of taxa. From the documentation associated with data we received we had also the quality criteria that has been used to compile the data collection:

> "The data that is included in the download is from all observations that: A: have a "Likely ID" (i.e. the suggested identification is the one that has received the greatest amount of agreements by people with the highest reputation scores, and is above a threshold in our reputation system) B: AND have been identified with a name that matches a name in the UK species inventory (or Catalogue of Life species

dictionary for overseas observations). Observations with species names that have been misspelled, or correct synonyms that are not in the NBN dictionary, are not included in the data download."

We can argue that criteria A somehow incorporates the incremental process in the datasets, but contrarily to the other citizen science projects they provided us only data 100% verified, while criteria B is too restrictive to spontaneous and less trained contributions. Therefore, all the observations that have not met those two criteria have not been provided to us. While this can be very helpful for the studies on biodiversity it is detrimental for our approach. It reduced the amount of observations and locations that volunteers deemed suitable to practice citizen science biodiversity recordings and the incremental process of collaboration has been eradicated from the data. Willing to provide us good quality biodiversity recordings they provided us a stripped and poorer version of iSpot data. They filtered data using quality criteria unfit for our purpose. In our case we can claim that while iSpot is an incremental project, its organisation disguised it for our study as additional. The resulting datasets contains 13.352 observations that has been labelled as situated in Greater London. Names of recorders were not included but they have been assigned a unique ObserverID. In terms of geographic references, iSpot observations are associated to two precision levels[62] of the BNG. Latitude and longitude were not provided. Since our work is conducted at the 1km level we needed to aggregate all observations at the 1km level.

5.6 Incremental projects

The third category of projects is made by incremental projects. They collect information and convey it in IT formats for web sharing. Here the interest of people for places can be achieved through the number of creators and editors of geographic information. In the following, there will be the introduction to the incremental projects used to elicit spatial preferences. They are: Wikipedia, OpenStreetMap, and Wikimapia.

The main purpose of Wikipedia is the creation of a crowdsourced collection of human knowledge. This implies that in Wikipedia the geographic component is often absent or treated as an ancillary component. In this respect, it is similar to Flickr whose main purpose is the storage and sharing of images whose geographic component is an ancillary feature. Therefore, Wikipedia is not a spatially explicit project.

OpenStreetMap and Wikimapia have been created as projects with tools to create geographic information features. They are therefore spatially explicit, and we call them Collaborative Mapping projects. Wikipedia and OpenStreetMap are not commercial open source projects and are run by two foundations that have been created by contributors. On the contrary Wikimapia is still centralised by its creators and is closed source.

[62]10,293 observations recorded with 100m (6-figure) precision while 3,059 observations recorded with 1 km (4-figure)

Alongside the main platform those three projects have some community resources. The open sourced platforms have seen a rapid growth of forums, mailing lists, online chat pages often relying on external services. Wikimapia holds a centralised forum and system where the community can orderly interact with the owners.

5.6.1Wikipedia

Wikipedia is the most successful non-commercial crowdsourced project. In the Alexa world raking of websites is the 7th most popular and the first non-for-profit[63]. Its purpose is to create the free encyclopaedia online through the contribute of people. After an initial phase, where everyone could contribute everywhere (Viégas et al. 2007, Wilkinson and Huberman 2007, Kittur and Kraut 2008) and where there was a direct relation between number of contributors and quality of articles, a certain level of review has now been included. Then the contributing activity became increasingly complex and it may have caused the decrease in participation (Suh et al. 2009, Simonite 2013). The main quality concept is related to verifiability (Wikipedia contributors 2011, 2017a) so its policy now tends not to rely only on a peer review process but on a verification one. This means that individual knowledge must be omitted unless verification is possible. This limits the possibility to report knowledge that is previously undocumented or unpublished[64]. This therefore enforces the bias in favour of the *Global North* (Graham et al. 2014) as it also happens for gazetteers (Graham and De Sabbata 2015).

Wikipedia includes more than 5 million articles with more than 29 million of registered editors while just around 120 thousand are actively contributing every month. The project collects knowledge in a broader term than just data. The creation of information is then not focused on the creation of geographic information only. Therefore, Wikipedia is a spatially implicit project. The possibility to add geographic information in terms of geodata usable for information systems is therefore not a main purpose and this information element is not widespread. Moreover, articles are referred with just a couple of coordinates using a point in space leaving the same geometric representation for large aerial units like a nation and smaller ones like monuments.

A process to integrate Wikipedia articles with OpenStreetMap is ongoing (Wikipedia contributors 2017b). Given its open nature, Wikipedia attracted IT enthusiasts that contribute with side projects, tools, and improvements. Therefore, there are several options in term of access, integration with semantic technologies and the use of Wikipedia itself as a rich information base. One semantic adaptation of Wikipedia called

[63] List of most popular websites. (2016.). Retrieved November 10, 2016, from https://en.wikipedia.org/wiki/List_of_most_popular_websites

[64] Interesting to document the following discussion on the OSM mailing list that also stresses how the two projects are so different in regards of local knowledge vs verifiability. https://lists.openstreetmap.org/pipermail/talk/2016-November/077096.html

DBpedia is at the core of the Linking Open Data (LOD) initiative (Auer et al. 2008, Bizer et al. 2009). The latter can lead to very interesting developments for the current work.

5.6.2OpenStreetMap

OpenStreetMap or OSM (https://www.openstreetmap.org) is a highly popular Collaborative Mapping project where contributors explicitly gather geographic information. Started in 2004 in UCL it went through an ever-growing success. In the Alexa global ranking of websites, it currently occupies a position around the 5,000th [65]. Two opposing visions on the role of OSM are animating a constant debate in the supporting community. The Wikipedia definition" OpenStreetMap (OSM) is a collaborative project to create a free editable map of the world." (Wikipedia contributors 2017c) focuses on the visual aspect that often motivates people to participate in a similar manner to participants to Panoramio, the possibility to contribute to a map that can be seen and used worldwide. While a data centric definition is the official and always proposed by the project itself "OpenStreetMap is built by a community of mappers that contribute and maintain data about roads, trails, cafés, railway stations, and much more, all over the world." (OSM contributors 2015). Both visions stress more the value of local knowledge, but the data centric vision tries to create overall standardised global taxonomies to include all localities. OpenStreetMap is also the most cited CGI project overall both in the European handbook of crowdsourced geographic information (Capineri et al. 2016) as well as in the more cited on Crowdsourced geographic knowledge (Sui et al. 2013b) as previously mentioned in Table 5-1.

In terms of community OSM has more than 3 million contributors registered[66] while the number of active contributors per month, the ones contributing to the database, is now a bit over 1% of the total number of accounts[67]. Similarly to Wikipedia, OSM gathered an enthusiastic community of IT experts and passionate that started developing every sort of tool to facilitate every stage of the data collection, storage, and visualisation of geographic information as well as the networking capabilities to keep the community together and to synchronise the efforts. The project has a sort of central coordination through a list of recognised standard tag system. While the purpose of the project is focused on the production of data the use of standardised tags can produce a visualisation of the contributor's work on the standard online map.

The visualisation on the online global map is a good motivational factor to contribute for volunteers not interested in the use of data. OSM has always been characterised by a degree of real world sociality and activity. Mapping parties and conferences along with

---

[65] On 13/11/2016 is 5,122 http://www.alexa.com/siteinfo/openstreetmap.org accessed 13/11/2016

[66] OpenStreetMap stats report run at 2016-11-04 00:00:08 +0000 http://www.openstreetmap.org/stats/data_stats.html accessed 13/11/2016

[67] http://wiki.openstreetmap.org/wiki/Stats#Contributor_statistics_reports

the need to gather real world “mundane” information has always been preferred to the collection of information from aerial pictures or from imports of data. Being the most open Collaborative Mapping project, every component is accessible for analysis. It made OSM also quite successful in term of academic research as can be seen on the target projects of most of the references mentioned for this family of projects[68].

Geographic information in OSM

OSM is structured to collect geographic information through three main primitives, nodes, ways, and relations. Nodes are the basic geographic primitive, ways are essentially sequences of nodes and relations are associations between primitives. The present work focused on the analysis of contributors to nodes. Ways and relations are not here used as geographic elements. The tags associated to ways have only been used to filter out nodes coming from imports or fully automated edits. Ways and relations are first and second levels of combination of nodes that keeps being the foundations of every entity in OSM. Every primitive is composed by three main parts of data. The locational data, the metadata on the last modification and the tags. OSM also stores lineage data, in this community called *historical data*, that includes all developments of its elements. OSM allows historical data to be downloaded through data dumps. Hence it is possible to reconstruct all the development of the database in time. The main unit of change of the project in time is the changeset. They are created once there is an individual upload of modified data. In the following a brief description of the information associated to nodes in OSM. Nodes represent points in space, they are the most numerous element of OSM. As of 11/04/2016 OSM included 3,592,183,347 nodes. They represent either punctual one-dimensional information or are the vertices of linear segments. The information conveyed with nodes consists of identification, temporal, positional and attribute information as in the node below coming from a file structured according the XML-OSM format.


```
<node id="288309905" visible="true" version="2" changeset="12838221" timestamp="2012-08-23T21:08:18Z" user="DyfalDonc" uid="696494" lat="51.4003803" lon="-3.1983722">
  <tag k="amenity" v="restaurant"/>
  <tag k="name" v="Captain's Wife"/>
 </node>
```


Figure 5-5: Example of a node element in OpenStreetMap

The unique identification of a node is given by the couple "node id" plus some information of the last change to whom the node has been subjected to. Therefore, either version, or changeset, or timestamp can play this role. In the example above the node has version two (version="2") meaning that the information that is shown in the map is no more the

[68] A scholar search on 13/11/2016 gave 750 outcomes with the term “OpenStreetMap” in the tile field

one with which this node has been created. The information displayed is associated with the first modification that occurred to the node with id 288309905. The information about the contributor can be achieved through the element *uid* (abbreviation for user id) or *user* that displays its chosen nickname. The use of the id for the contributor is a safer way to access contributor information given that the nickname can be changed while user ids not. Positional information is attached through the *lat* and *lon* fields (latitude and longitude) that are referred to the WGS84. Lastly there are the tags that associate some qualitative information to the specific node. Tags have the form of *k*, *v* pairs. K and v stand for key and value. Nodes that are created with the only purpose of supporting ways have not tags associated. For the current study is it important to stress two elements. The value of “source” tags and the modality of versions creation. Tags provide rich information on the origin of data. The tags whose k includes the word “source” tell us that the data or information pertaining to such primitive has been deducted from specific sources.

This is valuable to understand if the information has been provided by volunteers or by automated processes. Versions are created in a way that a change in the tagging of a way or a relation does not change the version of one supporting node. Similarly, the change in the position of a node that is supporting a way or a relation does not force the linked ways or relation to have a new version, the way will refer to the new node since the way is not linked to the version of the supporting node but only to its id.

On the contrary, the deletion of a node will force a version change on all elements linked to it. The possibility offered by OSM to download the historical dataset and the possibility to distinguish between edits originated by human actions and edits made by automated actions will help us in the definition of a strategy to derive people’s interest for places through the analysis of the *attractivity* contributor index as defined by Arsanjani et al. (2015a). The need to analyse all historical data implied the download and parsing of two different data dump files. The first recording all information stored since the beginning of the projects up to the licence change in 2014[69]. A second file to update the older one has been downloaded and parsed. Only nodes have been involved in the analytical phase. Tags from ways has been used only to filter out nodes created to sustain ways that have been originated through automated and unsupervised edits.

5.6.3Wikimapia

Wikimapia is the most popular Collaborative Mapping project whose data is available to use and download online. While it is a commercial project it allows some of its data to be downloaded[70]. Started in 2006 it aims at the creation an open content map annotating

[69] This is needed since with the licence change some not more compliant edits might be obfuscated in the most recent planet file

[70] Almost all data can be downloaded with the exception of road networks

georeferenced images with polygonal entities and textual information. The fact that this kind of information can be considered as derived from the aerial pictures used in the background may lead to the problematic use of the data collected. It can be considered as a work derived from the aerial pictures. Every entity has a specific page where it can be described using free text, tags, and links. Most of the time there is a very brief description and the link to the Wikipedia webpage. Very often Wikimapia plays the role of the link between the aerial images and local situated knowledge that received a poor geographic coverage in Wikipedia. At the beginning it underwent several problems due to data vandalism (Ballatore 2014). The project had to introduce contributor registration and a hierarchy of contributors to control contributors' behaviour.

In the Alexa ranking for websites it occupies a position around the 3,000[71]. It is widely popular in Russia (its creators are Russian) and India. The website offers a kind of coordination also through hierarchical organisation and created a loosely structured category list enhanced by the community. The project also offers a centralised forum. All features are centralised so there has not been a blossoming of side projects and tools. Its commercial aim and its not open platform does not promote the project amongst IT passionate. Currently there are almost 2,5 million contributors registered and the number of places recorded overcame 25 million in 2016[72]. but the activity and the number of objects added is slowly declining since 2011[73]. Wikimapia allows contributors to create mainly places that are areas. Contributors can associate tags to places and add a free text description and media content. Since 2009 contributors are also encouraged to add transportation data with roads, railways, ferries, and rivers. The public API allows access only to place data. Data has been downloaded through the API (http://wikimapia.org/api) while lineage/historical information that is visible but not downloadable has been accessed developing a web scraping code as mentioned in 4.7. The downloaded data includes both a centre point for the feature and all vertices for the polygonal entities. Given that historical data is associated to the polygonal entity and not to every individual vertex the analysis considered the position of the centroid only. The use of the 1km BNG grid help us to overcome this limitation.

5.7 Summary

This chapter described the main resources used in this work and the areas where the analysis has been performed. The study areas range from the large urban/cosmopolitan area of Greater London to the wide variety of human settlements and natural and cultural environments that characterise south east Wales. In terms of projects we have here

[71] 3,793 on 13/11/2016 http://www.alexa.com/siteinfo/wikimapia.org
[72] http://wikimapia.org/stats/action_stats/?fstat=2&period=2&year=2016&month=11
[73] http://wikimapia.org/stats/action_stats/?fstat=101&period=3&year=2016&month=11

described them and the challenges associated with their use. A data preparation process is needed for every project or family of projects. This ensures the removal of automated edits and the accounting for anonymous or undistinguishable contributors. This chapter also highlighted how the projects stand in the classifications defined in chapter three. To summarise this classification, we can show the two-dimensional grid of Table 5-12. We colour-coded the table assigning the blue colour to the photosharing projects, the green colour to the citizen science biodiversity recordings and the grey colour for the purely incremental projects. We also used the italics character for the two projects that are a sort of hybridisation between additional and incremental.

	Additional collaboration pattern	**Incremental collaboration pattern**
Spatially Explicit	Panoramio, iRecord, GiGL	*Geograph, iSpot,* OSM, Wikimapia
Spatially Implicit	Flickr	Wikipedia

Table 5-12: Classification grid for CGI projects used in this work

In the following chapters, we will show how the resources and methods shown in this and the previous chapter can be used to support CES focused research.

Results and analysis part

This part of the thesis is composed of three main chapters. Each chapter shows the analysis and the results of the three distinct types (or ‘family’) of crowdsourced geographic information. Every chapter introduces the analysis of the area and datasets as given at the beginning of the data preparation phases that have been described in the previous chapters.

Given to the heterogeneous nature of the data used and the multiplicity of data sources, while there will be an overall parallel sequence of phases in some cases some phases are reduced or not existing. When needed, there is an explanation of the data preparation phase. It consisted in the processes implemented to remove data inconsistent with the purpose of this work.

Following there is the description of the geographic harmonisation.

Next, a description of the different analytical steps with the individual grid cell comparison, the Moran’s I calculation and the local Moran’s I calculation are provided.

The last part of each chapter includes the visual inspection part comparing the spatial statistical outcomes with the territory.

Chapter 6 is devoted to the specific analysis performed over the photosharing projects in south Wales.

Chapter 7 is devoted to the description of the analysis performed using citizen science data in Greater London.

Finally, Chapter 8 shows the analysis performed using incremental projects in south east Wales.

6 Photosharing projects

The baseline situation for this analysis is covered in the following sections. First there is section 6.1 where we introduce the area where we set the first and third case study from a cultural point of view. Following section 6.1 focuses on the study area in demographic terms. Then section 6.3 shows the basic global statistics referring to the platforms. Thirdly, section 6.4 shows the distribution of contributors to the individual platforms. Following the analysis section with its first section (6.5) devoted to the differentiation between urban and non-urban areas.

The spatial analytical section (6.6) begins with a grid by grid analysis and is performed through statistical correlation analysis and scatterplots in section 6.6.1. The local spatial statistic part of this chapter (section 6.6.2) begins with the computation of local Moran’s I and the creation of maps with hotspots and outliers in section 6.6.3. To crown this analysis, section 6.6.4 describes the visual inspection on the characteristics of the places that are highlighted by the spatial statistical analysis.

The final part of the chapter is devoted to a side work that was performed using the same kind of data. Section 6.7 shows a study conducted in Greater London using photosharing but with a limited scope and less data sources involved. The chapter is concluded in section 6.8.

6.1 South Wales from a Cultural point of view

To ground our analytical work on the case study area we introduce the cultural component of the benefits humans receive from the ecosystems in Wales. We are not only aiming at benefits for resident people. Tourism based on long walks and sceneries provide benefits also in monetary terms for the local economy, for the physical and mental wellbeing. Those benefits are enjoyed by Welsh people and by tourists.

Seen from an exclusively monetary term the UK NEA reports a 3% figure for the national output that can be connected to wildlife-based activities (Russell et al. 2011). Moreover, it reported that 50% of tourist spending in Wales is by day trip visitors, 32.5% by longer-stay UK holidaymakers, 8% by international visitors and 6% by business tourists.

Those beneficial effects are spread on almost all Welsh territory with the two extremes of uplands and coastal margins mostly characterised by a powerful mix of landscapes and sceneries. Overall Wales is well known for its landscapes with 24% of its surface listed as AONB, Areas of Outstanding Natural Beauty. Conversely, the coastal ecosystems are regarded as under threat from tourism while having Wales a high percentage of EU Natura 2000 designated areas. The costal landscapes include impressive Sand Dune systems and spectacular Sea Cliffs. The intensification of Woodland cover started in all Wales from the 1990's and now reached 14% of land area.

This also in response to the need to reclaim and convert the areas that has been hugely scarred by the industrial extraction of coal. Between the two extremes we have the peculiar 'valleys' environment. The land classified as good agricultural soil is a rarity, therefore the opportunities offered from the sceneries and local identities are a really relevant component for local development.

The peculiarity of the landscapes and the linguistic peculiarity which unfolds with Welsh toponyms is a strong contributor to the Welsh national and local identities. The creation of the nation-state rhetoric in the 19th century often stressed the links between local nature, local language, local practice, and landscapes. Wales has plenty of components to flag its identity.

Summing up, Wales as a whole is in general characterised by three elements who falls inside the remit of CES:

1. Visual sceneries
2. Cultural identity
3. Sustainable and healthy tourism

Those three elements are inextricably interwoven and the emergence of one of them is often associated if not based on the underlying presence of the others (Church et al. 2011). This characterises all the UK but is heightened in Wales.

The co-evolutionary aspects of nature and society are witnessed by long-lasting remnants of human-nature interactions that have also symbolic and religious meanings. The localisation of religious cultural heritage has always been connected to natural areas which has always a symbolic value in cultural practices. This is also witnessed by several remnants of the Cistercian legacy. In recent years has been designed the Cistercian Way[74], which is a long-distance footpath linking sites associated with Cistercian monasticism in Wales. This suggests how the eminently interactive relations between the components of CES as proposed in the UK NEA is well fitting our study.

In our study area we see all the elements listed above.

In Figure (5-2) we have the uplands mostly in the county of Powys and Carmathshire there are elements of the inner plateau and their incisions, are included in the Breacon beacon National park which includes also AONBs and other designation for biodiversity conservation. Similarly, the southernmost margin of our study area coincides with the coastal margins and marine areas. There are several areas which are subjected to designations and protection for visual and natural reasons. For visual reasons there are AONBs (Gower peninsula in Swansea County), for nature conservation there are several EU Natura 2000 sites along the coastline up to the Severn estuary.

[74] http://www.cistercianway.wales/

In between those two areas we have the former South Wales coalfield which covered most of the Valleys. This landscape stretches from the uplands towards the sea up until the latitude of Newport and Swansea. This is a soil not favourable for agricultural use and in the past centuries has provided coal mines that shaped the landscape, local cultures, and support for the industrial development. Once the mines have been closed this left a landscape with multiple signs and in need of remediation. This also shaped local identities and the heritage of it is witnessed also by the initiative of the Valleys Regional Park, an informal network of amenities, open air activities, and recreation opportunities connected by a wide network of rights of way.

The rest of the territory with the Vale of Glamorgan and Monmouthshire the only areas with favourable land for agriculture and Carmarthenshire mostly characterised by Grassland and Marshes which provide the basis for local produce derived from grazing. The local identities we can find in the area are also characterised by an inextricable link between natural heritage history and cultural heritage. In our study area we have also 5 stations of the Cistercian Way.

6.2 Resident population and distribution of contributors

This work used spatial units that were not designed around population or social thresholds: an isotropic tessellation based on 7,255 1km BNG units. For the detection of environmental spaces, we examined the relationship between the number of contributors that took and posted pictures in the same non-urban grid unit. The study area has an average of 276 inhabitants/km^2. The OECD classification used found only 925 urban grid units (~14% of the total). The analysis focused then more in detail 6,330 non-urban grid units found following the procedure described in section 4.5.2. Proximity and accessibility characterise CES, as does the movement of the receptor toward the area (Costanza 2008). If people did not move to pursue cultural enjoyment, this work would have found only pictures shared in the urban grid units, while in fact this study found that people share pictures from non-urban areas. Therefore, there is an indication of the phenomenon and a measure of its magnitude for specific areas.

6.3 Basic global statistics of photosharing projects

From this point onward, unless otherwise specified, the statistics and maps shown refer to the data aggregated to 1 km^2 of the British National Grid. Considering all projects as they unfold in the selected study area, some basic statistics to compare them have been extracted in Table 6-1, below.

	Flickr	Panoramio	Geograph	All projects
Maximum images per km^2	**29,473**	1,833	1,058	32,241
Total individual images	**444,399**	49,966	117,850	612,215
Percentage over total images	**72.59%**	8.16%	19.25%	100.00%
Grid squares with images	4,330	3,533	**7,175**	7,201
Coverage over total grids	59.68%	48.70%	**98.90%**	99.26%
Total contributors	**9,661**	4,725	1,117	15,503
Percentage over total contributors	**62.32%**	30.48%	7.21%	100.00%
Maximum images per contributor	11,574	2,329	**43,480**	----
Average images per contributor	46	10.57	**105.5**	39.49
Average contributors per km^2	1.33	0.65	0.15	2.14

Table 6-1:Basic statistics of photosharing projects in case study area (7,255 grid units). In bold noteworthy values.

The total amount of pictures is strongly influenced by the Flickr community, which stood for almost 73% of photos (444,399 photos) and 62% of contributors (9,661 contributors). From a geographical point of view, Geograph covers almost 99% of all the BNG grid units in the area compared to 60% and 49% for Flickr and Panoramio respectively. Geograph's incentives that promote spatial coverage also influences contributor productivity, with an average contributor providing more than 105 pictures compared to the 46 pictures supplied by Flickr's more numerous but less productive contributors. The most prolific contributor to the Geograph platform (43,480 pictures) provided 30% of the overall platform, which equalled 87% of all contributors to Panoramio (49,966 pictures). The behaviour of Geograph contributors, with its tendency to spread more spatially, stresses the fact that the desire to gain points together with the desire to share photos of places worth seeing, drives the motivation of contributors. The behaviour of Panoramio contributors is often included in the range defined by Flickr and Geograph. The proportions of pictures taken in non-urban grid units were 51% for Panoramio, 55% for Geograph, and just 31% for Flickr, yet 87% of grid units in the study area are classified non-urban. In general terms, Geograph, with its successful campaign to cover the whole territory, reflects the study area more closely in terms of spatial coverage and accessibility, but follows Flickr in terms of the quantitative production of pictures. High productivity on the part of some individual contributors can skew the data in terms of the number of pictures and geographic preferences. Therefore, a statistic based on the number of contributors is more suitable for our study as stressed previously.

6.4 Spatial patterns of individual projects

To start the analysis, we first look at spatial patterns of the individual projects. The different projects have quite different number of contributors. Therefore, maps depicting the total numbers of contributors for each project while interesting can be less useful in the development of this work that aims to compare the projects. To map contributions, we derived grid by grid the grid contributor ratio (Cr). This is the percentage of contributors for each grid unit compared with the total number of contributors to the specific project in the study area. As an example, in Flickr there are 9,661 contributors in the study area while in the BNG grid ST1575 in Cardiff there are 20 contributors. The Flickr contributor ratio for grid ST1575 is hence Cr=20*100/9,661. Therefore, the Flickr contributor ratio Cr for the BNG grid unit ST1575 is 0.21.

Panoramio's contributors, in Figure 6-1, spread their effort quite uniformly covering with 8% of pictures almost 49% of the grid units in the area. Although this uniformity is an evidence for movements in space it never reaches interesting peaks obtaining its largest values in the urban area of Cardiff and some other points along the coastline. The highly designated area in the delta of the Loughor is uniformly covered. The spatial analysis reveals in section 6.6.3 the existence of spatial patterns in the area.

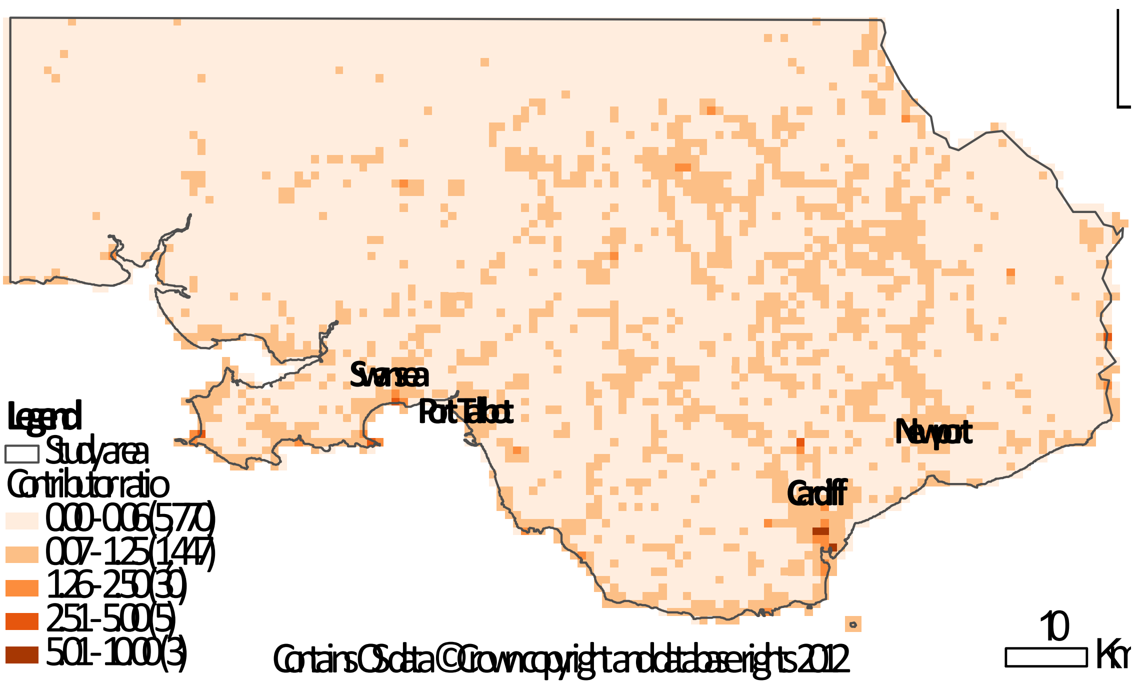

Figure 6-1:Spatial distribution of Contributor ratio for Panoramio in south Wales

Flickr contributors spread a bit more in space and post a lot more pictures. In this case in Figure 6-2 hotspots and trends are easier to detect. The trend along the coastline towards west and some hotspots are here more noticeable. Apart from the expected hotspots in urban areas there is the tendency to cover almost entirely the coastline towards the open sea while the highly designated area in the delta of the Loughor, received relatively less attention.

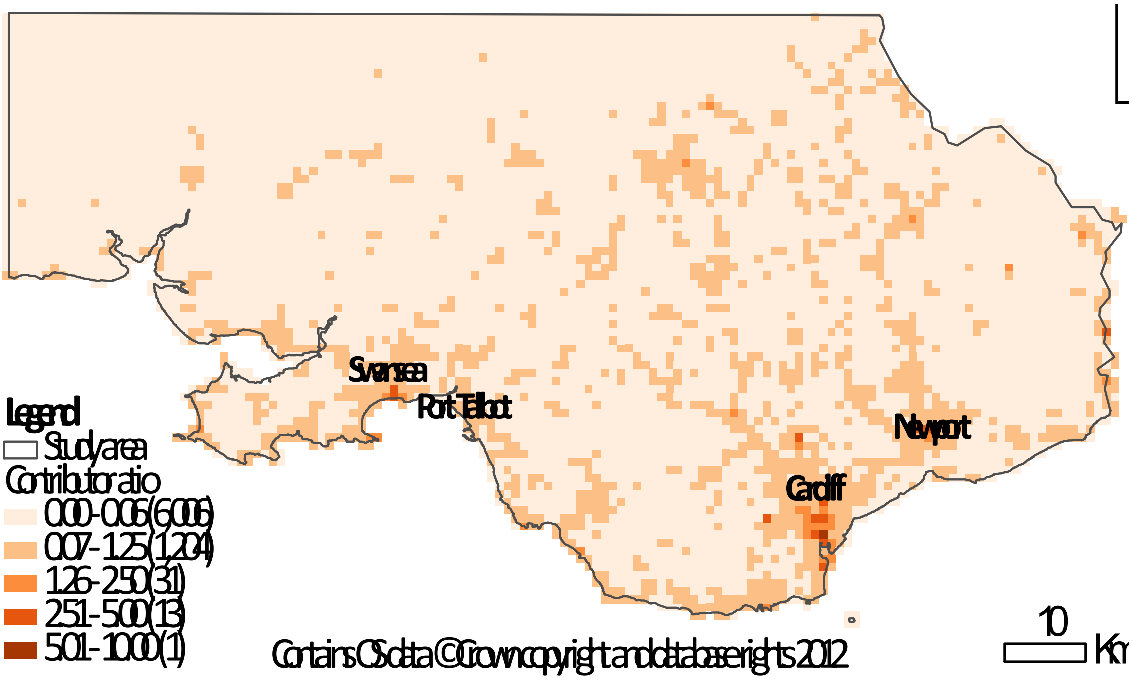

Figure 6-2:Spatial distribution of Contributor ratio for Flickr in south Wales

A small cluster can be found in proximity of the highest peak of the Brecon Beacons, the Pen y Fan. On the west of the study area a small cluster has been found in correspondence of the National Botanic Garden of Wales. In the central part of the study area there is an area of pictures from five different hikers describing their path while penetrating towards the core of the Brecon Beacons National Park. The Geograph project is quite richer than Panoramio, it is characterised not only by a larger number of pictures but also by the fact that it covers almost all the case study area as seen on Table 6-1, above. The distribution of numbers of contributors occupies higher classes in the map in Figure 6-3**Error! Reference source not found.**.

The top two classes having the highest number of contributors especially contain 43 grid units while in Flickr and Panoramio there were 14 and 8 respectively. The coastal trend disappears in this map apart from some areas that are very popular among all photosharing projects.

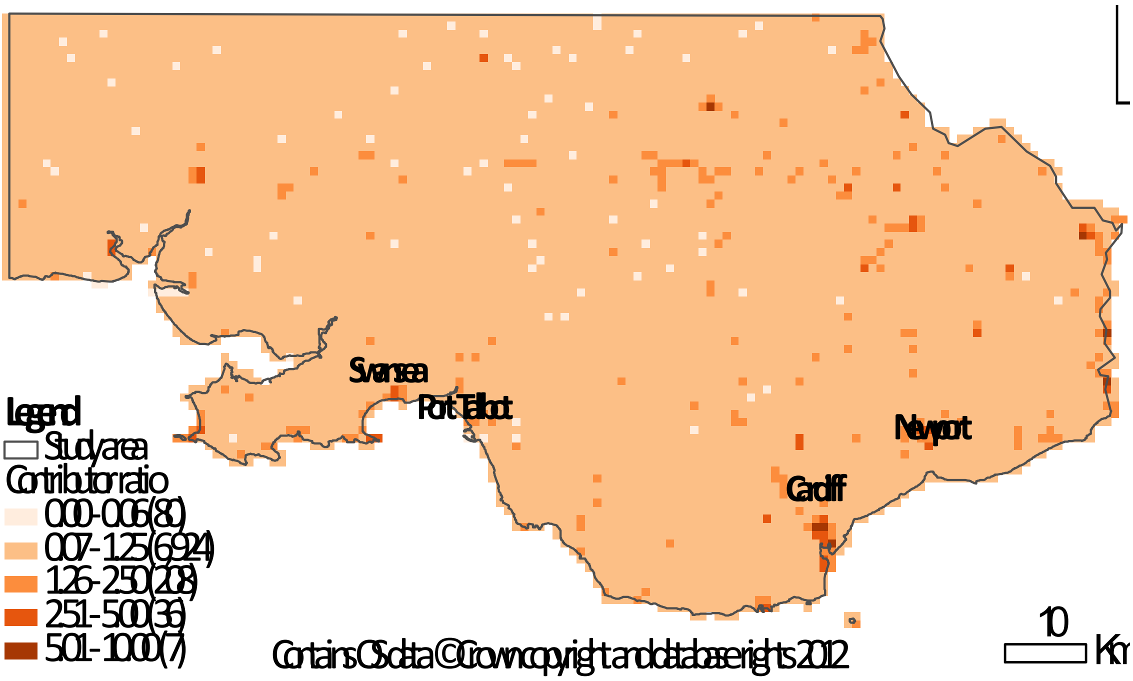

Figure 6-3:Spatial distribution of Contributor ratio for Geograph in south Wales

The trend in the main peak of the Beacon Beacons National Park is blandly recognisable. Amongst the grids with most contributors there are here 4 that are not located in the urban area of Cardiff and three are not in urban areas at all. One of them can be found in correspondence of the Hay Bluff a high and very panoramic hill close to the Wales-England border. A small cluster is also visible around the Welsh National Botanical Gardens.

6.5 Urban and non-urban areas

Following the procedure described in section 4.5.2, the classification of all the 1km BNG units in urban and non-urban led to 925 urban BNG units and 6,330 non-urban. Below in Figure 6-4 the map with the urban and non-urban classes of grids.

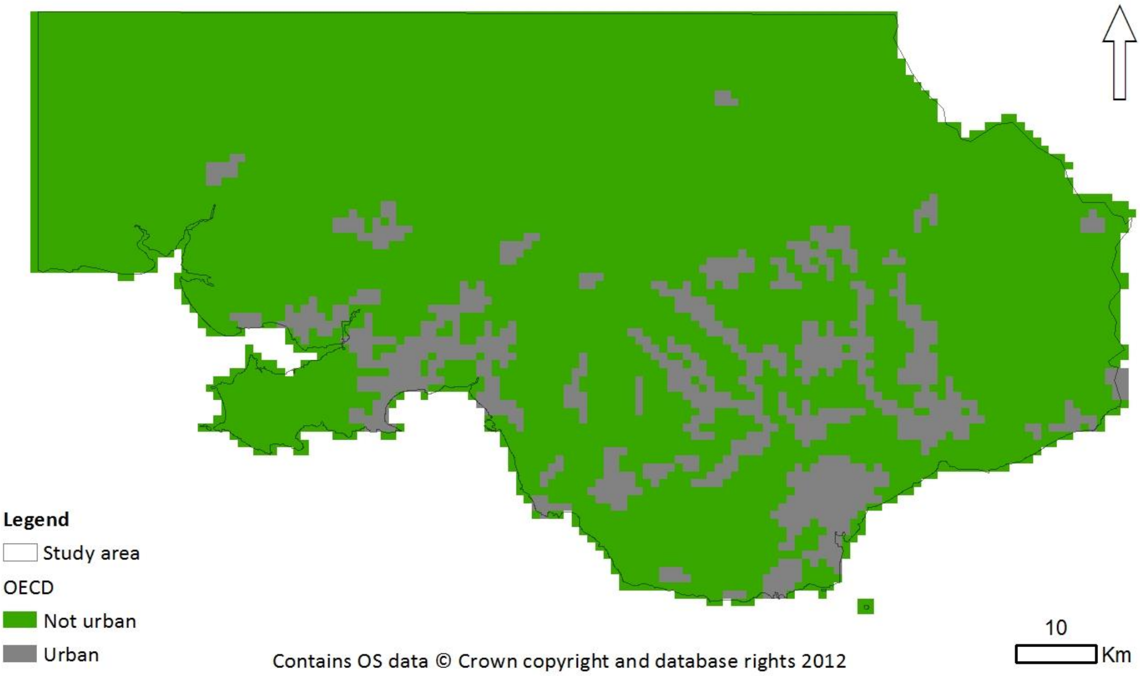

Figure 6-4 Urban and non-urban areas according to the OECD methodology.

The behaviour of the three projects is evidently different when urban and non-urban areas are considered as discerning categories. Examining the relation between urban and non-urban grids, we can see that all projects tend to have more contributors in urban areas, as expected. But while Flickr has in urban areas an average number of contributors that is nine times the average of non-urban areas, for Panoramio this relation is four times for Geograph is even just two times as can be seen on Table 6-2.

Projects	Non-urban contributors		Urban contributors		All study area contributors	
	Max	Avg.	Max	Avg.	Max	Avg.
Flickr	350	3.22	1,779	27.47	1,779	6.32
Panoramio	181	1.92	433	8.02	433	2.69
Geograph	70	3.91	77	8.01	77	4.43
All Projects	595	9.05	2,242	43.49	2,242	13.44

Table 6-2: Contributors productivity in urban grid units, non-urban and all grid units in the study area

Therefore, Flickr appears overwhelmingly a project more focused on urban areas while Geograph is the most non-urban project; with the highest average number of contributors in non-urban areas. Noteworthy the average number of contributors for Panoramio and

Geograph in the urban environment is very similar. Most clusters of contributors for all three platforms cover urban areas, with a smaller number located in non-urban areas (the vast majority in protected areas). Here, the number of contributors for all projects will be used as the unit of measurement to derive clusters and outliers. The characterisation of the local environments that are highlighted by clusters and outliers has also a dedicated section (6.6.4). The focus of this study is more on non-urban areas. As a cue on the overall focus of all projects on urban areas none shared pictures from one only urban grid unit, then 99.9% of urban grid units have at least one geo-located image. There are 5,255 grid units (87% of non-urban grids) with no resident population but with pictures shared online. Each non-urban grid unit had an average of nine contributors sharing pictures, and a maximum of 595 contributors submitting pictures from the same grid unit.

In the above sections, this work offered the basic statistical analysis and maps. They came after the aggregation of number of contributors in BNG grid units. In the following sections, more advanced form of spatial and statistical analysis will be used to answer the research questions.

6.6 Spatial analysis

We now turn to two levels of analysis. At first, the analysis of the relation between projects on a grid by grid basis will ignore the topological relations between grids. In the next level of analysis, we take into consideration also the topological component. Therefore, there will be an analysis of the spatial patterns, hotspots, and outliers using all combined platforms. The correlation between project considering every grid unit separately, as explained in section 4.5.5 is investigated through the development of scatterplot diagrams and Rank Spearman's statistics in section 6.6.1. The scatterplot diagrams have been derived only for the non-urban grids while the Rank Spearman's test has been derived for the overall study area and for the non-urban grids. The global analysis for local clusters and outliers in section 6.6.2 uses the local Moran's I method.

6.6.1Relationships between projects

To compare the three platforms in a grid by grid basis using the number of contributors to each project two methods have been used. The design of scatterplot diagrams and the estimation of the Spearman's rank test. It is run through couples of projects and gives a measure of the correlation between the two ranked collections ρ and it measures the statistical significance of it. The large sample ensures that we will have always a high significance ($p<0.001$) the value of the ρ and the scatterplots will help to assess the significance of the correlations. The scatterplot method offers a visual impression of the correlation grid by grid. To design the scatterplot, the contributor ratio (Cr) was used. The visual analysis of scatterplots, while it might lead to misleading interpretation due to overlapping points is very helpful to figure out outliers. For the calculation of Spearman's rank test, the use of this ratio is not needed since this statistical method ranks the values for every project and compares the ranks but does not consider the relative magnitudes[75]. In Table 6-3, below a summary of the Spearman's rank test applied to all study area grid units.

All study area	Flickr	Panoramio
Panoramio	ρ=0.662301; p<0.0001	
Geograph	ρ=0.6668766; p<0.0001	ρ=0.5898477; p<0.0001

Table 6-3: Spearman rank test for all grids of photosharing case study

As can be expected, from the comparison of the maps the three projects are highly positively correlated. All of them share the main hotspots in the urban areas that are highly positioned in all ranks. Flickr and Geograph overall have the best correlation while Panoramio and Geograph show the least, while still relevant, correlation. In Table 6-4, below, a summary of the Spearman's rank test applied to the non-urban grid units.

Non-urban only	Flickr	Panoramio
Panoramio	ρ=0.6123573; p<0.0001	
Geograph	ρ=0.6080886; p<0.0001	ρ=0.5381275; p<0.0001

Table 6-4:Spearman rank test for non-urban grids of photosharing case study

There is still a relevant positive correlation between the three projects and in this case the lead in terms of correlation is to the couple Flickr–Panoramio, while still the couple Panoramio-Geograph lags behind the other two couples. Flickr therefore keep being the most correlated project and his worst performance (correlation with Geograph in the non-urban set of BNG units) is still better than the best performance of Geograph and Panoramio best performance. The scatterplots give a better idea of the similarity of evaluation amongst the different platforms. We computed six scatterplots connecting two project a time for urban and non-urban grid units. To help understand the relevance of such diagram is here helpful to remind how most of the grid units are without contributors

[75] As a verification of this the test run using the number of contributors and their percentage, as calculated for the scatterplots, gave exactly the same results in terms of trend and significance.

both in Flickr and in Panoramio. Given to the incentives to spread in space implemented in Geograph, it has almost 99% of all grid units with contributors, while Flickr does not reach 60% and Panoramio not even 49%. For the non-urban areas, Geograph coverage is closer to 99%, while the percentages fall to 54% and 43% for Flickr and Panoramio, respectively. Therefore, there will be the majority of Flickr and Panoramio points standing for grid units with zero contribution that are spread along the horizontal and vertical axis when compared with Geograph which has contributors in almost every grid unit.

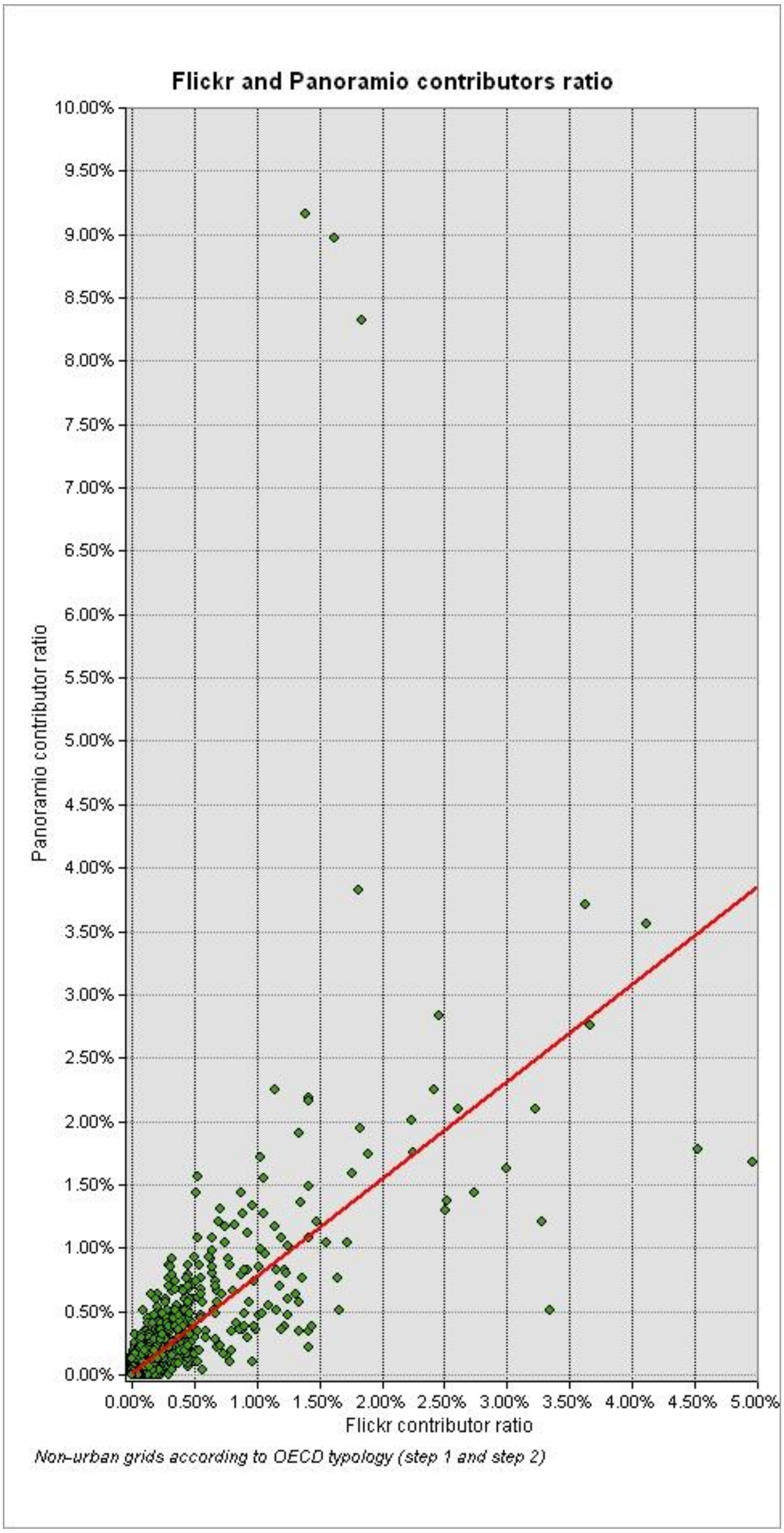

Figure 6-5: Scatterplot with Flickr and Panoramio contributor ratio considering all grids

The scatterplot for all grid units involving Flickr and Panoramio showed clear correlation between the two platforms (Figure 6-5) as expected given the result of the rank Spearman's test. Most interestingly almost all grids with no Flickr contributors are also with no Panoramio contributors, those grids are all overlapping in the origin of the diagram. For the rest of grid units, if we omit few outliers, Flickr and Panoramio show very similar percentages of contributors who took pictures in all grid units. The above scatterplot underlines the similarity between the two platforms as suggested by the line representing the trend. The diagram also stressed the points of divergence between the two platforms. All of them are situated in urban areas. The trend is a bit unbalanced towards Flickr.

When drawing the scatterplot considering only non-urban grid units (Figure 6-6) the convergence between the two platforms is stronger and the trend underlines a small prevalence of Panoramio's contributors for non-urban grids.

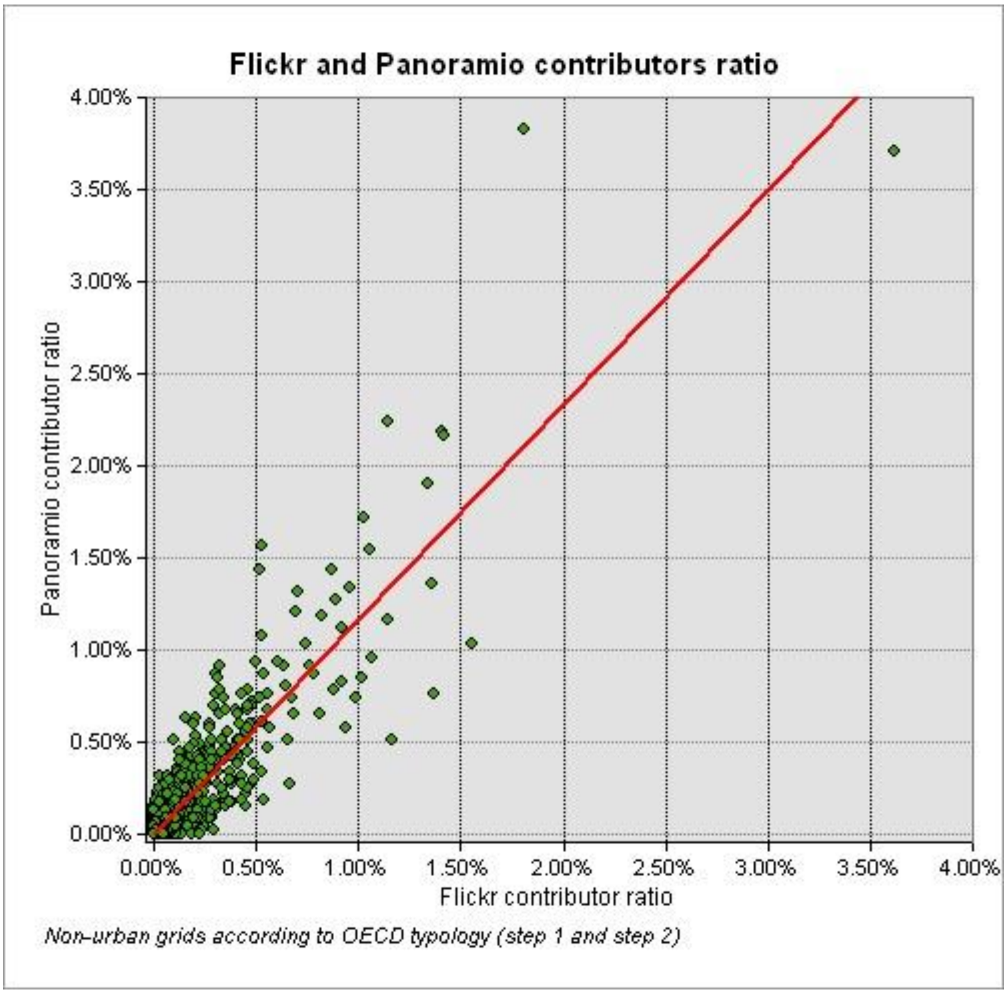

Figure 6-6:Scatterplot with Flickr and Panoramio contributor ratio considering non-urban grids

The points further away from the trend represent in the case of Flickr grids that are contiguous to Cardiff but not classified as urban while in the case of Panoramio they are places with specific natural and cultural elements. They will be better described in section 6.6.4. Even more than in the previous scatterplot most of the grids have no contributors in both projects and are concentrated in the origin of the diagram. The presence of a so high number of zero values in both platforms can explain the inverse results between the scatterplots and the rank spearman test. The test creates ranks in the two lists of contributors per grids and then compares if the same grid unit is in a similar position in the rank. In our case the data is zero inflated, half of the couples have zero values, therefore the test cannot rank them.

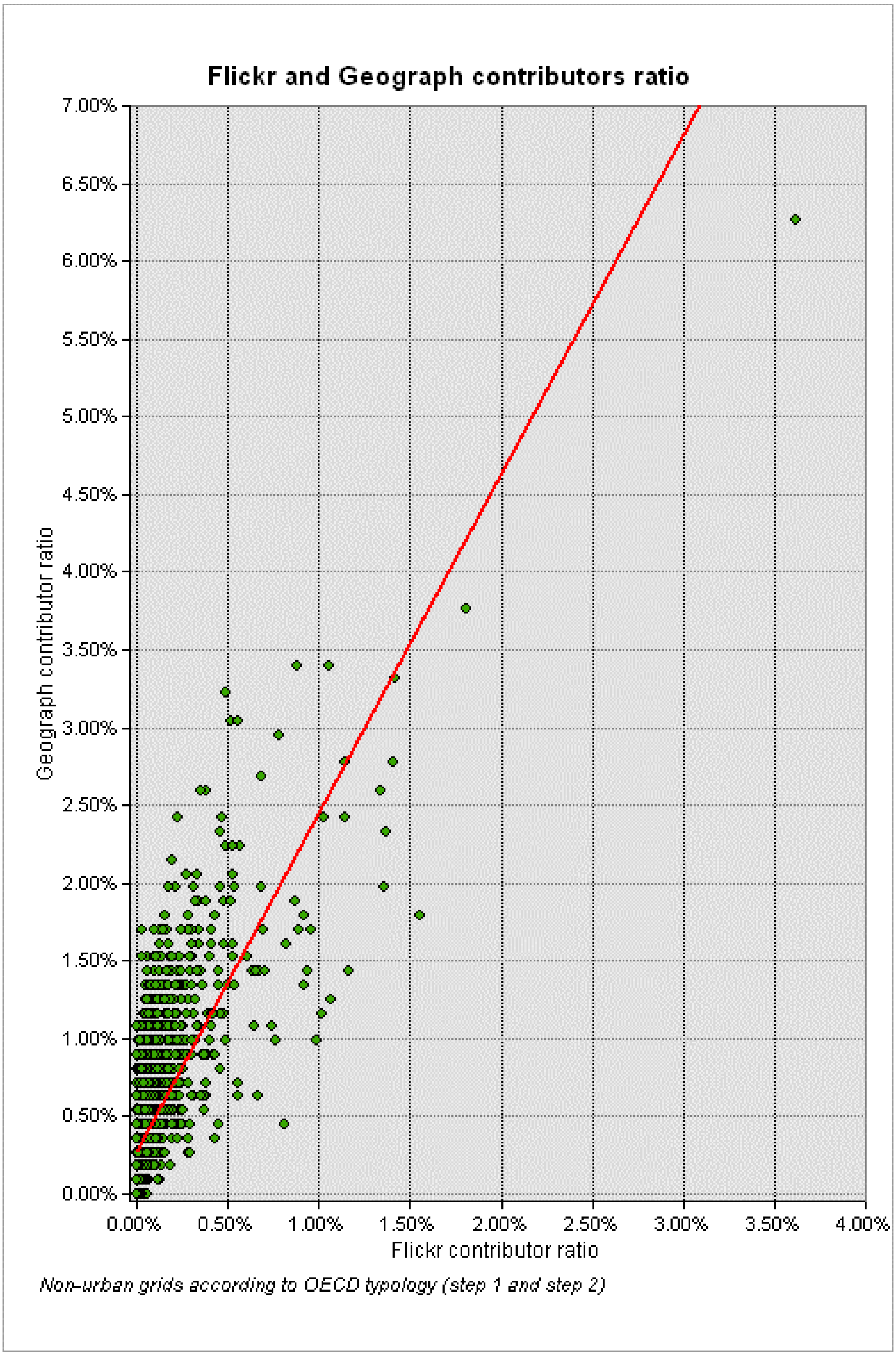

Figure 6-7:Scatterplot with Flickr and Geograph contributor ratio considering non-urban grids

When there are more couples of no zero values, as happens including the urban grids, the ranking is more accurate and the correlation is proved stronger even with more visible outliers. The strong correlation between Flickr and Panoramio implied that those two platforms behaves similarly when compared to Geograph. The relationship between Geograph and Flickr is strongly skewed toward Geograph (Figure 6-7) when comparing non-urban grid units. This emphasized the selectiveness of people using Flickr and Panoramio when dealing with areas worth sharing visually, compared to Geograph contributors, whose aim is to cover all areas. Here we can see that most of the grids with no contributors in Flickr got contributors in Geograph therefore we see the spread of points along the lowest part of the vertical axis. The scatterplot that represents the overall comparison between Flickr and Geograph (Figure 6-8) shows more outliers and a trend that keeps favouring the number of contributors from Geograph

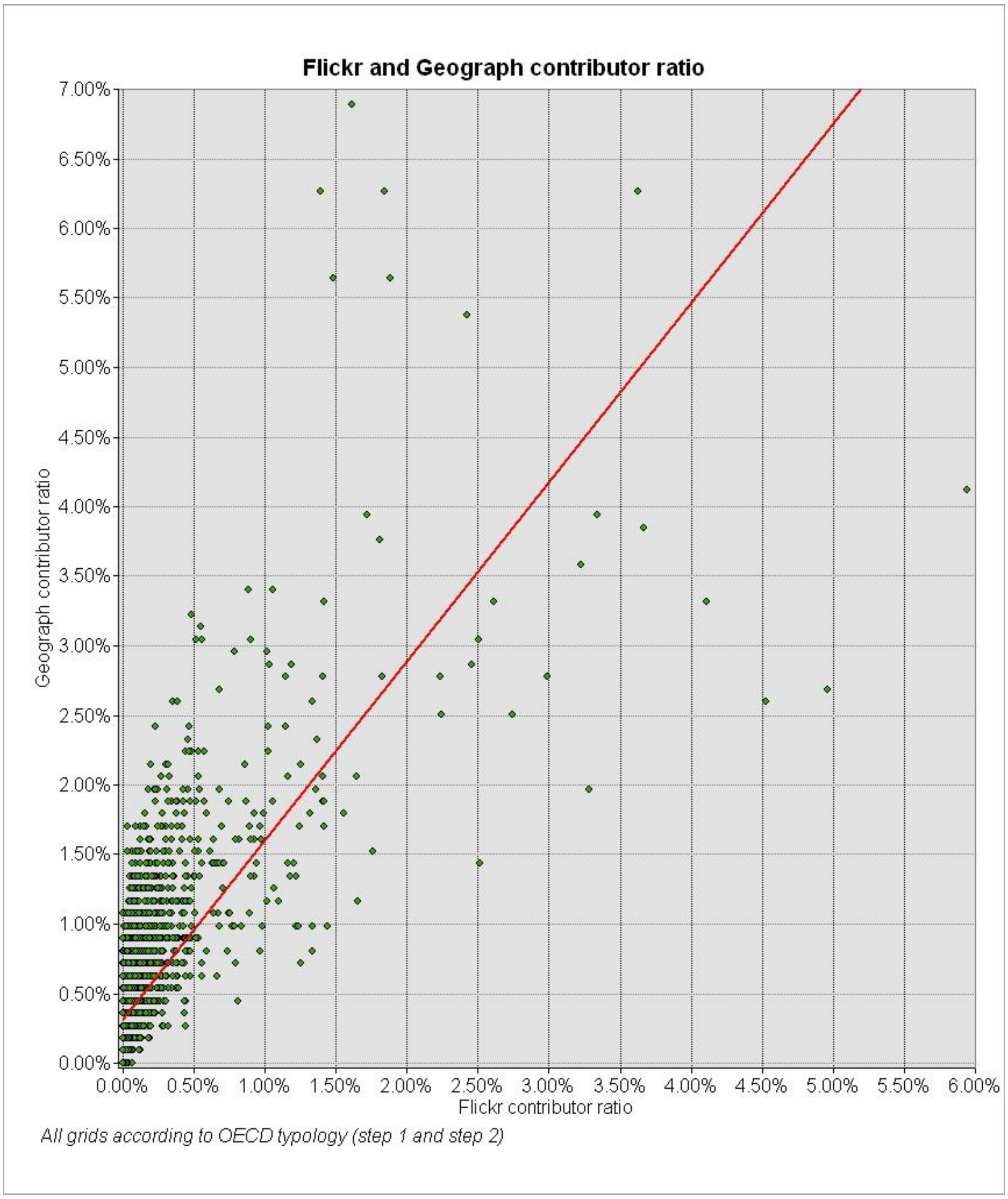

Figure 6-8:Scatterplot with Flickr and Geograph contributor ratio considering all grids

The full comparison considering Geograph and Panoramio is in Figure 6-9. As expected the results resemble the above diagram relating Geograph and Flickr.

Figure 6-9:Scatterplot with Panoramio and Geograph contributor ratio considering all grids

The relation between Geograph and Panoramio considering only non-urban grids is in the following Figure 6-10. The most urban outliers are excluded from this representation and there is the usual trend that stresses how Geograph covers more territory with less contributors.

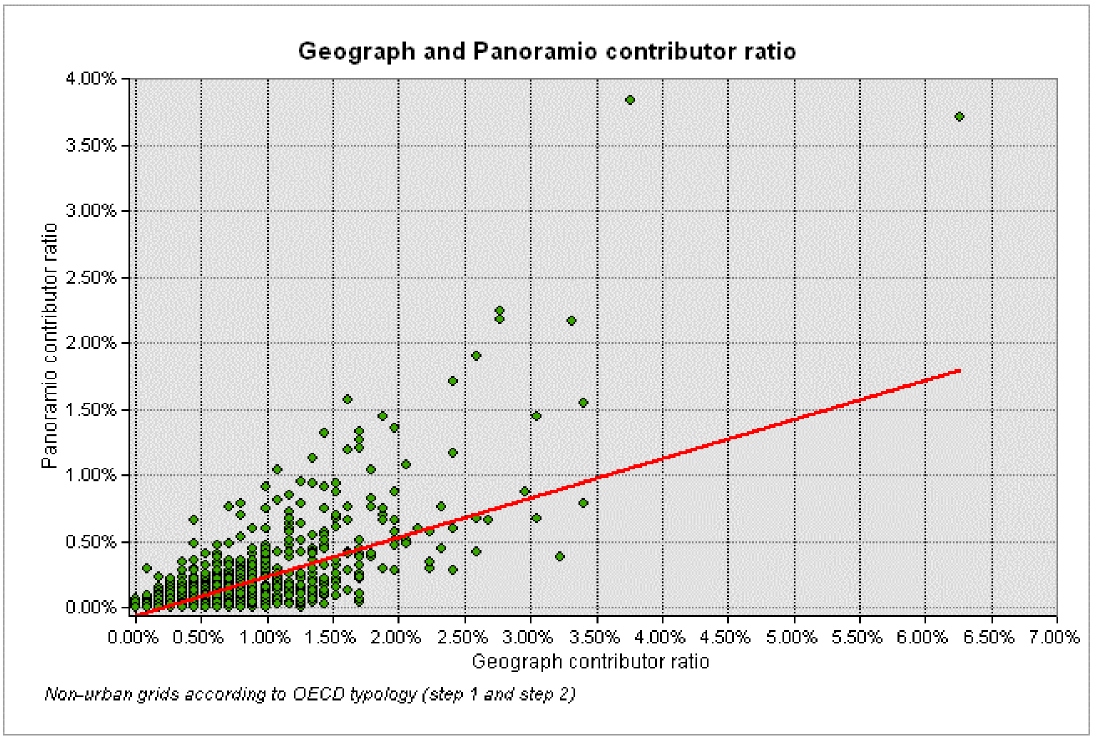

Figure 6-10:Scatterplot with Panoramio and Geograph contributor ratio considering non-urban grids

6.6.2Spatial statistics

The topological element is here considered with global and local statistics. This section shows the computation of the global and local statistics that accounts for spatial autocorrelation. The calculation of spatial local autocorrelation helps the identification of environmental spaces that are statistically and spatially relevant. The identification of grid units that individually or in spatial aggregations have values that are in stark contrast with the values of surrounding grids have been found using the values given by the total amount of contributors to all photosharing projects. By analysing the spatial distribution of contributors, we observed the movement of people from urban to non-urban areas when they take and share pictures. These people are behaving as receptors of the CES. We then considered the number of users of the three photosharing platforms, focusing on the distinction between urban and non-urban grid units (Figure 6-11). The Moran's I global spatial autocorrelation index, calculated over the distribution of total contributors for every grid unit using a queen contiguity[76] ($I = 0.0651$, $\rho = 0.001$), reveals that there is < 1% probability that the autocorrelation has been generated by a random distribution. While previously the use of the contributor ratio was functional to the detection of correlations, in this case the number of contributors is preferred as in this case every individual contributing is considered regardless of the project to which he contributed.

[76] Two contiguous polygons share an edge and/or a vertex

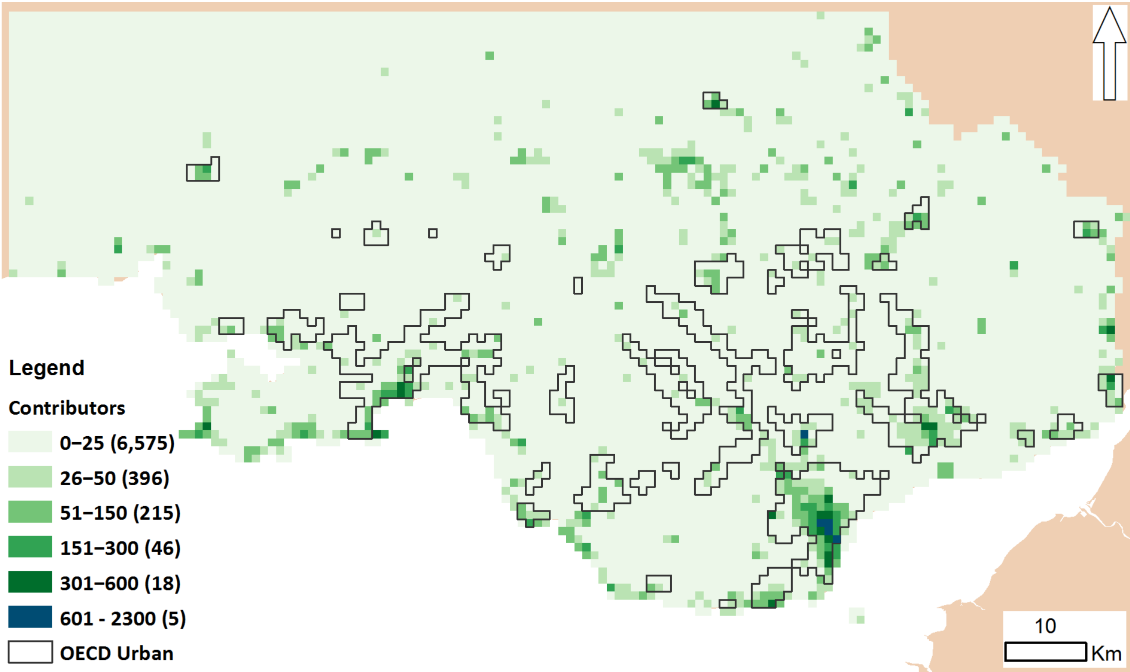

Figure 6-11:Total contributors from photosharing projects and classification of BNG grids according to OECD procedure

There are visible clusters of grids with elevated number of contributors in both urban and non-urban areas. In general, as can be seen in Figure 6-11 above, there are several areas where the aggregation of contributors to photosharing platforms is particularly evident.

6.6.3Spatial trends: clusters and outliers

This study derived clusters and outliers, which will comment on in section 6.6.4. Here, we are more interested in the non-urban areas. To detect spatially statistically significant grid units, such as spatial clusters and outliers, this study used Moran's I local indicator of spatial association (LISA) (Anselin 1995) as implemented in ArcGIS 10.2.2 spatial statistic tools. This brought to the finding of both clusters of grid units with high numbers of contributors and grid units that differed markedly from surrounding grid units. There were 295 grid units with significant local autocorrelation. They can be grouped into 40 spatial clusters of grid units having a high number of contributors, while the other 36 grid units are considered individual or outlying because they are surrounded by grid units with lower numbers of contributors. Filtering out clusters and outliers pertaining to the urban areas resulted in 20 clusters with high values and 17 outliers (isolated grid units characterised by values higher than their surrounding units).

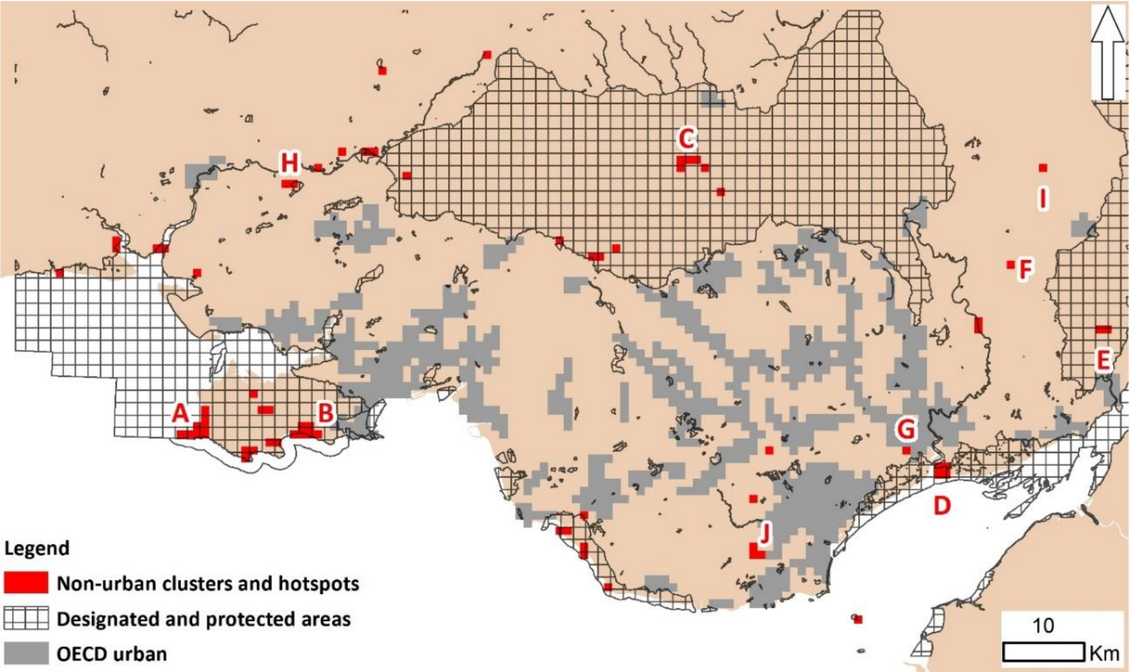

Figure 6-12: Non-urban clusters and hotspots, as identified by Moran's I local indicator of spatial association (LISA). Letters indicate clusters of grid units discussed in the text

Therefore, we have delimited spatial clusters and found that approximately 45% of them are non-urban. We analysed these clusters and outliers to assess if those patterns or hotspots are associated with natural and cultural features.

6.6.4Visual inspection: Sites of special photographic interest

Among all the statistically relevant elements derived using LISA, eight grid units form the largest cluster, which partly covers Rhossili Bay on the Gower Peninsula. The most easily accessible part of the bay has the most contributors (Figure 6-12, cluster A). In 2014, a popular touristic website ranked the bay as the ninth best beach in the world (TripAdvisor, Top 25 beaches — world, http://www.tripadvisor.co.uk/TravelersChoice-Beaches-cTop-g1). This cluster is second in terms of the number of pictures and seventh in terms of average number of contributors. It includes the second and fifth most popular grid units, the first of which is the most popular grid unit in Panoramio, and the second of which is the most popular grid unit in Geograph. The cluster includes part of the Gower AONB, the Carmarthen Bay SPA, three SSSIs (Gower Coast: Rhossili to Port Eynon; Rhossili Down; and Sluxton Marsh, Whitemoor), and the Gower Coast NNR.

Names and tags help to understand the cognitive and affective values associated with this cluster. Amongst the 1,360 pictures shared in Panoramio, 1820 in Flickr and 497 in Geograph most of them refer to views of cliffs and the long beach. There are recurrent adjectives like magnificent, magic, beautiful, as well as feelings and activities aroused due to contact with the local environment. There is also contemplation of the landscape as well as natural phenomena such as a large number of pictures aiming at the sunset (also referred to as "golden hour") that mostly in this area happens with the sun going down towards the sea. Most of the pictures are targeted at the natural environment or to some heritage/scenic elements such as wrecks. Several pictures, mostly in Flickr, have names and tags in foreign languages.

Six grid units form the second largest cluster (Figure 6-12, cluster B), which covers Three Cliffs Bay (three SSSIs, one AONB, and one NNR). This cluster includes the 14th most popular grid unit. To understand the main reasons to share pictures of this area we can see how the 588 pictures shared in Panoramio, 1,472 in Flickr and 259 in Geograph are described by users. The very scenic three cliffs bay is often portraited along with the individual beaches that constitutes the bay. Secondly the views of the ruins of the Pennard Castle are often portrayed with a natural background.

Two clusters of four grid units are in two diverse environments. One, which includes the third and tenth most popular grid units, covers the highest peaks of the Brecon Beacons (Figure 6-12, cluster C), which give the name to the National Park and the SSSI partially covered by the cluster. This cluster includes the third and tenth most popular grid units.

We have here 505 pictures shared in Panoramio, 2,007 in Flickr, and 290 in Geograph. Most pictures are referred to views or the various main natural features of the area like the summits of Pen y Fan, of the Corn Du, of the Cribyn, and the small lake Llyn Cwm Llwch. Some images are also named explicitly to stress the magnificence of the views or also stressing the sense of achievement "we made it, what achievement!". In this area you have often pictures aimed at meteorological phenomena such as snow, and mist. Several pictures represent ascending and descending phases of groups of hikers. The second cluster is situated in the Severn Estuary SSSI and covers the Newport Wetlands SSSI and NNR (Figure 6-12, cluster D). A mixture of landscape and close-up pictures like those used for environmental observation characterises this cluster. We have here 114 pictures shared in Panoramio; 1,738 in Flickr, and 102 in Geograph. In Panoramio we have mostly pictures of the landscape, some pictures of species and some group picture to remember enjoyable visits. In Flickr there is an overwhelming majority of pictures portraying birds, bugs, and wildflowers. Here landscapes are portrayed in comparatively a minority of pictures, tags also report how most pictures are family or fieldtrips. The lighthouse also got some attention. In Geograph we have only pictures targeting landscapes or human made artefacts. We can see how the three platforms provide a blending of preferences. The value stays in the integration and in the comparison of platform.

Two grid units form the most popular cluster (Figure 6-12, cluster E). This cluster has the highest average number of contributors in any platform; it contains the overall most popular grid unit on both Flickr and Geograph, and the second most popular grid unit on Panoramio. This grid unit contains a feature that is cultural in a natural area: Tintern

Abbey, in the Wye Valley AONB. It is very close to the River Wye, which has a dedicated SSSI. The remaining grid unit in the cluster owes its popularity to the views over the abbey and the natural area surrounding it. In this cluster the large majority of pictures depict the ruins of the Tintern Abbey in the context of the surrounding natural environment. Even photo albums devoted to car or motorbike rallies devote several pictures to the cultural and natural landscape. Those gatherings use the facilities around the ruins to gather, consume, and then enjoy the view of the local environment. Tintern Abbey is also included in the Cistercian way.

Looking at outliers, the second most popular statistical element overall, which is also the sixth most popular grid unit, includes the ruins of Raglan Castle (Figure 6-12, cluster F). To this cluster contributed also events organised in the area such as re-enactments and the fact that the ruins has been used as scenery for some open-air plays. Another outlier is Tredegar House and its gardens, which is the 20th most popular grid unit (Figure 6-12, cluster G). Three kinds of pictures contributed to the success of Tredegar House: those of the house itself; those of the park, with flowers and birds; and those of an annual vintage car show, attended by several Flickr contributors. Pictures here in Panoramio range from panoramic views of the heritage house and gardens, to group pictures plus some close-up images of natural elements such as birds, trees, and flowers. In Geograph we have heritage and nature focused pictures as well as pictures of events held in the gardens such as an "agricultural show" held in 2010. The social component of gatherings is enhanced in Flickr

The cluster containing the National Botanic Garden of Wales is also a statistical outlier (Figure 6-12, cluster H). Here pictures are largely aimed at close range shots of natural elements, but pieces of landscape art and buildings also contribute to the attractiveness of this place. Both Raglan Castle and Tredegar House are not included in conservation areas, but do constitute cultural heritage. The grid unit containing Skenfrith Castle and the adjacent St Bridget's church with his imposing tower (Figure 6-12, cluster I) constitutes a statistically significant outlier that does not include conservation areas as well but that inspired several visitors, yet again we have a blending of heritage, nature, and spiritual values.

All this evidence points to the blending of natural and cultural elements that are considered worth sharing visually. The range spans from predominantly natural features such as coastal landscapes, waterfalls, and open views to predominantly manmade cultural features such as castles, Neolithic dolmens (Figure 6-12, cluster J), and National Trust houses. Cultural features are more likely to be included when they form part of conservation areas and have gardens used to host events.

It is evident how not only the features of the landscape either natural or cultural constitute an element of interest, but also social gatherings and cultural visits are well documented.

6.7 Photosharing in Greater London (Panoramio and Geograph)[77]

To test in a different, more metropolitan, and cosmopolitan environment the relevance of photosharing projects to detect environmental spaces a side project on Greater London has been developed.

6.7.1 Material and methods

Photosharing data used come from the photo collections of Geograph and Panoramio. The study used a grid of 100 meters to aggregate and count the number of contributors. The size of the grid is smaller since at the urban scale we are interested in finding out if photosharing can help detecting the engagement for nature also expressed through concentration of pictures and users in private gardens[78] and small public gardens rather than only taken in large parks or reserves. Another factor considered lies with the horizontal positional accuracy of GPS (William J. Hughes Technical Center 2016) that is generally in a range between 4 and 20 meters but in urban areas can decrease when the receiver is situated close to buildings. Senaratne et al. (2013) measured the distance from a target and the correctness of the geotagging. Their study focused on hi visibility touristic built landmarks while in our case the focus will be on natural areas, less visible, less advertised so we expect to have better proportion of correctly geotagged images[79]. This study includes two main phases. At first the centroids of every grid unit were used to compute a Kernel Density Estimation with a 200m bandwidth to have a visual image of spatial clusters created by contributor's aggregations. This visualisation only using KDE has already been done (Hollenstein and Purves 2010, Comber et al. 2015, 2016). Hollenstein and Purves (2010) calculated the bandwidth only on Parks in London and found that a search radius of 270 m was optimal. Furthermore, KDE allows a better visualisation of the Greater London area that includes 166,750 100 m grid units compared with the 7,255 1 km grids used in south Wales. As a second and more in-depth analysis, this study investigated the density of contributors in relation to the presence of surface water in the same 100m grid unit. This in-depth analysis has been performed only on two of Greater London’s Boroughs as representatives of inner and

[77] This work is derived from a project in fulfilment of the Master in Research (MRes) in Biodiversity, Evolution & Conservation in UCL. It has been performed under the candidate supervision and assistance by James Koh. The student performed under the candidate supervision the data analysis and have drawn maps and diagrams. The candidate individually performed the data collection and filtering, the writing of all the text in 6.7 and subsections.

[78] But in this case the number of users will necessarily not focus on the same features but on similar features in close gardens

[79] In the mentioned paper, the corrected geotagged pictures for two of the main Berlin's landmarks like the Brandeburg Gate and the Reichstag were respectively 51% and 62%

outer London. To investigate the correlation the Wilcoxon rank sum test has been performed.

This sub project used 25,954 photos shared by 3,974 contributors from Panoramio and 190,938 photos shared by 1,259 Geograph contributors. This relation is in line with the statistics found in south Wales (6.3) where 49,966 photos shared by 4,725 contributors from Panoramio and 117,850 photos shared by 1,117 Geograph contributors. In Table 6-5, below the comparison of the two study areas and the productivity of contributors. The relation between the number of images with the number of contributors gives the productivity of contributors. Geograph hold always a higher productivity.

		Panoramio	Geograph
South Wales (7,255 km^2)	contributors	4,725	1,117
	contributors/km^2	0.65	0.15
	images	49,966	117,850
	Images/km	6.89	16.24
	productivity (Images/contributors)	10.57	105.51
Greater London (1,667.5 km^2)	contributors	3,974	1,259
	contributors/km^2	2.38	0.76
	images	25,954	190,938
	Images/km	15.56	114.51
	productivity (Images/contributors)	6.53	151.66

Table 6-5:Comparing photosharing in south Wales and Greater London

Compared to south Wales there are less images and less contributors from Panoramio while there are more images and more contributors to Geograph. But this is due to the smaller area considered which is less than a quarter of the south Wales case study. We have that averaging to the smaller surface we have that for both projects in terms of images and contributors per km^2 there are quite higher values. It is to stress here how the gulf between the two platforms widened in Greater London and how the productivities of the two projects diverged with Geograph having an increase and Panoramio having a decrease. This can be explained by the fact that a largely urban area is less prone to offer scenic views as the ones considered more fit for the usual Panoramio contributors.

6.7.2Results

The first stage of this analysis delivered two contributors' density maps covering Greater

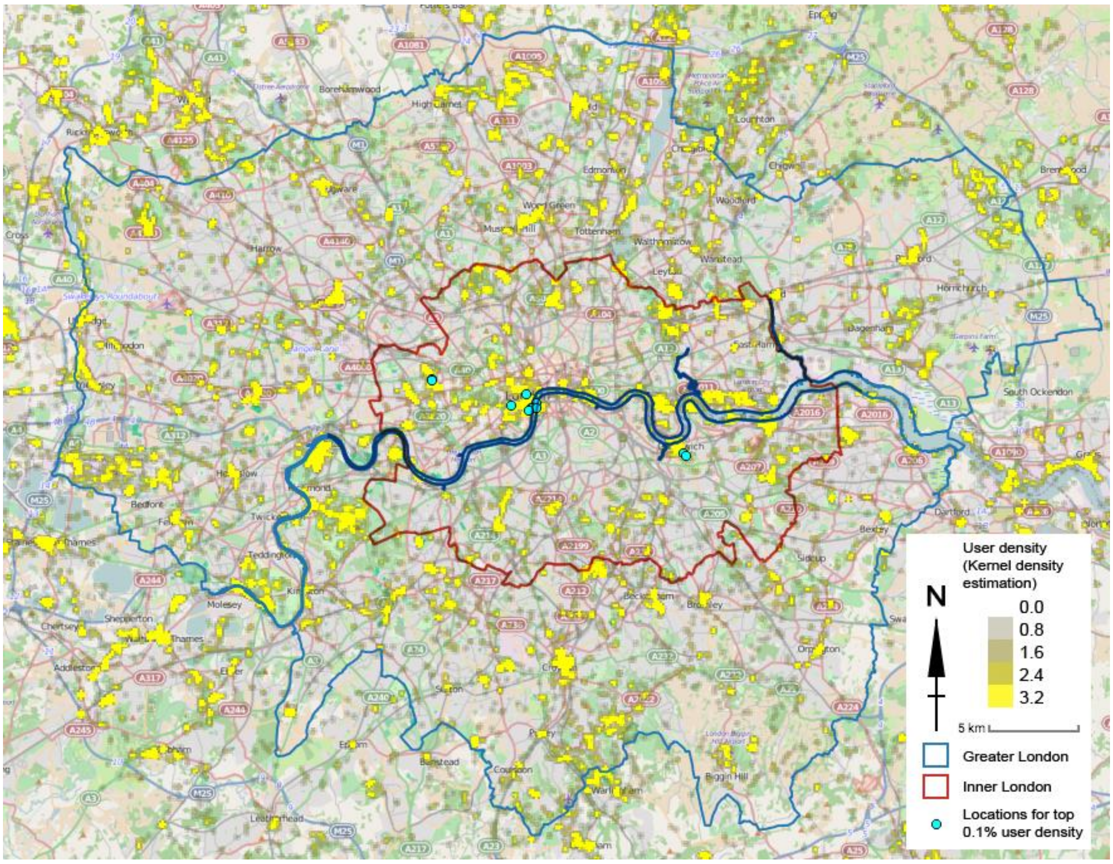

Figure 6-13: Contributors density distribution for Panoramio in Greater London

London Figure 6-13 for Panoramio and Figure 6-14 for Geograph. The visual comparison of the two maps underlined how Geograph had an overwhelming majority of contributors in the core area of Greater London, characterised by higher accessibility, higher population density, and a relevant flux of tourists. While the density of contributors to Panoramio is more widespread across Greater London. In the maps there are also the location of the top 0,1% grid units and the aggregation of London Boroughs and the City that are referred to as *inner London*. Here are almost all hotspots and the wide area with more contributors to Geograph. One interesting results of this analysis stresses how amongst the top 0,1% grid units we have mostly cultural iconic elements, following we have natural areas or parks, also in this case iconic ones. But surprisingly we have a considerable amount of rail stations even when they have not iconic value.

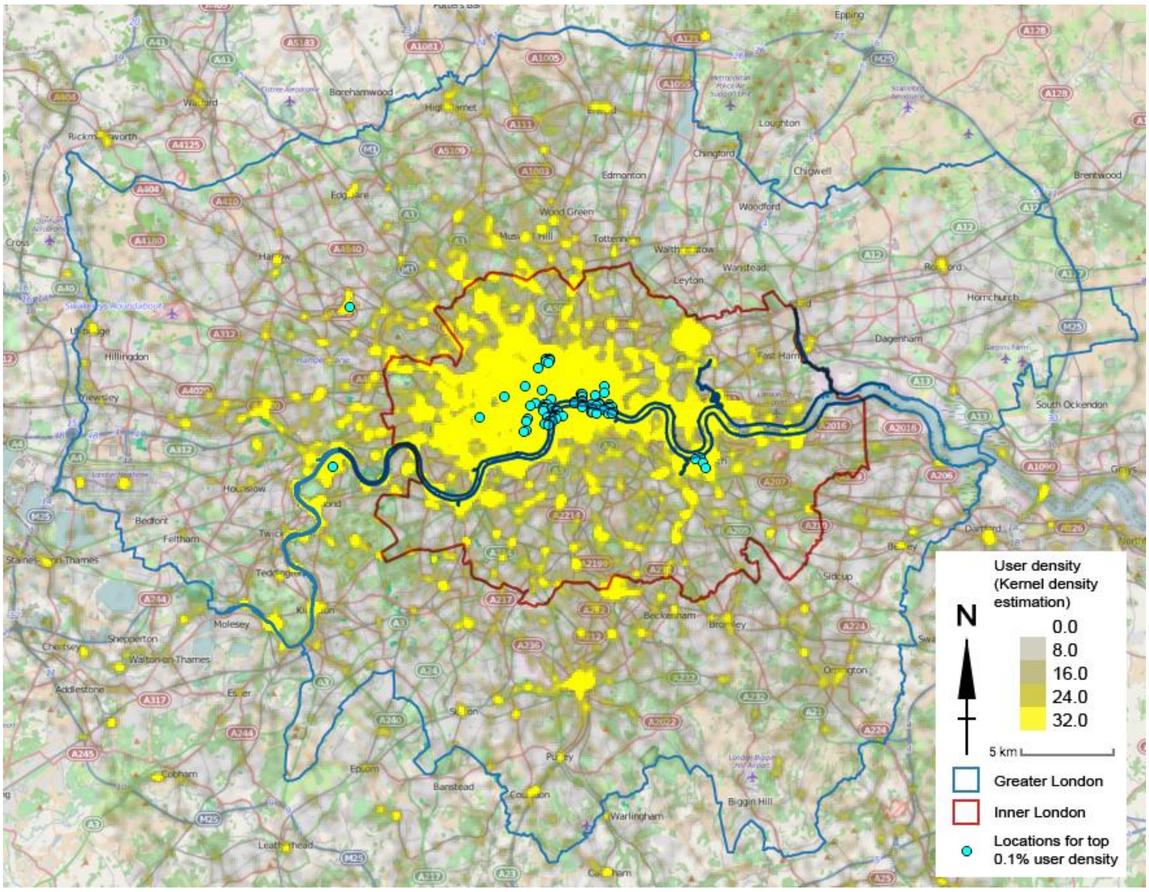

Figure 6-14:Contributors density distribution for Geograph in Greater London

The great disparity between inner and outer London suggested the need to create two specific studies delivered to two of Greater London's Borough with one included in Inner London and another outside of it.

Hackney as a representative of inner London and Enfield out of it have been chosen as shown in Figure 6-15. The specific analysis on Hackney and Enfield focused on the relation between the number of contributors and the presence of water bodies. Water bodies, as also found in south Wales, constitute a significant element of the experience of ecosystems. The Wilcoxon rank sum test applied to both projects separately gave the following values. For Panoramio there is a significantly higher density of contributors in proximity to water bodies both in Enfield, $W(8475) = 7157508$, $Z = -10.15$, $p < 0.001$, and in Hackney, $W(2062) = 304583.5$, $Z = -6.78$, $p < 0.001$.

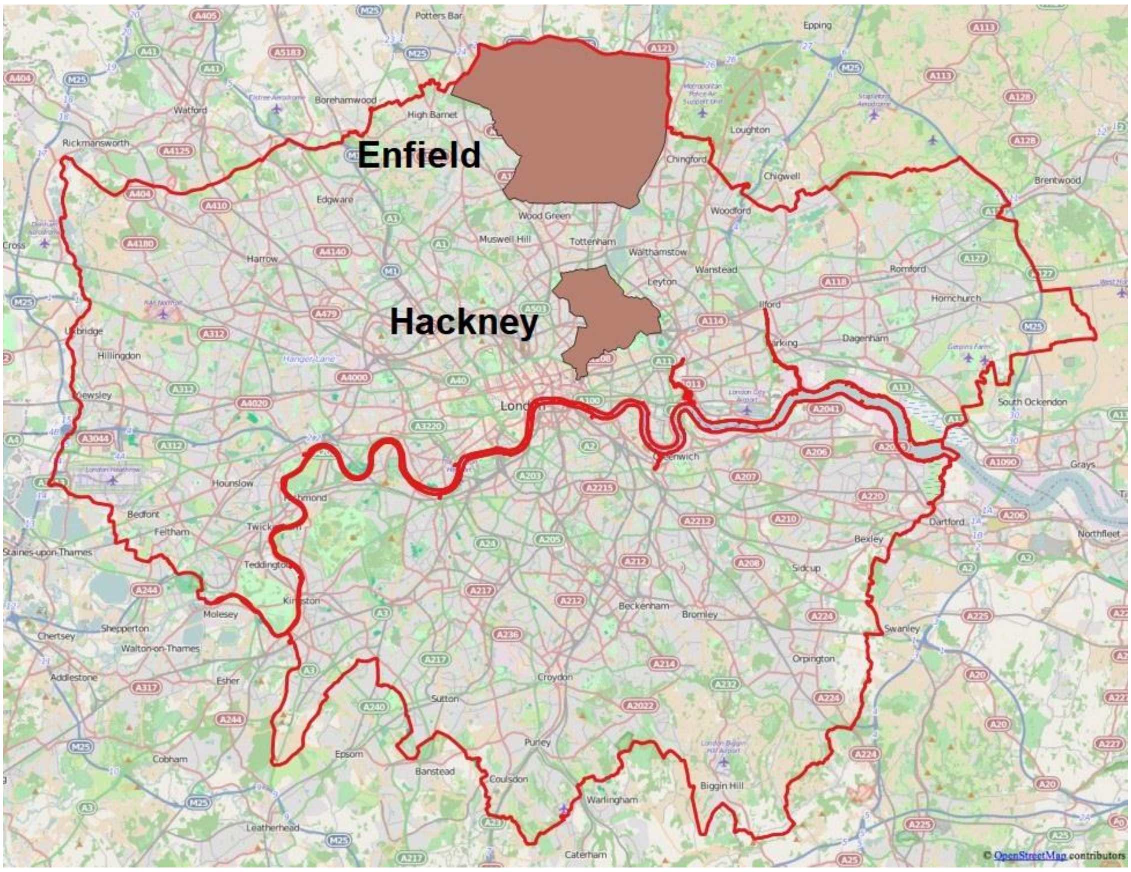

Figure 6-15:Hackney and Enfield in Greater London

The diagrams in Figure 6-16 and Figure 6-17 show that while the majority of grid units is without water bodies all grids with the highest number of contributors to Panoramio contain water bodies

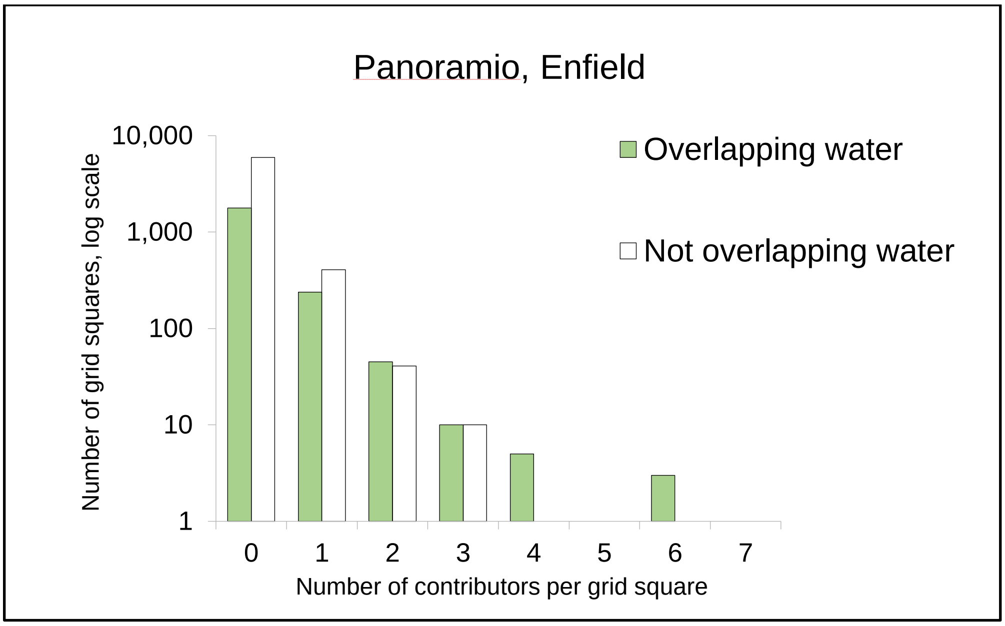

Figure 6-16: Panoramio density of contributors per 100m grid cells in Enfield

.

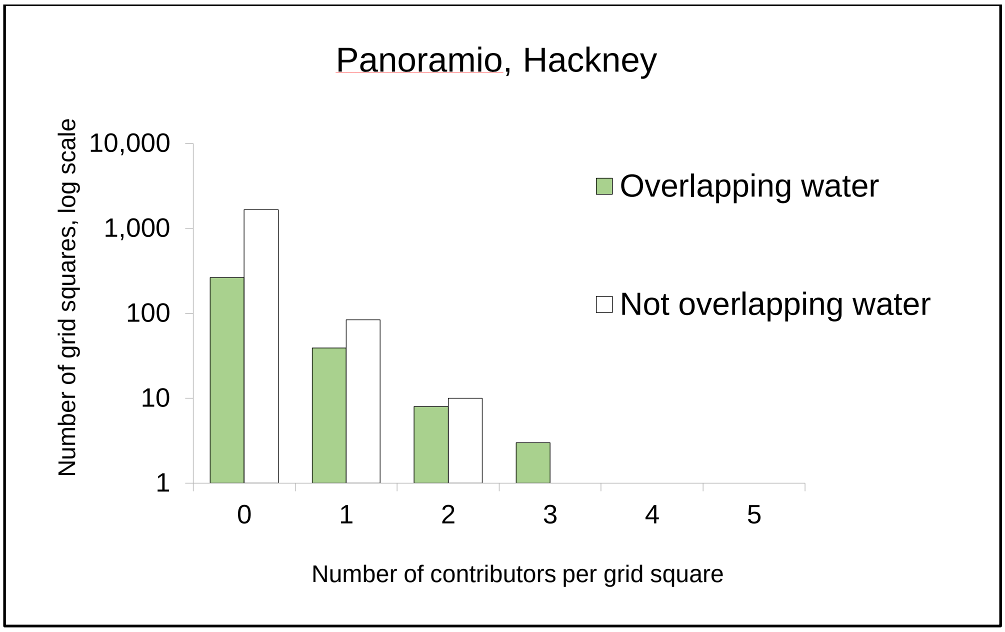

Figure 6-17: Panoramio density of contributors per 100m grid cells in Hackney

For Geograph there are significantly higher density of contributors in proximity to water bodies for Enfield, W(8475) = 7114616, Z = -5.77, $p < 0.001$, and Hackney, W(2062) = 380730.5, Z = -11.63, $p < 0.001$. The statistical evidence in this case is less supported by the diagrams in Figure 6-18 and Figure 6-19. Here the trend shows how while still the vast majority of grids do not contain water bodies. The grids with higher number of contributors are almost equally shared by grids with or without water bodies.

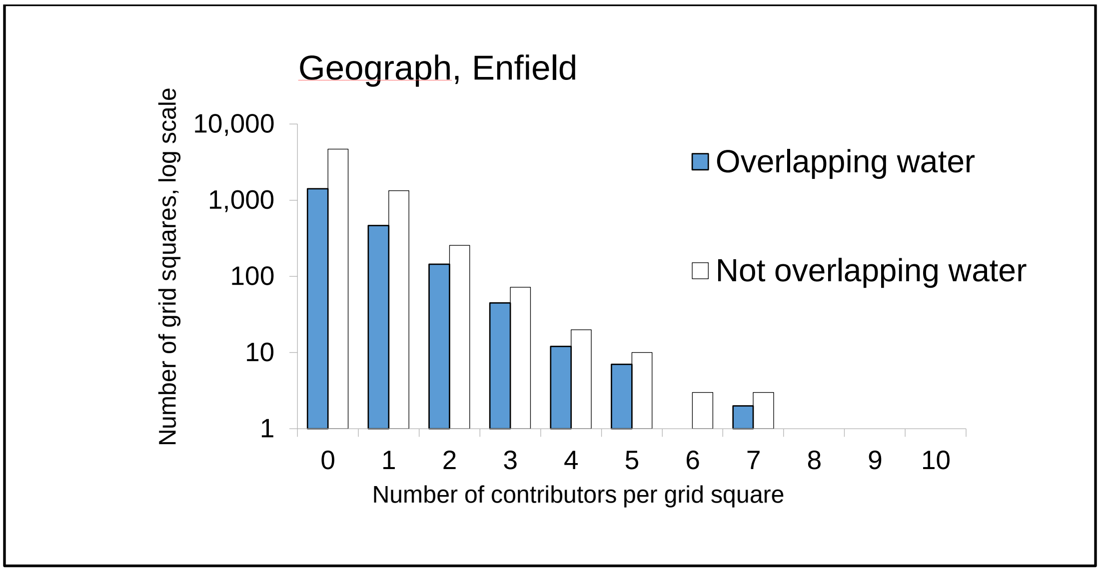

Figure 6-18: Geograph density of contributors per 100m grid cells in Enfield

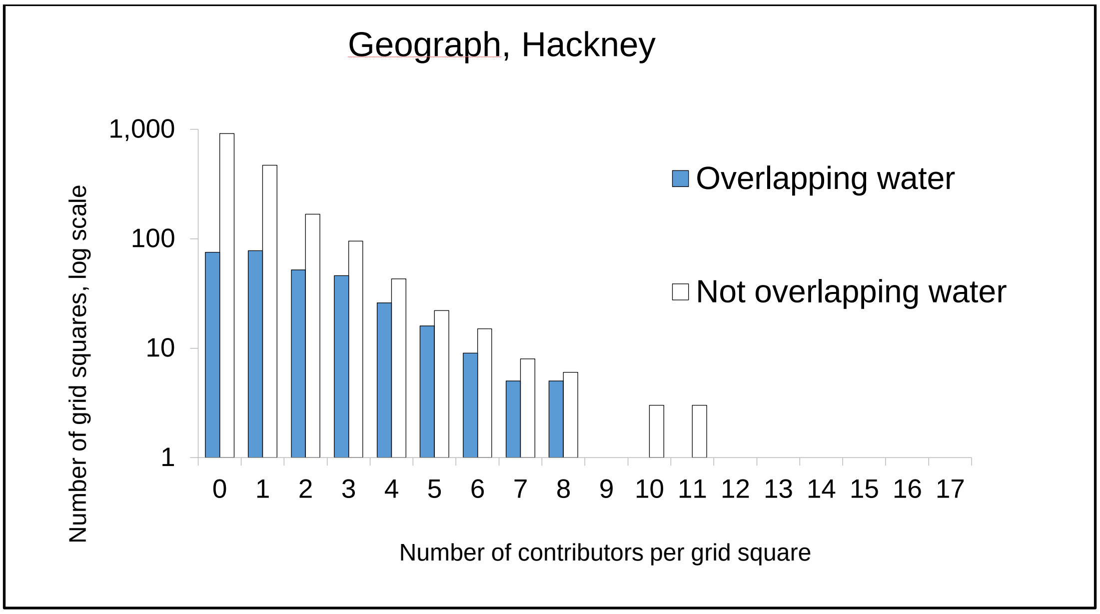

Figure 6-19: Geograph density of contributors per 100m grid cells in Hackney

6.8 Summary

This chapter analysed the use of geolocated photographic collections as a potential source of information for the detection of environmental spaces. The main project and a sub project proved that the information source used, and the methodology designed gave as result interesting insights; the photosharing projects have individualities and characteristics; those individualities nonetheless always stress areas of high significance as natural and cultural areas.

7 Environmental Citizen Science - biological recording projects

The analysis of environmental citizen science implied the analysis of three biological recording projects. Each of them is aggregations of communities, as explained in chapter four. In this chapter, the analysis involves data coming from three organisations. We have examined a range of projects that goes from iSpot, through iRecord and including GiGL. Contrarily to the previous projects that can be assumed as centralised, in this chapter there is the comparison among three sources of data that while focusing on the collection of environmental information by volunteers they play distinct roles. One is a service that supports data collection (iRecord), the second is a service that created a supportive community to help identify observations (iSpot), and the third is in charge of collecting, connecting, digitising, and developing the exploitation of all environmental observations ever recorded from a wide range of sources by an heterogenous community of observers (GiGL). Therefore, the current range of data sources includes two extremes. On one side, iSpot is like the centralised projects of the previous chapter. It has a platform that supports the correct identification of environmental observations. It has a community and a specific scope. On the other extreme, GiGL is a champion of data digitalisation and aggregation; it aggregates data whose large majority pre-existed the institution itself. It coordinated the uniformed digitisation and storage of data created from diverse communities, even hundreds of them. Every organisation considered in this chapter is a data holder, while GiGL has a very limited activity as a data collector. Therefore, all those organisations, while sharing similar purposes, embed a wide range of practices, motivation, and focus. The analysis that follow will be affected by this heterogeneity.

In the previous case study, we used the OECD methodology to classify 1 km BNG units. This filtering procedure has been introduced since from our literature review we found out that in urban areas photosharing collections focus almost exclusively on social and cultural features. Photosharing projects being about generic pictures they may include any sort of human visual engagement for their environment, hence they can also include artefacts but most often other humans regardless of the surrounding environment. Conversely environmental observations are inherently focused on wildlife and citizen science has been already documented as producing and caring cultural practices (Fish et al. 2016b). Therefore, in this chapter we did not classify and filter out BNG grids.

The baseline situation for this analysis is shown in the following sections. The first section (7.1) introduces the area of Greater London from a nature-centric point of view, then **Error! Reference source not found.** is a section showing the basic global statistics referring to the projects. Following, section 7.2 characterises every individual project. Then, section 7.3 is devoted to the mapping of the individual platforms. The following analysis section (7.4) begins with the analysis of the correlation between the three organisations (section 7.4.1). The local spatial analysis, in section 7.5, is followed by the identification of hotspots and outliers in section 7.5.1. The visual inspection on the characteristics of the places that are highlighted by the spatial statistical analysis (in section 7.5.2) anticipates the summary (section 7.6).

7.1 Nature and London

In the context of this work, this case study focuses on the wellbeing benefits that can be derived not only participating in citizen science environmental recording projects. The purpose of this project is to understand the places people choose to spend their spare time to engage with nature observations. Through the detection of the intensity of a cultural practice we aim to understands the features of the environmental spaces that enabled them.

Greater London has an enormous potential to deliver health and wellbeing benefits through nature. They have been already calculated in the reports aiming at the economic valuation of ES during both phases of the UK NEA. In almost all scenarios which don't rely only on market values Greater London is always the source and destination of an increasing number of visitors (Sen et al. 2011). The presence of the largest metropolitan area in the UK makes every improvement in London's environment an investment in welfare and general wellbeing with high returns as can be seen in (Bateman et al. 2014). The proximity and accessibility of large amount of population explain the importance of the investment in London's environment. To provide an idea of the extent of London potential for engagement with nature we list now data that has been collected and aggregated by GiGL(2016). According to GiGL roughly 47% of London is green and 2.5% is blue. Green areas include 33% of vegetated green space and 14% of vegetated garden private green space. Almost 40% of Greater London surface is open space and it is used in a wide range of ways as can be seen in Table 7-1, next page. This classification show how open-spaces are linked to several benefits to human wellbeing. Amongst the more natural areas several holds statutory protection status. There are 2 Special Protection Areas (SPAs), 3 Special Areas of Conservation (SACs), 2 National Nature Reserves (NNRs), 37 Sites of Special Scientific Interest (SSSIs)although seven are designated for their geology, and 143 Local Nature Reserves (LNRs).

Land Use	Area (ha)	Percentage
Parks And Gardens	9,207	5.77%
Natural and Semi-natural Urban Greenspace	8,859	5.56%
Green Corridors	5,671	3.56%
Outdoor Sports Facilities	10,718	6.72%
Amenity	6,575	4.12%
Children and Teenagers	72	0.05%
Allotments, Community Gardens And City Farms	995	0.62%
Cemeteries And Churchyards	1,390	0.87%
Other Urban Fringe	12,893	8.09%
Civic Spaces	74	0.05%
Other	3,063	1.92%
Unknown	2,601	1.63%
Total:	62,118	38.96%

Table 7-1:Open spaces in London (GiGL 2016)

But in the classification of Table 7-1 we have also less natural areas that are able to contribute to human wellbeing with a range of benefits linked to exercising, learning, and socialising. As already mentioned also private spaces and more specifically gardens can contribute to human wellbeing.

7.2 Basic global statistics of citizen science platforms and data preparation outcomes

Table 7-2 provides some basic statistics to compare the spatial and numeric extent of all data used in this chapter. Those are the original data as given us by the organisations. The number of grid cells for which records had been submitted at a 1 km resolution or better was GiGL 1,780; iSpot 899; iRecord 1,071, corresponding to 91%, 46% and 55% respectively of the total 1,950 cells in Greater London. The total numbers of environmental observations are overshadowed by the GiGL data collection, which represented almost 98% of the environmental observations collected (1,102,746) and almost 74% of contributors (6,705 citizen scientists). From a geographical point of view, GiGL covers almost 91% of all the grid units in the area compared to 55% and 46% for iRecord and iSpot. Despite having fewer total records than iSpot, iRecord's volunteers covered a greater geographic area than iSpot's.

	GiGL	iRecord	iSpot	All projects
Total observations	1,102,746	12,899[80]	13,352	1,128,997
Percentage over total observations	**97.67%**	1.14%	1.18%	
Total contributors	6,705	1,498	887	9,090[81]
Percentage over total contributors	73.76%	**16.48%**	**9.76%**	
Maximum observations per contributor	**45,299**	486	2,207	
Average contributors per km^2	3.44	0.77	0.45	4.53
Surveys	176	29	1	206

Table 7-2:Basic statistics of citizen science organisations in case study area. In bold noteworthy data

As has been shown in chapter 5, several data preparation processes have been implemented over GiGL and iRecord data collections, to filter out observations which we deemed unfit for our purposes because they had not been associated to a unique identifier for the observer or because their spatial precision is worse than 1km. iSpot data came after a very stringent selection process implemented by the organisation itself. Its excessively filtrated dataset deprived our analysis of the data about collaboration patterns and this might cause an underestimation in some cases. Moreover, the filtering imposed by iSpot on its data stripped of observations that did not meet strict biodiversity recording quality criteria, obliterating from the data we received the locations of the observations with associated uncertain taxonomic identification.

Geographic harmonisation was requested since iSpot included data referred to two precision levels of the BNG. To achieve this, we added a field to the table provided us and a Python query has been developed to calculate the required BNG coordinates. As shown in chapter 5, iRecord data had several levels of complexity. To prepare data for our analysis, we had to remove some observations and implement a geographic harmonisation. The removal of observations mostly concerned the ones that were shared with iSpot and GiGL, and the ones with a lower precision then the one on $1km^2$. After removing the surveys shared with GiGL and iSpot and omitting the observations with

[80] Some of the observations were contributed to iSpot and GiGL therefore in the following calculation the remaining 992 observations will be considered

[81] In the unified database, they are 9,081 but maybe some contributed to more than one project. For this calculation, we disjointed them. The difference is 0.1%

precision level lower than 1,000 meters iRecord data collection in terms of precision gives the distribution of values as can be seen on Table 7-3, below:

Precision	Count	%
1,000	2,715	26.21
100	1,886	18.21
50	2,662	25.70
10	2,693	26.00
1	402	3.88
Total	10,358	100

Table 7-3: iRecord precision levels of analysis dataset after data preparation

Precision 1,000 is now the dominant class even only for 22 observations. After the removal of double counted surveys and referring all observations to the 1km BNG grid, the further analysis has been conducted with a GIS. After the cleaning of duplicate recordings and the selection process of observations with a spatial precision of 1km or more as described in chapter 5 the resulting datasets had the general stats as in the following Table 7-4. This includes the updated values of Table 7-2 and in addition some more lines helpful to understand the relation between the three platforms. We see how the relations between the three organisations is unchanged and that one of the most interesting elements of this table is the fact that the two digital platforms (iSpot and iRecord) can attract a larger number of contributors, who are less productive in comparison with GiGL, that represent a collection or digitisation of mainly not natively digital projects.

	GiGL	iRecord	iSpot	All projects combined
Total observations	1,014,769	10,358	13,352	1,038,479
Percentage over total observations	**97.71%**	1.00%	1.29%	
Total contributors	6,705	1,237	887	8,829
Percentage over total contributors	75.94 %	**14.01%**	**10.05%**	
Maximum observations per contributor	45,299	486	2,207	
Surveys	176	28	1	205
Maximum observations per km^2	58,999	332	693	59,481
Grid squares with observations (1,950 all grids)	1,744	976	887	1,773
Coverage over total grids	89.44%	50.05%	45.49%	90.92%
Average observations per contributor	151.34	8.37	15.05	
Average contributors per km^2	3.44	0.63	0.45	4.53

Table 7-4:Basic statistics of citizen science organisations in case study area after data preparation. In bold noteworthy data

The mutual relationship remains intact, with GiGL collection overshadowing the other two organisations in all categories and other small but uninfluencing adjustments that leave the previous evaluations holding their value.

7.3 Spatial patterns of individual projects

To start the analysis, we first look at spatial patterns of the individual projects. The different projects have quite different number of contributors, thus maps depicting the total numbers of contributors for each project, while interesting, are less useful in the development of this work that aims to compare the projects. Therefore, to map contributions we derived the grid contributor ratio (Cr). This is the percentage of contributors for each grid unit compared with the total number of contributors to the specific project in the study area. As an example, in GiGL there are 6,705 contributors in the study area, while in the BNG grid TQ2277 covering partially the London wetland centre there are 1,263 contributors. The GiGL contributor ratio for grid TQ2277 is hence Cr=1,257*100/6,705. Therefore, the GiGL contributor ratio Cr for the BNG grid unit

TQ2277 is 18.75, as expected in a place with several visits and ongoing educational activities we have a quite high number of people contributing. As another measurement of biodiversity appreciation, we also calculated for every group of surveys the volunteer productivity (Vp) This is the mean number of records per volunteer operating in the same grid square.

7.3.1 Contributor ratio

The distribution of contributor ratio varies considerably across grid cells. As can be seen in the following maps. GiGL is the richest of all our collections and has several areas with a large relative number of volunteers collecting environmental records as in Figure 7-1.

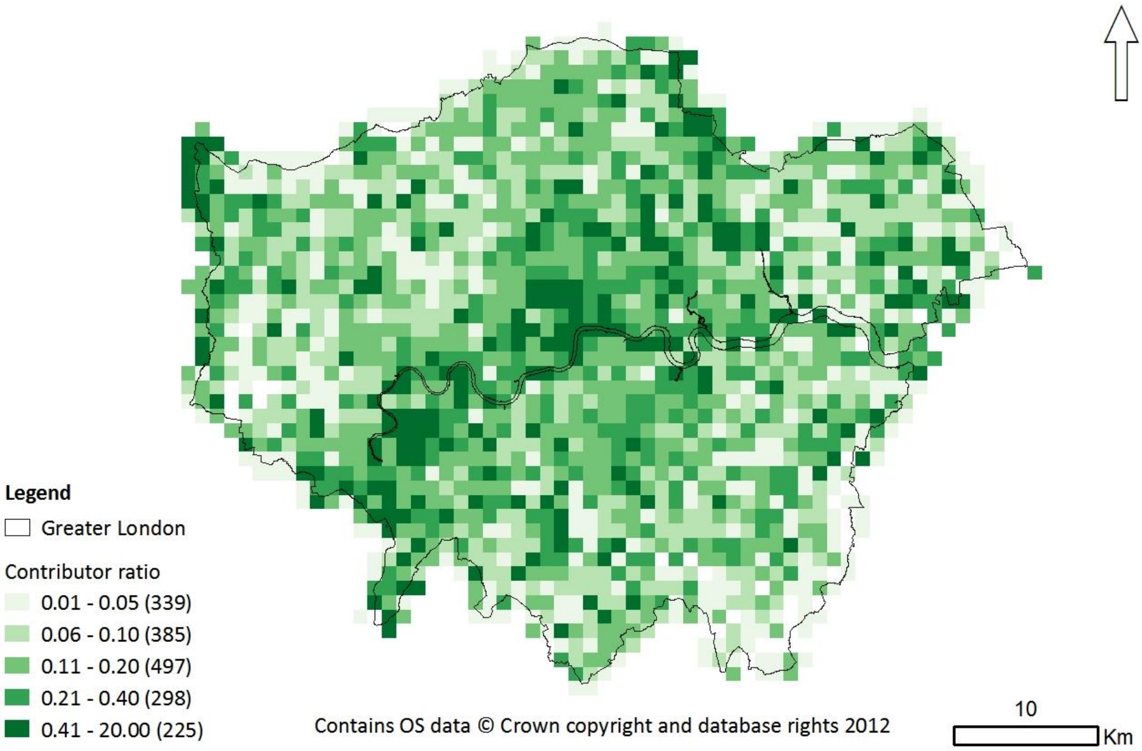

Figure 7-1: GiGL Contributor ratio. Spatial distribution of contributors to GiGL as a percentage of all contributors to GiGL in Greater London. The parentheses include the number of grid units in the range

GiGL surveys have volunteers that with their activity covered almost all Greater London. Areas with high number of contributes can be found scattered all over Greater London, apart from the south-east area. This part belonged to Kent until Greater London was created in 1965. Remembering the peculiar nature of GiGL, it is because Kent's Local Environmental Records Centre at the time of the collection of data for this work did not share all their observations with GiGL[82]. There are clear patterns of higher numbers of volunteers contributing where there are the Royal Parks, Richmond Park, Greenwich Park, along the Thames and most of the open areas whether they are parks, play areas, open fields. When compared with the maps showing iSpot and iRecord contributors' ratio

[82] GiGL employees gave us this explanation.

(Figure 7-2 and Figure 7-3 respectively), we see how the two most recent projects have similar distributions amongst them with a clear prevalence of inner London areas.

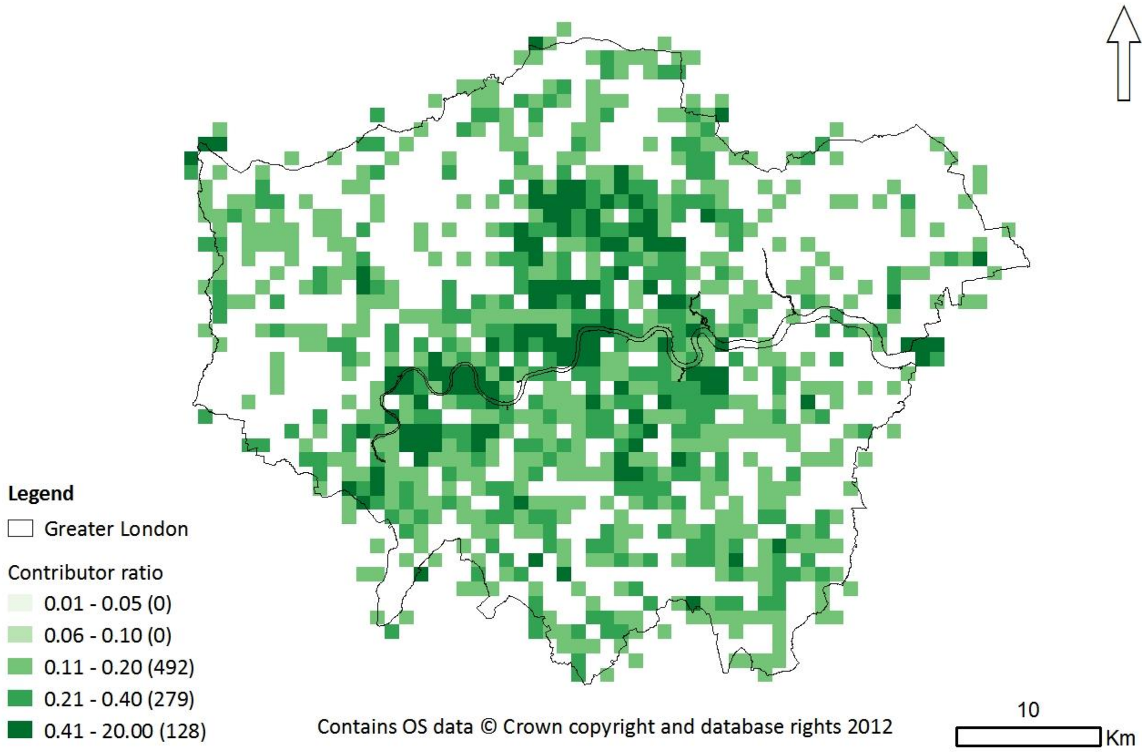

Figure 7-2:iSpot Contributor ratio. Spatial distribution of contributors to iSpot as a percentage of all contributors to iSpot in Greater London. The parentheses include the number of grid units in the range

For iSpot we have a distribution of Cr that is very similar to the one of GiGL for the area surrounding the River Thames from east up towards Greenwich Park. The spatial coverage has areas with considerable contributor ratio that were not spotted in GiGL. They are in Harringay and the Lea valley where there are the long strings of reservoirs and the Hackney Marshes in north east London. For iRecord, the most popular areas in terms of contributor ratio covers the already mentioned green areas mostly along the River Thames. iRecord has some areas of higher Cr in Southwark that did not emerge mapping GiGL and iSpot.

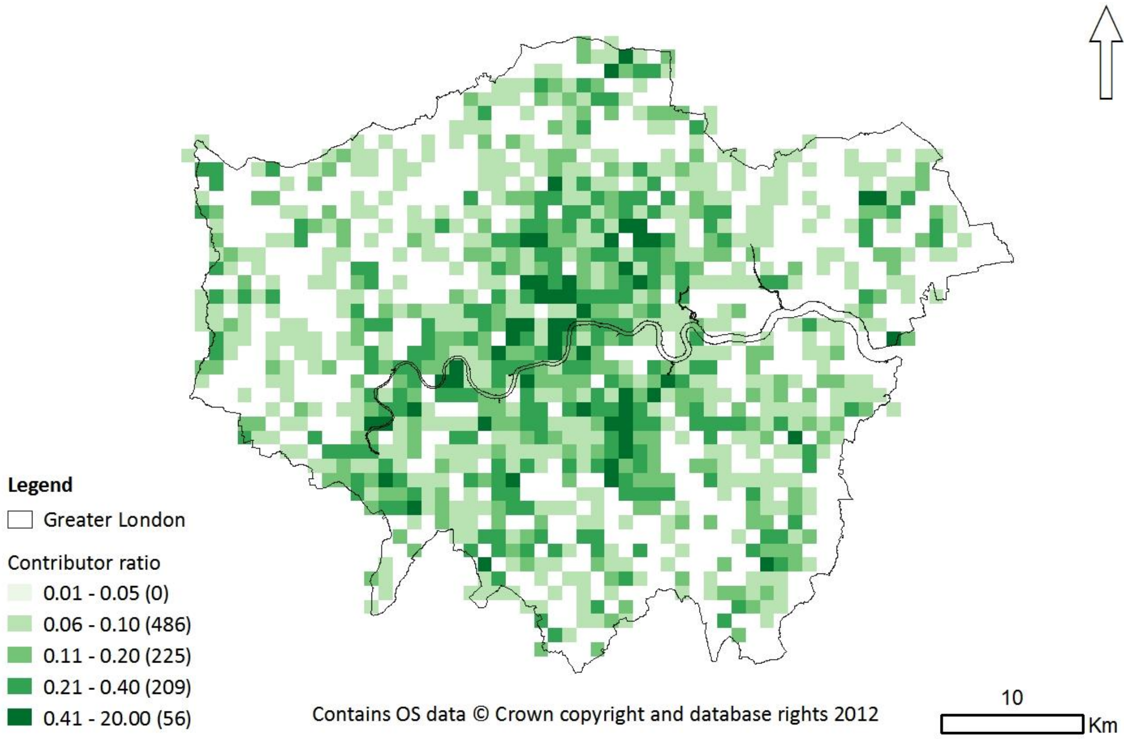

Figure 7-3: iRecord Contributor ratio. Spatial distribution of contributors to iRecord as a percentage of all contributors to iRecord in Greater London. The parentheses include the number of grid units in the range

7.3.2Volunteer productivity

While contributor ratio helps comparing the three projects we calculated also the spatial distribution of volunteer productivity or Vp (mean number of records per volunteer) recorded. The distribution of volunteer productivity recorded varies considerably across grid cells. In terms of volunteer productivity GiGL shows (Figure 7-4) an almost homogeneous behaviour on all Greater London with some grids of main productivity. What is more striking here is the disappearance of some areas that are characterised reversely by high contributor ratio; the Royal parks and all the areas that were high in contributors' ratio are comparatively low in productivity. Figure 7-4 shows in darker colours places where a very limited number of citizen scientists, or even one only, are submitting enormous amounts of observations.

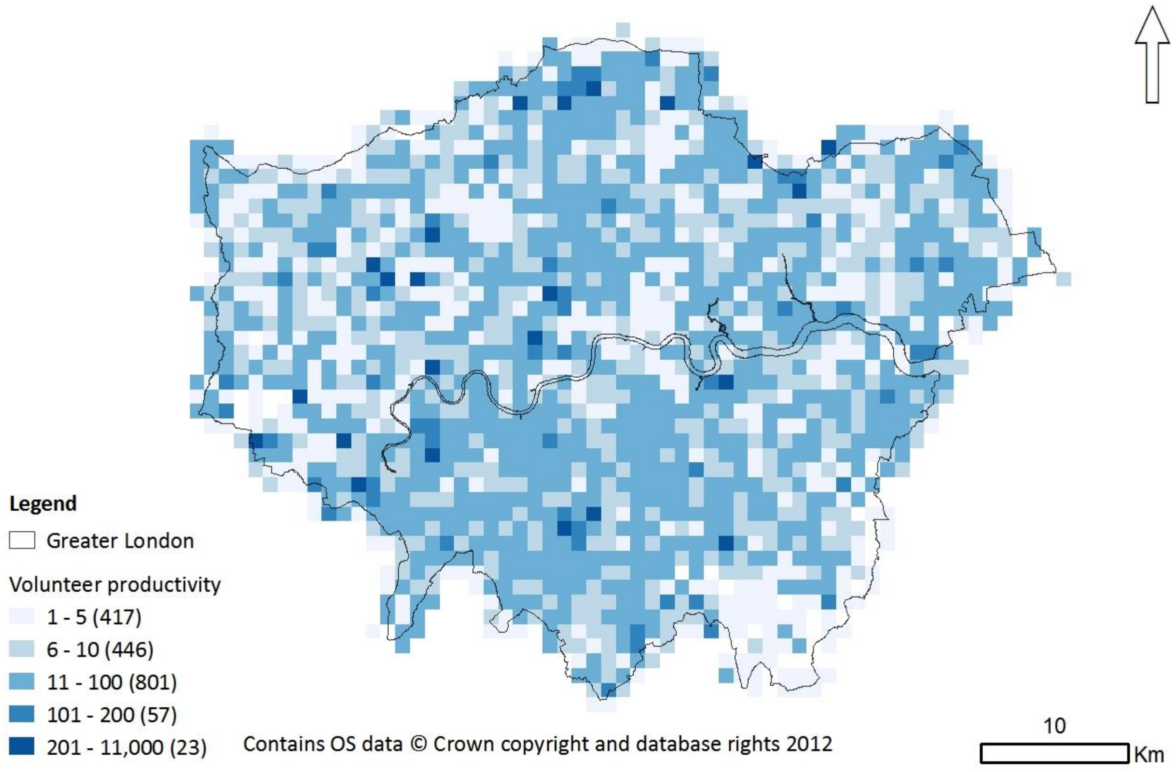

Figure 7-4: GiGL volunteer productivity. Number of observations per volunteer for each 1 km OSGB grid in Greater London

There are two volunteers that have been able to contribute with tens of thousands of observations of moths from their gardens as seen in (Table 5-4).

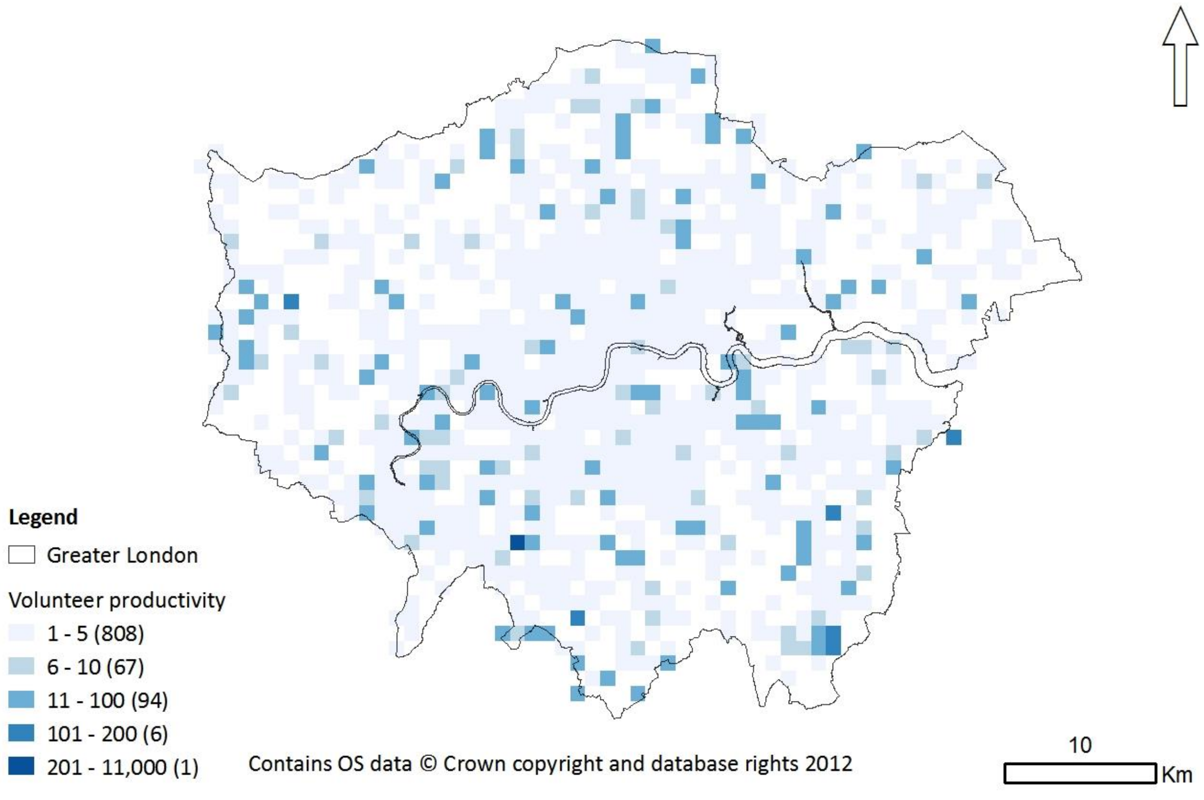

Figure 7-5:iRecord volunteer productivity. Number of observations per volunteer for each 1 km OSGB grid in Greater London

The examination of volunteer productivity was designed to highlight areas where volunteers are more committed, and for GiGL it revealed that home locations and gardens play a significant role in engagement for cultural practices. A similar quite patchy image is given by the volunteer productivities for the other two projects.

The Vp of iRecord contributors in Figure 7-5 does not show visually any significant clue but the grids with higher values of volunteer productivity are also here characterised by the fact that a highly committed private citizen can produce hundreds of observations from private gardens or private land and woodlands. Vp can highlight the high commitment of few or even one only contributor in an area. The small cluster in Orpington, south east of Greater London close to Kent is as an example due to one only volunteer that submitted also several observations from the UCL main open area.

Figure 7-6 shows the volunteer productivity for iSpot; also in this case the main hotspots are created by very committed volunteers (usually one for the specific grid) who collect a considerable amount of environmental observations in the same area.

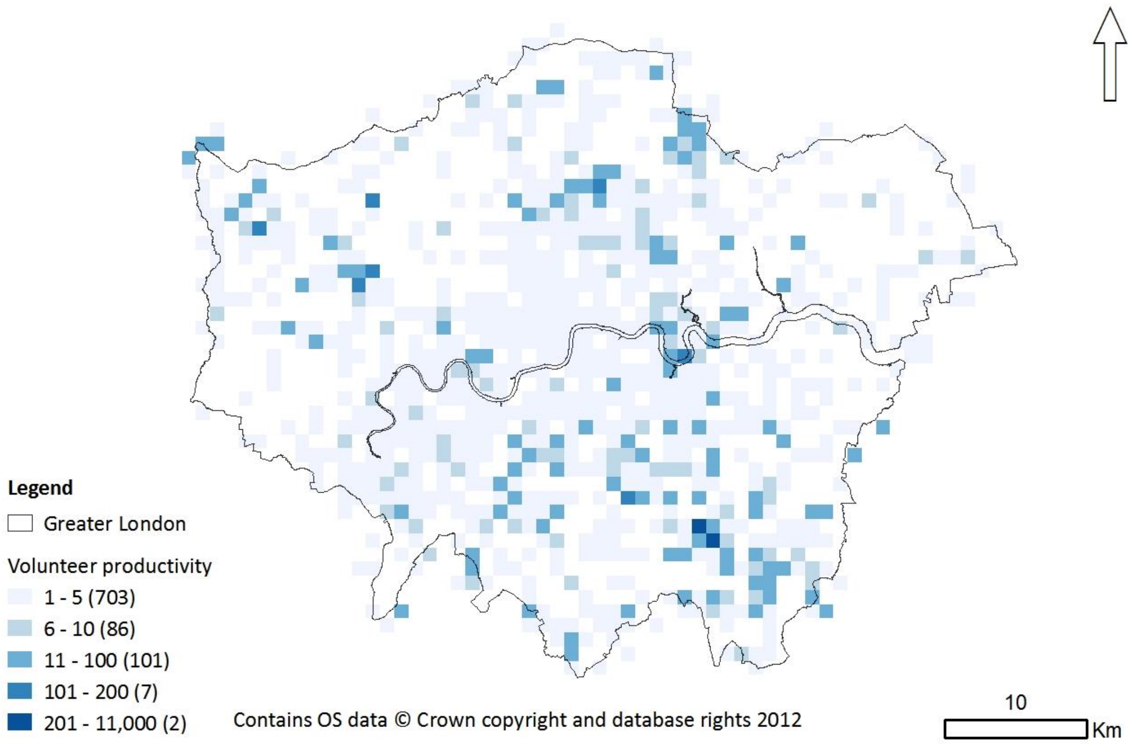

Figure 7-6:iSpot volunteer productivity. Number of observations per volunteer for each 1 km OSGB grid in Greater London

The cell in which the London Wetland Centre is sited contains both the highest number of records and the highest number of volunteers. It is interesting to note, however, that despite attracting many volunteers, the productivity in the grid cell containing the London Wetland Centre was comparatively low.

7.4 Spatial analysis

In the following subsections, two levels of analysis are carried out. At first, the analysis of the relation between projects on a grid by grid basis will ignore the topological relations between grids. In the following level an analysis of the spatial patterns, hotspots and outliers using all combined platforms will also take into consideration the topological component. The correlation between projects, considering every grid unit separately, is investigated through the development of scatterplot diagrams and rank Spearman's statistics (in section 7.4.1). The global analysis for local clusters and outliers (7.5.1) uses the local Moran's I method.

7.4.1 Relationship between projects

To compare the three platforms in a grid by grid basis using the number of contributors to each project as a comparison metric, two methods have been used: the design of scatterplot diagrams and the estimation of the Spearman's rank test. It is run through couples of projects and gives ρ as a measure of the correlation between the two ranked collections and it measures the statistical significance of the result obtained. The scatterplot method provides a visual impression of the correlation grid by grid. To design the scatterplot, the contributor ratio (Cr) was used. The visual analysis of scatterplots, while it might lead to misleading interpretations due to overlapping points, is very helpful to find outliers. For the calculation of the Spearman's rank test, the use of this ratio is not needed since this statistical method ranks the values for every project and compares the ranks but does not consider the relative magnitudes.

The spatial distributions of volunteer productivity were significantly correlated between datasets: GiGL and iSpot ρ (1778) = 0.235, $p < 0.001$; GiGL and iRecord: ρ (1778) = 0.271, $p < 0.001$; iSpot and iRecord: ρ (1069) = 0.333, $p < 0.001$.

The correlation coefficients are not particularly high, but the results suggest that the same grid cells are broadly more or less productive for all datasets for both number of records per volunteer. Although the total number of records per grid cell summed across datasets was not uniform across the study area the spatial distribution of records was significantly correlated between datasets (Spearman's Rank correlation); GiGL and iSpot: ρ (1778) = 0.373, $p < 0.001$; GiGL and iRecord ρ (1778) = 0.402, $p < 0.001$; iSpot and iRecord: ρ (1069) = 0.544, $p < 0.001$. The total number of volunteers recording in each grid cell was also not uniform. The number of records per grid cell was highly correlated with the number of volunteers, ρ (1801) = 0.795, $p < 0.001$. The number of contributors per grid cell was quite mildly correlated among the three platforms: GiGL and iSpot ρ=0.435, $p < 0.001$; iSpot and iRecord ρ=0.4310229, $p<0.001$; iRecord and GiGL ρ=0.422, $p<0.001$. The scatterplot diagrams, given the mild results of the Spearman Rank test, do not give a strong specific image of the spatial correlation between organisations. But as explained in the previous chapters, while they cannot be very effective for correlations with elevated elements, they are very useful once we focus on the detection and characterisation of outliers.

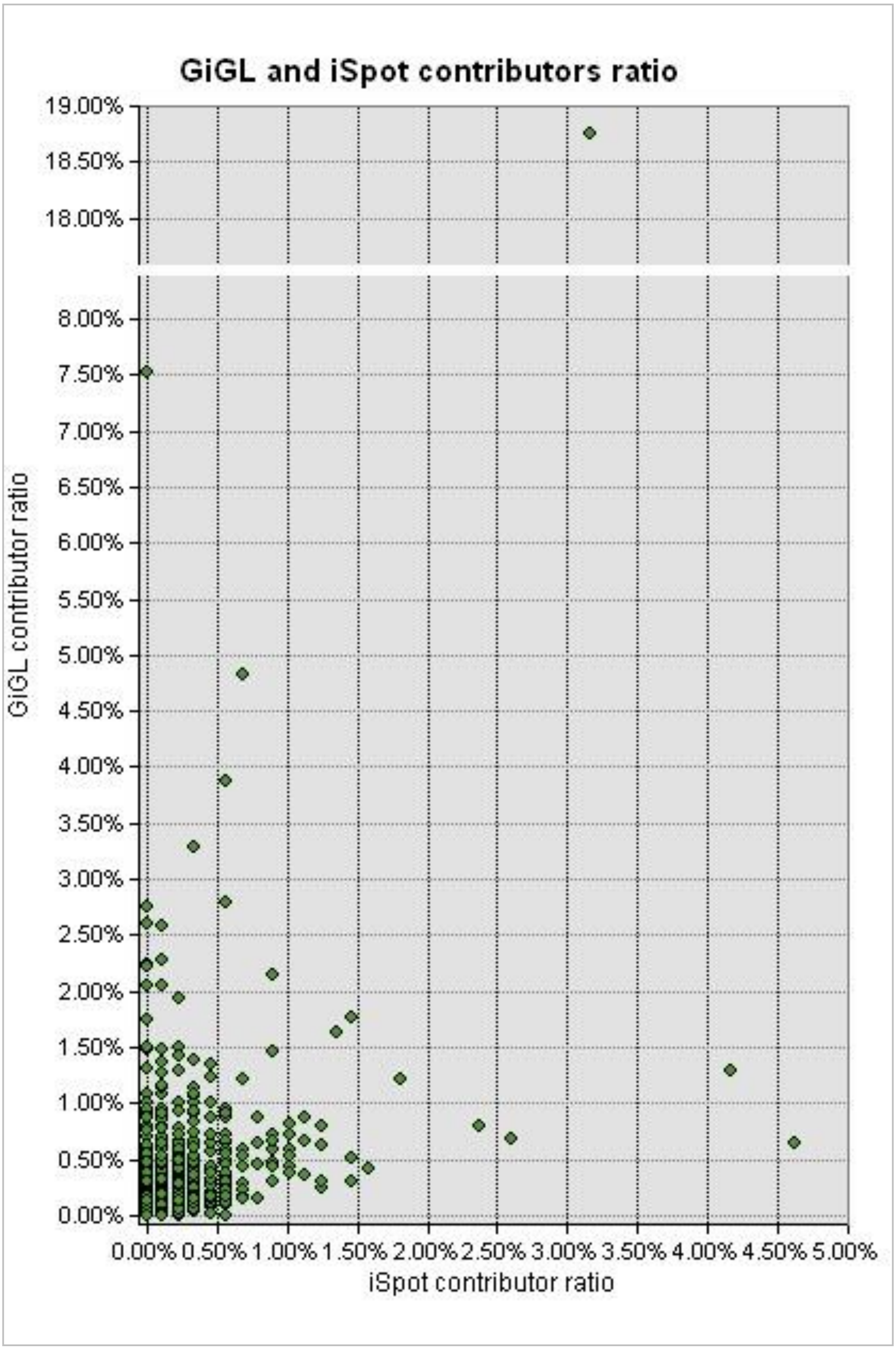

Figure 7-7: GiGL and iSpot contributor ratio scatterplot

The scatterplot diagram between GiGL and iSpot Cr in Figure 7-7, shows how important the London Wetland Centre[83] and the River Thames are for both projects. Amongst the five main outliers for both projects, we have two grids including the London Wetland Centre. Especially the grid (TQ2277), which is the top performing grid unit for GiGL, while is the third for iSpot. Almost all the other outliers have diverging values with relatively high values for GiGL's Cr, accompanied by low Cr values for iSpot. The second highest Cr value for GiGL (TQ5280) including the Rainham Marshes in the borough of Havering, a wet area close to the Thames, is even located in a grid where there are no iSpot observations at all. While looking at the grids with significant Cr for iSpot but relatively low values for GiGL, we have another grid covering the London Wetland Centre (TQ2276) and St. James' Park and Westminster Abbey in Westminster. The grid unit including the River Thames and the back part of Westminster Abbey also among outliers with significant contribution from iSpot (TQ3079). The neighbouring grids with high significance in one or another project witness how the reliance on one platform only can give a partial picture of a phenomenon. This picture is enriched by the aggregation of platforms and the geostatistical analysis. Another outlier at a closer look seems not really fit for environmental recordings. It is the grid TQ3281. It covers the Barbican centre, parts of St Paul's Cathedral's gardens and a highly urbanised area in the City of London up to Leadenhal Market. Focusing on both statistics and original geospatial data we have in this grid 384 observations mainly coming from the iSpot dataset[84]. The iSpot dataset contributed with 56 but several of them are misplaced since we have here observations with locational descriptions that refer them to places in Kent[85], South Gloucestershire[86], Devon[87], and other areas of London out of this grid unit. We can conclude that the geolocalisation of observations is wrong and they do not relate to this grid unit. This proves how important is the motivation to collect and at the end to manipulate geographic information which is not the main aim of the data collection.

In Figure 7-8 we have the scatterplot for the relation between iRecord and iSpot. Apart from the most evident iSpot outliers mentioned above, we can focus here on the five grid units where the contributor ratio of the two projects is not too dissimilar where iSpot Cr>0.5 AND iRecord Cr>0.5.

[83] http://www.wwt.org.uk/wetland-centres/london/

[84] Going to see some comments I think they misplaced some observations of LNHS 2010 especially from Kent. Some observations have comments about the surrounding environment and report a woodland which does not exist in this part of London.

[85] 16 observations with location_name set as Somerfield Rd, Maidstone

[86] 1 observation with location_name set as M4, Bradley Stoke

[87] 1 observation with location_name set as Mansands Ln, Kingswear

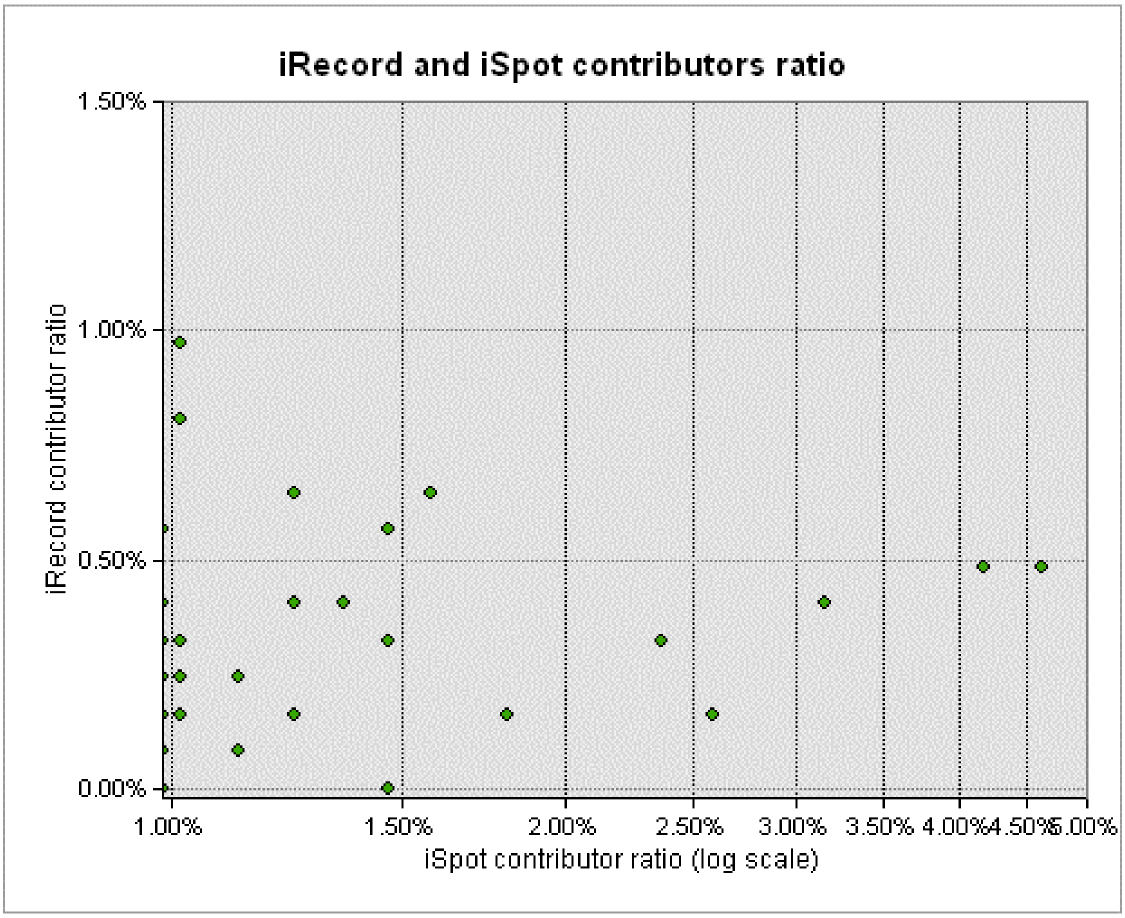

Figure 7-8:iRecord and iSpot contributor ratio scatterplot

Those areas include parts of royal parks (Regent's and Hyde park) and surrounding areas, Hampstead Heath, and the Stoke Newington area around the Abney Park Cemetery where observations are quite widespread and not confined only to the open spaces.

Another noticeable outlier covers a part of the Lea Valley where the Walthamstow Marshes meet the Springfield Park: here, most of the observations are along the Lea River. The only grid south of the River Thames here overlaps with Syndenham Hill Wood and Dulwich Wood (TQ3472). iRecord has a particular spatial aggregation of grids covering the southernmost section of the Borough of Camden (TQ2882, TQ2883) which includes part of Regent's Park but also the grid unit TQ2982 which includes also UCL's main quad where an individual who contributed to the creation of some iRecord spatial patterns in terms of Volunteer productivity in Orpington Borough, giving also here a noticeable contribute. In Figure 7-9, we have the scatterplot diagram that compares the contributor ratio of GiGL and iRecord. The mild correlation between the two data collections is visible here while in terms of outliers they have been already described in the sections above. A grid unit not mentioned earlier covers the area of Islington, including all of King's Cross station and the mostly residential and office areas in the

Pentonville area. Here several GiGL observations are made in everyday environments, walls, pavements together with some recordings erroneously geolocated in this grid while actually recorded in a small nature reserve behind King's Cross Station[88]. The analysis

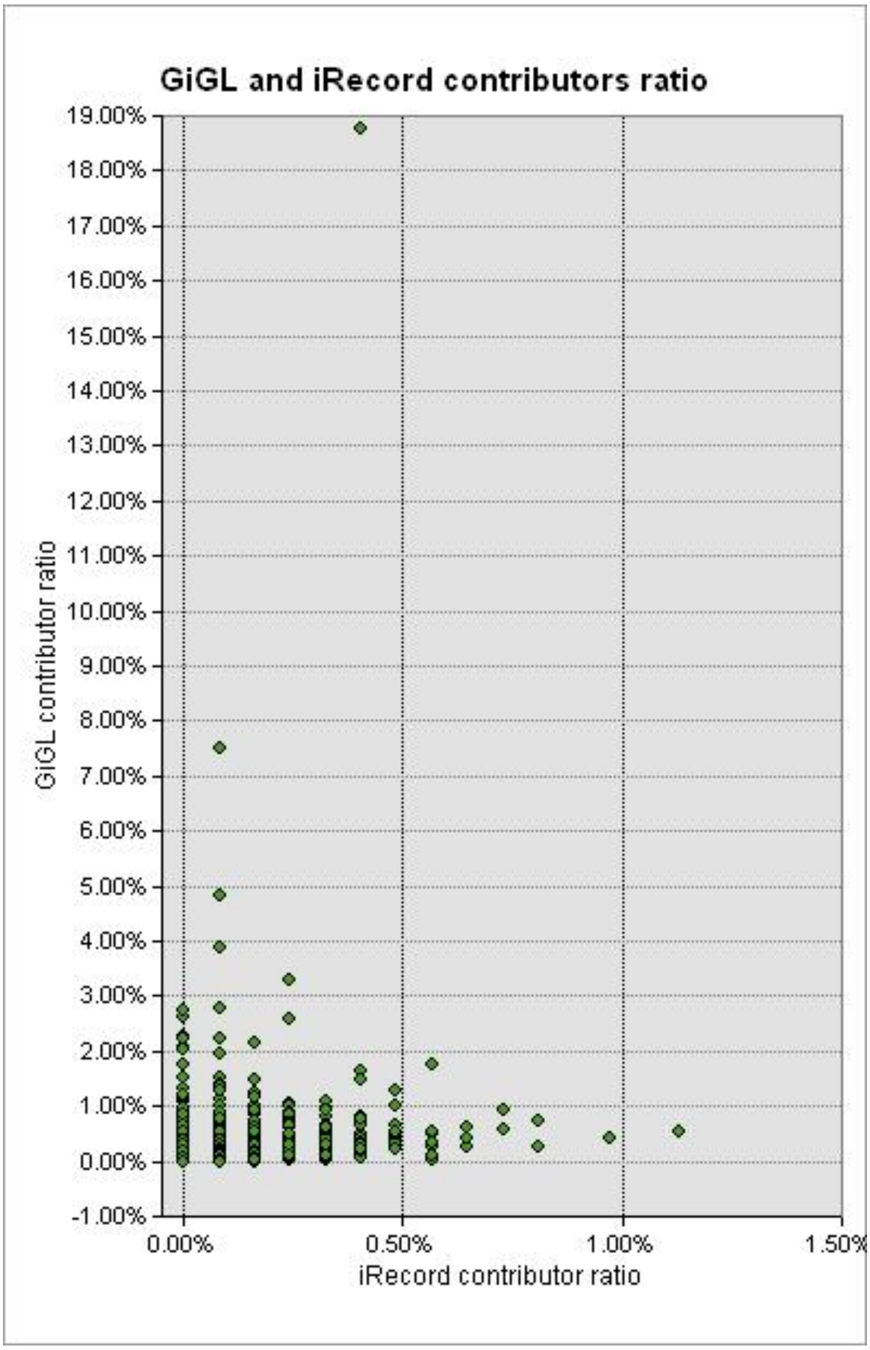

Figure 7-9: GiGL and iRecord contributor ratio scatterplot diagram

of the data at the individual cell level helped us to stress some interesting outcomes. We noted that 9 of the 10 cells with the highest number of records and all the 10 cells with the highest number of volunteers contained open water ('bluespace'), suggesting that such sites are particularly popular with recorders.

[88] Carmley Street Natural Park (http://www.wildlondon.org.uk/reserves/camley-street-natural-park) more correctly falls inside the grid unit TQ2983

7.5 Spatial statistics

This section is devoted to the identification of environmental spaces that are statistically and spatially relevant. The identification of grid units that individually or in spatial aggregations have values that are in stark contrast with the values of surrounding grids have been found using the values given by the total amount of contributors to all citizen science projects.

The relation between projects as presented in Figure 7-10, shows how the three platforms we used have such large differences in terms of number of observations and number of volunteers that considering the simple addition of contributors and then studying their distribution will equate most of the time to the analysis of GiGL volunteers' distribution as shown in Figure 7-10, below.

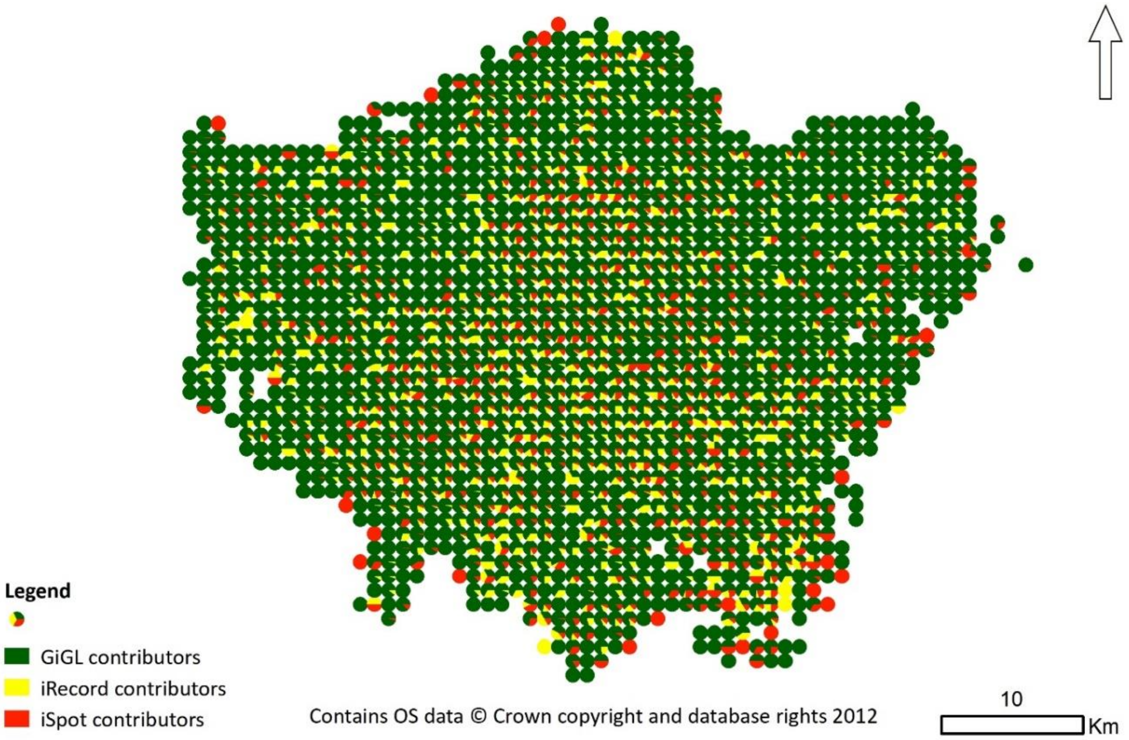

Figure 7-10: Proportion of contributors amongst Citizen Science organisations

It is noteworthy that the green colour associated to GiGL overshadows all other organisations for the vast majority of Greater London. While the summing up of all contributors can therefore be seen as the creation of a bias that favours GiGL, we must keep in mind that GiGL data is a collection of 176 surveys run by several heterogeneous organisations, as well as the fact that iRecord is a more coordinated family of 29 surveys built around technological affordances.

Willing to characterise the overall phenomenon of citizen science as a cultural practice that help us to detect environmental spaces, we will then sum up all observations collected by our three organisations, reminding that in such a way we are summing up 176 GiGL surveys, 29 iRecord surveys and iSpot. A study comparable to the other case studies analysed in this work might be run considering all 206 surveys individually.

The map of total number of volunteers per grid units is shown in Figure 7-1. It is evident how some grid units particularly exceed the values registered in neighbouring grid units as well as some spatial trends.

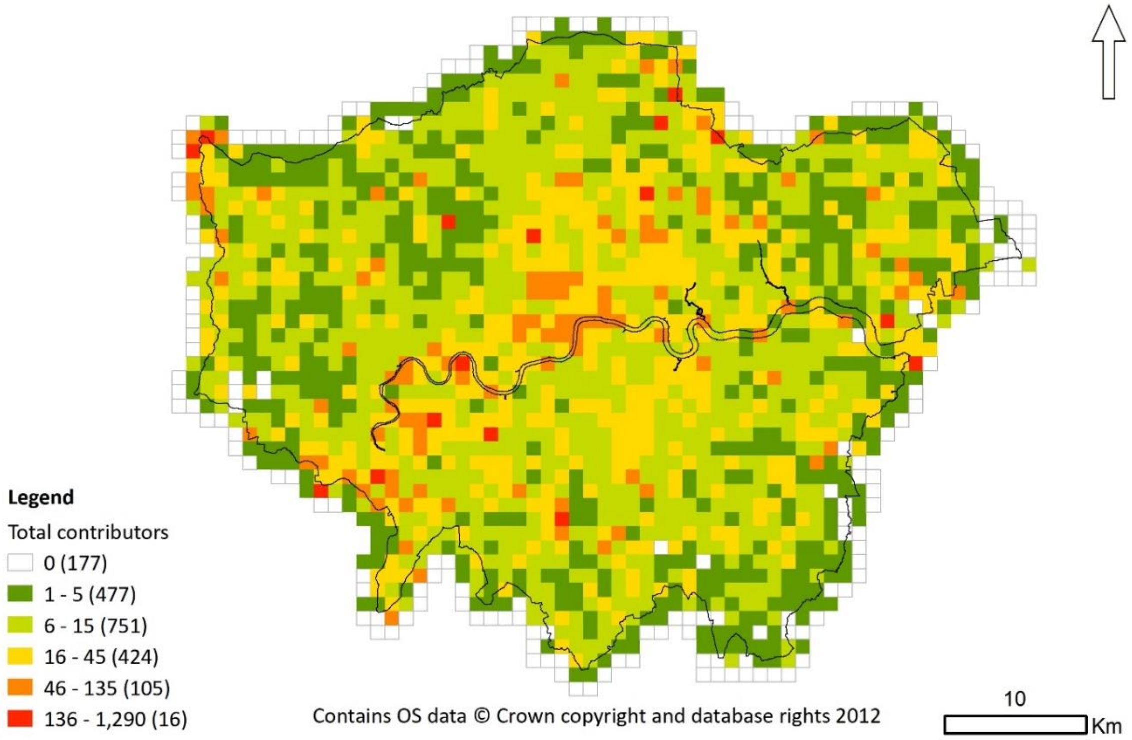

Figure 7-11:Total contributors to citizen science biological recordings per BNG grid unit

We ran several tests to find the most appropriate methodological arrangement to stress the evident spatial clusters as can be intuitively seen in Figure 7-11. In terms of the size of the analytical window to use, the most effective distance range resulted to be 10 km. Then the most significant Moran's I global spatial autocorrelation index, calculated over the distribution of total contributors for every grid unit using a linear decay of influence[89] (I = 0.034380, z=17.756409 $\rho < 0.0001$), reveals that there is largely less than 1% probability that the autocorrelation has been generated by a random distribution. While previously the use of the contributor ratio was functional to the detection of correlations, in this case the number of contributors is preferred as in this case every individual contributing is considered regardless of the project to which they contributed[90].

7.5.1Spatial trends: clusters and outliers

To detect spatially statistically significant grid units, such as spatial clusters and outliers, this study used Moran's I local indicator of spatial association (LISA; Anselin 1995) as implemented in ArcGIS 10.2.2 spatial statistical tools. The computation highlighted 120 grid units that constitute in clusters, or individually a statistically relevant spatial unit,

[89] Inverse distance

[90] The sum of contributor ratios from the three data holders provided even better indicators of autocorrelation but that number was considered a false positive for the reasons explained above (we have more than 200 surveys and not just three communities).

when analysing the distribution of the number of contributors. In Figure 7-12 we have them plotted on the Greater London Map and against all Green and main Blue spaces[91]. The grids with high values in clusters or in isolation define 44 areas of different sizes. The largest five of them are all in areas along the rivers Thames or Lea. Several other clusters and hotspots are in areas which include parks, but several large green areas do not emerge as a significant high value clusters. We have also three grid units which are particularly poorer in terms of contributions to citizen science compared to neighbouring areas. In the following section, we are going to identify and characterise some of the areas that have been identified through the spatial statistical analysis.

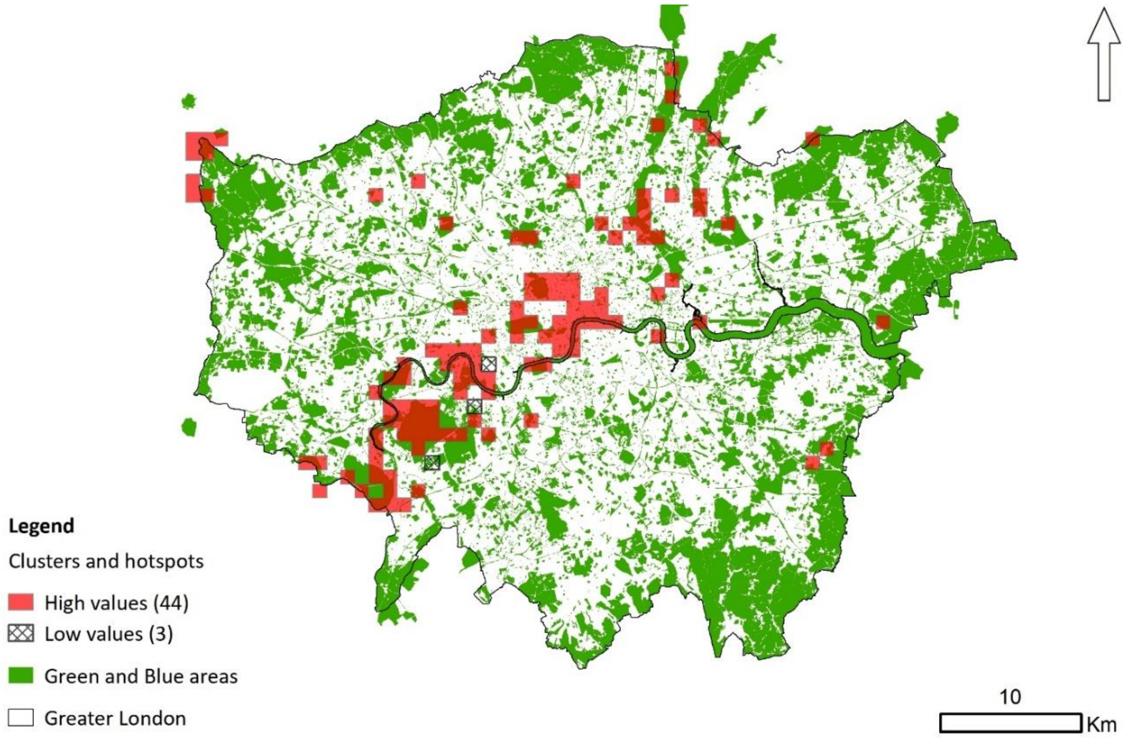

Figure 7-12: Citizen Science clusters and hotspots

7.5.2Visual inspection: sites of environmental connectedness

Amongst the 44 clusters and hotspots characterised by high values of contributors, there are some clusters which are particularly large. Some of them can be grouped in regional areas. The main visible trend between clusters and hotspots is the one that follows the River Thames. The River Thames touches 12 clusters. Ten of them are situated in the western half of the study area, beginning with the first westernmost gateway of the River Thames to Greater London in correspondence with Bushy Park in Kingston upon Thames. This is an interesting pattern and it may be due to several reasons[92]. The

[91] Data on green and blue spaces kindly contributed by GiGL

[92] The most obvious is related to the presence of large green areas facing or in the immediate proximities of the River Thames. More historical ecological and cultural explanations are worth investigating.

pattern ends its eastwards development with a regional group of clusters made up by three clusters that cover most of central London. This group of clusters includes also the larger cluster found in this work. It is 22 km^2 and covers the River Thames from Pimlico downstream up to the Tower of London; on the north side of the river it expands to cover all the central area of London where the congestion Charge is applied with the exclusion of Mayfair and Fitzrovia and the southernmost part of Marylebone[93] in the west, and with the exclusion of Clerkenwell and Fitzrovia in the east. This large area also includes the Regent's Park and all the grids at the north of Euston Road. The central regional group of clusters includes also two other small clusters. One which includes most of Hyde Park and the grid mostly is characterised by the presence of Paddington station. In this last grid, the observations are mainly bird sightings and are not concentrated in a specific area. The Royal Parks, meanwhile, for which specific conservation groups and events are created, are almost entirely involved in the definition of areas characterised by high volumes of contributors to citizen science. We also associated to those two previous clusters, the one covering the River Thames and Battersea Park. The Thames trend includes also all grids that cover The London Wetland Centre, the Castelnau and Barnes areas and all grids with parks and reserves near the River Thames. The second largest cluster of the study area is also part of this trend and it covers entirely Richmond Park and connects with part of Wimbledon common.

Another local agglomeration of clusters is the one that follows the River Lea valley with a main cluster of 6 grid units and a couple of individual grids very close to this in Hackney at the west side and another in Walthamstow in the other side. The Hackney side grids includes one grid already mentioned for its high number of observations in the Abney Park Cemetery[94] (TQ3386). Particularly rich in volunteers is also the westernmost part of Greater London, facing Hertfordshire, with several grids that fall inside the green belt area that includes a much larger green space, the Colne Valley with several green and blue spaces. A nice case of engagement is registered in Harrow where a survey part of GiGL collection, the Harrow Schools Survey 2007, with the participation of school pupils[95] determined two grid units that stands out to the surrounding ones thanks to that survey. We have two grids one with 465 observations and another with 475 in those grid units and 2,169 in total in Greater London having as recorder a *schoolkid*. This suggests how initiatives and organisation play a relevant role in the delivery of engagement for nature. The Brent reservoir attracted a lot of volunteers to record mainly birds. This grid

[93] Roughly all the area with postcode W1

[94] Several recordings were erroneously geolocated out from the actual cemetery but inside the grid unit that covers most of it.

[95] Several observations have as recorder "Name (Surname) Harrow-Schoolkid"

unit stands out even in the presence of a lot of observations recorded with collective rather than individual identification names[96].

The spatial statistical test also highlighted three grid units that constitute a significant downfall compared to bordering grid units. Two are in the SW postcode area with one in Fulham and the second in Wandsworth, a third one is in Kingston upon Thames. The first, with only 5 contributors, is characterised (TQ2477) by a very limited amount of sport/recreation oriented public spaces and is largely composed by a low density urban development and some tower blocks. In Wandsworth (TQ2374), just 3 km away from the previous mentioned grid unit, we have a very similar area where 156 observations have been submitted by 4 contributors but amongst them we have a collective identification ("Member of Public"). We have also some recordings that while in the GIS dataset provided by GiGL are placed in this BNG grid unit they have associated another grid in the specific field[97]. The grid in Kingston is characterised by extensive open spaces: two golf courses cover roughly 33% of the surface of the grid unit; and a low-density neighbourhood with plenty of trees, swimming pools, and private gardens characterise the rest of the area.

7.6 Summary

In this chapter to detect and map CES, we analysed environmental Citizen Science and in particular biological recording projects. The methodology has been tailored to the datasets provided which amounted to the collections of several surveys conducted by a wide range of communities over a long time. Two organisations, iSpot and iRecord, fall perfectly within the technologically enabled domain of big data having been created around IT affordances. The third project, GiGL, in more general terms can fall into the conceptualisation of "Big Science", managing collecting and digitizing enormous amounts of data to inform policy and research. This analysis revealed how the most recent IT enabled data collections tend to focus on areas with more population density, can engage comparatively more people who collect on average fewer observations. Blue areas and wet areas are mainly amongst the most successful areas when surrounded by green spaces. This study also stressed the important overall role played by environmental organisations which have a specific spatial focus. This chapter also proven how very committed individuals can produce an incredible amount of data, skewing the analysis. This confirmed the correctness of our approach in counting the number of contributors rather than the number of observations.

[96] "Welsh Harp Conservation Group" and " Welsh Harp Conservation Gp", " Volunteer - Butterfly Conservation", " Recorder - BTO Migration Watch" and others more

[97] TQ537850, TQ412928, TQ519812, TQ361867. They all belong to the same survey Essex Field Club

8 Incremental projects

The case study involving incremental projects is the most challenging from a technological point of view. The need to analyse engagement in this case study required an innovative approach towards data since our use of those projects to evaluate connectedness to places is almost unprecedented. From a technological point of view, while it is not complex to download and store the lineage of Wikipedia articles with associated geographic information, it is computationally intensive to perform a similar work for OSM and requires the design of a bespoke web parser for Wikimapia. Moreover, OSM and Wikipedia are the main targets of the Web 2.0 and Geoweb 2.0 enthusiastic public. Therefore, there are several projects, applications, and initiatives that tend to improve them. This implies also the creation of automated systems to improve the quality of geospatial data or of spatial knowledge.

To detect digital cultural practices that are related to the local space we carried out a series of data preparation processes. We had therefore to remove all unsupervised edits performed without local supervision. This case study is set in ta subset of the area analysed in chapter 6. To have an introduction to the cultural aspects of this area please refer to section 6.1. The baseline situation for this case study is covered in the following sections. The first section 8.1, describes the data preparation processes. The second section 8.2, shows the basic global statistics as they have been obtained from the projects and how they have been changed after data preparation. Thirdly, section 8.3, is devoted to the mapping of the individual projects. In section 8.4. we turn to the description of the differentiation between urban and non-urban areas, here we also describe the study area in demographic terms. The analytical phase begins with section 8.5. in section 8.5.1 a grid by grid analysis is performed through statistical correlation analysis and scatterplots. The local spatial statistic part of this chapter (section 8.5.2) begins with the computation of local Moran's I and continues with the creation of maps with hotspots and outliers in the following section 8.6. To complete this analysis, section 8.7 describes the visual inspection on the characteristics of the places that have been revealed by the spatial statistical analysis. This chapter ends with a summary in section 8.8. This case study uses part of the area of the first case study area. While our initial intention was to reduce the size of the grid units zooming at a 100m grids we abandoned this option after few considerations. Firstly, the 100 meters grids work very well for an analysis aiming at OSM nodes only and in an urban setting mainly where features are relatively small sized, but it fails when we consider the non-urban areas characterised by larger entities. Secondly, Wikipedia and Wikimapia store information about spatial entities in centroids of aerial units. Therefore, in natural areas it is highly unlikely that aerial entities are such small as to be contained in a 100m grid. Thirdly, the information collected by contributors in the incremental projects that we are considering here is not solely targeted at natural elements. Therefore, we will use the OECD classification of 1km BNG grid units already used in the first case study.

8.1 Basic statistics: The preparation and harmonisation process

In Table 8-1, below, some basic statistics to compare the spatial and numeric extent of all data used in this chapter. Those numbers represent the original data including also automated edits and unsupervised imports.

	OSM	Wikimapia	Wikipedia	All projects combined
Total individual objects	1,175,844	1,651	1,525	1,179,020
Percentage of total objects	**99.73%**	0.14%	0.13%	100.00%
Objects and revisions	1,561,948	3,964	145,116	1,711,028
Percentage of total obj. and rev.	91.29%	0.23%	8.48%	100.00%
Total contributors	1,128	465	26,378	27,971
Percentage of total contributors	4.03%	1.66%	**94.30%**	100.00%
Objects and rev. per contributor	1,385	9	6	61

Table 8-1:Basic statistics of incremental projects in case study area. In bold noteworthy data

Table 8-1 shows how the three projects are different in terms of size, contribution patterns, and engagement of contributors. Before evaluating the impact of automated edits on those three data sources it is evident how OSM has overwhelmingly a larger number of objects. We must remember that we are considering every individual node of OSM while for the other projects we are dealing with centroids, a unique punctiform element for every information unit or article. They may refer also to very large entities such as administrative units. To make the numbers of entities between OSM and the other two projects comparable we derived the number of OSM's ways and nodes contributing to such ways as they are in the downloaded and unfiltered history (lineage) file of OSM. We have 71,914 ways built using 593,038 nodes. Therefore, we have 654,720 units of information in OSM (ways and POIs) compared to the total amount of information units from all projects. It still leaves OSM the most information-rich project but this way it contains just 55.53% of all information units and not over 99%. Wikimapia holds more information units than Wikipedia but like OSM is overpowered by the large numbers of contributors to Wikipedia.

This case study required an extensive research to prepare the data after collection to remove unsupervised, imported, or automated edits. We want to emphasize how in this case the intensity of human engagement for an area has been measured using the number of different contributors providing their work for all entities included in the same 1km BNG grid unit. This has been called *attractivity* by Arsanjani et al. (2015a). Therefore, we exclude automated contributions inserted by code. We wanted edits made by humans that are aware of the surroundings. Or at least that interact with the information units in a way that is contextualised. From now on, we will simply call "objects" the units of information used either there are nodes (in OSM) or centroids (in Wikipedia and Wikimapia). In Wikipedia (Wikipedia contributors 2017d) bots are used to import great amounts of data but also to check for grammar and formatting errors. The removal of their contribution is mandatory for the scope of this study.

OpenStreetMap is also affected by imports from external sources. Imports are excluded from this analysis since they generally cover quite large areas and do not come from a direct relation between the contributor and the area. Some imports are country specific and are listed in the following page (http://wiki.openstreetmap.org/wiki/Category:Import-uk) some of them are global imports therefore they are listed in the global list as follows: (http://wiki.openstreetmap.org/wiki/Import/Catalogue). The use of the term "import" in OSM is quite confusing since it includes several practices. With the term imports the community refers sometimes to fully automated processes, while at other times refer to the use of information from other sources that is picked and adapted to the map (like all OS Opendata products). For this work only entirely automated edits (also called in the community one-time imports) have been excluded from the analysis. Bots have been identified in all three incremental projects as follows: In Wikipedia, the flagging of edits created by bots followed two steps. We first compared the usernames of contributor in the data we downloaded and the list of Wikipedia's bots.

The list at the following address (https://en.wikipedia.org/wiki/Wikipedia:List_of_bots_by_number_of_edits) is the most exhaustive but it is not updated regularly. In it among 902 names 876 include the “bot” string. Secondly, we run a string matching query over the usernames of contributors in our area. This to find the usernames with the substring "bot". This to check if any other user that operated in our selected area is a bot whose username has not been inserted in the list previously cited. We had to check matching strings individually since the "bot" sequence may refer to real people such as someone using his real surname such as Abbott.

The use of bots in OSM is quite restricted therefore the search for bots followed a similar procedure as described above but using the list of OSM bots (http://wiki.openstreetmap.org/wiki/Bot). In OSM our target to remove bulk uploads from the analytical phase has also been previously done by Mashhadi et al. (2015).

Wikimapia has a policy hostile towards automation, they implemented a ban list (http://wikimapia.org/user/tools/ban/?action=list) which is also deemed to include everyone performing map vandalism or implementing automated edits. Automatic edits are therefore automatically removed from Wikimapia itself. But the project makes use of a bot to correct text. We located its contribution and removed it for our evaluations.

Another element that can include spurious element in our research is caused by some system malfunctions. Neis and Zipf (2012) reported that until a certain time, editing software in OSM created false positives. The software increased the version of a node even when the node was not actually changed. To amend to this false positive a specific code has been designed. This software was deemed to compare all couples of consecutive versions of a node. Comparing position (geometrically) and tags (semantically) it was able to flag when two versions were identical therefore the most recent version was deemed as coming from the malfunctioning editor and excluded the most recent version from our analysis. We found out that in our area this malfunction created even 30 identical version of the same node. The impact of the removal of automated edits on incremental projects is addressed in Table 8-2, we here extend Table 8-1 and compare the two data collections, before and after all the data preparation processes. The columns in grey contain the % of reductions due to the filtering of automated edits. Only OSM objects coming from unsupervised edits were entirely removed. Given the disproportionate contribution to the number of objects and revisions of OSM its reduction almost equates to the overall Wikipedia objects and revisions and is almost 27 times the contribute of Wikimapia in the area. The filtering impact is overall very low. In terms of internal percentages of projects, it does not reach the 7% of objects and revisions in both Wikipedia and OSM. Also in terms of contributors the filtered-out contributors are just around 1% of both Wikipedia and OSM. Figure 8-2 shows how the filtering does not change the relations between the data sources. In absolute terms, the filtering process involves numbers that while are irrelevant internally for the larger projects involving the removal of only 105,727 objects and revisions in OSM and 10,029 in Wikipedia, they are comparable with the overall Wikimapia project. The number of bots that contributed to Wikipedia for example amounts to 247 which equates to 53% of all contributors to Wikimapia in the area. This elevated number of contributors removed allowed Wikimapia to increase its share of contributors having lost just one of its own contributors. Even more significant is the removal of those bots equating to just 0.94% of Wikipedia's total number of contributors is increasing its total share of contributors since OSM lost 15 contributors that equates to 1,33% of its total number of contributors. Moreover, considering the different case study areas and time ranges of data collections all those removed objects, revisions and contributors are also relevant in comparison with the other projects considered in the previous chapters.

	Total individual objects	Total individual objects (%)	Objects and revisions	Objects and revisions (%)	Total contributors	Total contributors (%)
OSM	1,175,150	99.73	1,456,221	91.29	1,113	4.02
Impact (%)	-0.06	0	-6.77	0	-1.33	-0.02
Wikimapia	1,651	0.14	3,865	0.24	464	1.67
Impact (%)	0	0	-2.5	0.01	-0.22	0.01
Wikipedia	1,525	0.13	135,087	8.47	26,131	94.309
Impact (%)	0	0	-6.91	-0.01	-0.936	0.004
All projects	1,178,326		1,595,173	100.00	27,708	
Impact (%)	-0.06		-6.77	0	-0.94	

Table 8-2:The impact of automated edits on incremental projects

Another important factor to consider is the locality of impact. While an overall impact of 7% looks negligible, if this is not uniformly spread on all over the study area it can have locally relevant impact.

8.2 Basic global statistics of incremental projects

From this point onward, unless otherwise specified, the statistics and maps shown refer to the data aggregated to 1km^2 of the British National Grid. Considering all projects as they unfold in the selected study area, some basic statistics to compare them have been extracted in Table 8-3. As we have seen in the previous section the filtering process is not influencing the relation between projects so it will be no more mentioned, but we still will use filtered data since automated edits may induce local patterns that are not of our interest. For instance, in our study it is not relevant if in coastal areas data about bathymetry was imported or similarly if an editor with local interest but with bad writing skills was automatically corrected by bots dedicated to detecting and correcting spelling and grammar mistakes. To correctly interpret Table 8-3 we must stress the fact that we are analysing elements that are quite heterogeneous also in spatial terms. While OSM is analysed at an individual point level, Wikipedia and Wikimapia are geographically represented by centroids. This difference is given to the fact that in OSM information is associated with points and the literature on OSM focused on them. We had access to all points describing Wikimapia entities (40,712 points) but those points have only visual value since edits are not associated to every one of them and Wikimapia does not record changes in their individual position.

	OSM	Wikimapia	Wikipedia	All projects combined
Total individual objects	1,175,150	**1,651**	1,525	1,178,326
Percentage over total objects	99.73%	0.14%	0.13%	100.00%
Objects and revisions	**1,456,221**	3,865	135,087	1,595,173
Percentage over tot obj. and rev.	**91.29%**	0.24%	8.47%	
Grid squares with objects	2,712	709[98]	846	2,713
Coverage over total grids	**99.96%**	26.13%	31.18%	
Total contributors	1,113	464	26,131	27,708
Percentage over total contributors	4.02%	1.67%	**94.31%**	
Max. objects per contributor	178,983	657	556[99]	

Table 8-3:Basic spatial statistics of incremental projects in case study area after data preparation

Considering the above and the fact that OSM is the most explicit project in the willingness to create digital geographic information, and therefore the number of geographic nodes and modifications associated with OSM is more than 10 times the number of Wikipedia articles and 50 times the number of Wikimapia places. OSM has 30 times the number of points that Wikimapia uses to describe spatial entities. Similarly, can be interpreted the difference in spatial cover where OSM nodes are found in almost all grid units while the

[98] If we consider not only the centroids but also all points used to draw an entity we have 1,807 grid squares that are 40.04% of the total.
[99] This is the filtered value of the top human contributor. The top two overall Wikipedia contributors were bots with values that reach 900 and 600 contributes.

other two projects have their centroids found only in a third of grid units. This must be considered in the specific context of Wikimapia data production. It is always visualised and edited with a background image, so it usually collects *platial* information that is not represented in the aerial imagery used as background. While OSM is intended as a dataset that represents all elements in space and is not intended as complementarily to aerial imagery. Moreover, Wikimapia data about infrastructure networks is not available to download while OSM was primarily devoted to such kind of data that is not only available to download but makes a considerable portion of OSM data.

In terms of contributors Wikipedia affirms its strong superiority. The most popular UGC website in the world which is also partially a CGI project in the area has 26 times and 60 times the contributors of OSM and Wikipedia respectively. This is highly significant and must be considered thoroughly interpreting the results. The huge differences shown in Table 8-3 stresses how an approach based on ratios rather than total numbers is appropriate to compare the projects involved. It also stresses how the handling of those three projects must be tailored and how they give quite different results if considered independently.

8.3 Spatial patterns of individual projects

To start the analysis, we first looked at spatial patterns of the individual projects. The different projects have quite different number of contributors, therefore maps depicting the total numbers of contributors for each project while interesting can be less useful in the development of this work that aimed at the comparison of projects. Therefore, to map contributions we derived the grid by grid contributor ratio (Cr). This is the percentage of contributors for each grid unit compared with the total number of contributors to the specific project in the study area. As an example, in Wikipedia there are 26,131 contributors in the study area while in the BNG grid ST1575 in Cardiff there are 5 contributors. The Wikipedia contributor ratio for grid ST1575 is given by Cr=5*100/26,131. Therefore, the Wikipedia contributor ratio Cr for the BNG grid unit ST1575 is 0.02. Wikimapia is the project with less contributors overall. Its centroids cover the least surface with just 26% of the grid units and only 1.67% of contributors. Another weak element of Wikimapia is the late introduction of user registration and the adoption of usernames. Therefore, there is not the possibility to distinguish between early contributors. This drove to the fact that the most prolific contributor in the study area is "Guest" with 657 entries (17% of all contributions). Moreover, before the full implementation of user registrations, contributors have been recorded with the nickname Guest associated with an IP address. While this is a step further to distinguish between different contributors it may still lead to incorrect associations[100]. In total, we have 195 Guest or Guest+IP address contributors (42% of total contributors) in Wikimapia in the case study area. They provided 964 edits (25% of the total).

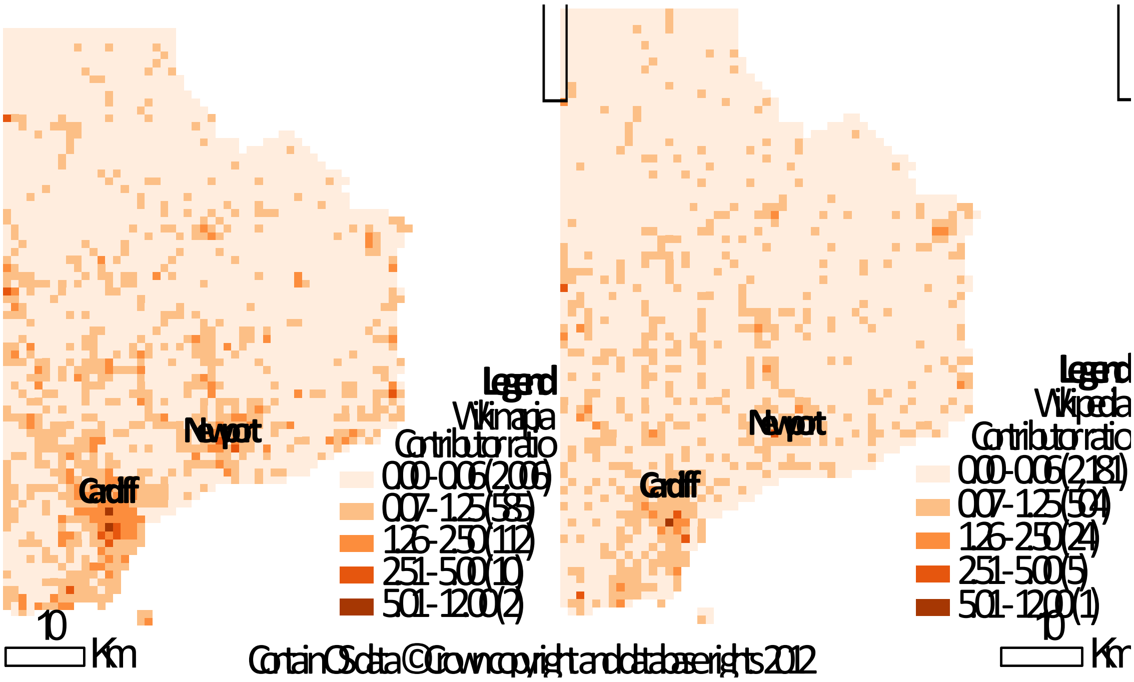

Figure 8-1:Wikimapia (left) and Wikipedia (right) Contributor ratio

In Figure 8-1 the maps depict the Contributor ratio for Wikimapia and Wikipedia. They have been coupled in Figure 8-1 since they both are bearers of geographic information associated to centroids. Considering that from Table 8-3 they also have comparable covering of all the territory they might be quite similar in terms of spatial patterns. In fact, the five categories chosen have been populated in the same order and magnitude.

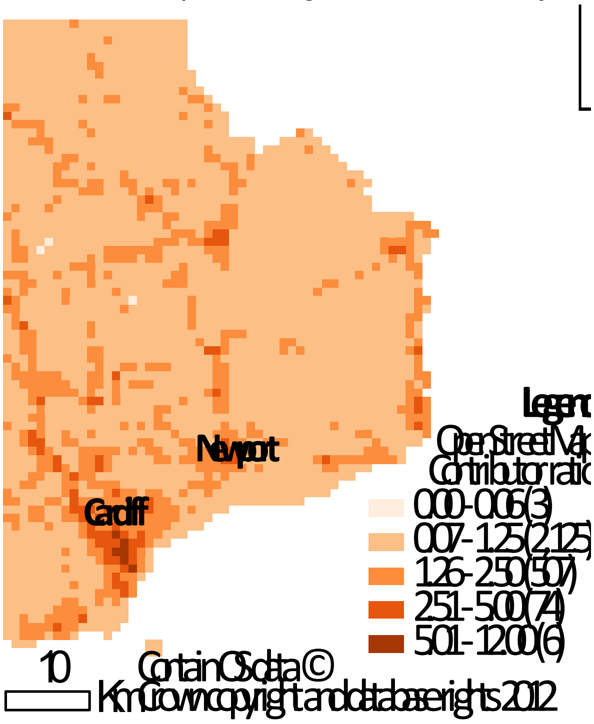

Figure 8-2:OpenStreetMap Contributor ratio

In Figure 8-2 we show the Contributor ratio of OSM. It is evident how the explicitly spatial project and the analysis conducted over the individual point elements provides a higher level of basic engagement while highlighting similar but larger spatial patterns. The distribution of contributor's ratio is

[100] The same household can have a dynamic IP address or the same user may access from different places with different IP addresses associated to them.

pushed one level up compared to the other two projects with the most populated category being the second where the Cr of OSM is in the range between 0.07 and 1.25.

8.4 Residential population, Urban and non-urban areas

The study area as briefly described in chapter 4 is included in 2,715 1km BNG grid units where approximately 1,137,961 inhabitants live with a total average population of around 420 inhabitants/km^2. The OECD classification used identified 486 urban grid units (~ 17.9%) and 2,229 non-urban. Our analysis will focus on the non-urban areas since the selected incremental projects aim at the collection of human knowledge in general terms and do not focus on nature as in the case study in chapter 7. Therefore, in urban areas contributors will more likely focus on elements of no interest for this study.

Figure 8-3 shows the subsection of south Wales chosen for this analysis and its grid units. The behaviour of the three projects is evidently different when urban and non-urban areas are considered as discerning categories. Examining the relation between urban and non-urban grids, we can see that all projects tend to have more contributors in urban areas.

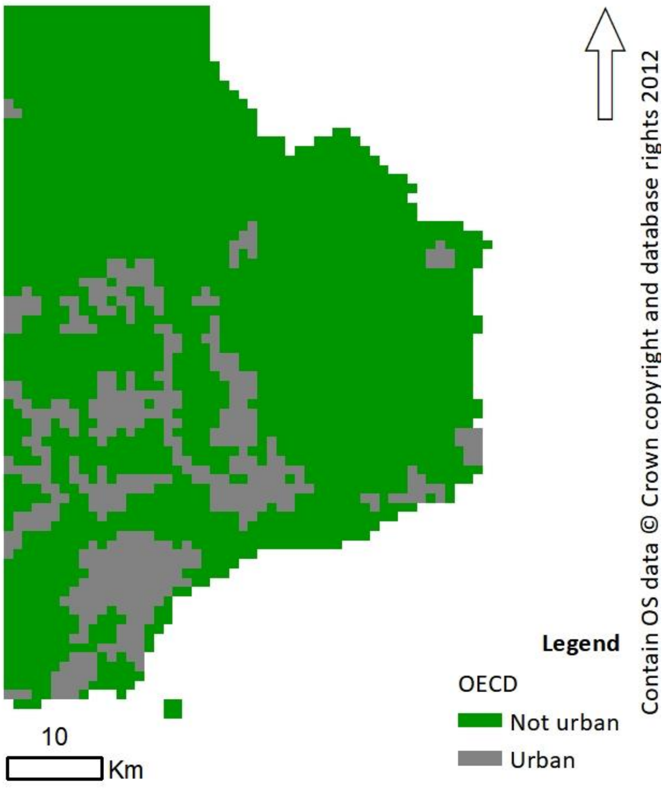

Figure 8-3:Urban and non-urban areas according to the OECD methodology

Examining the statistics as shown in Table 8-4 Wikipedia is leading as expected in terms of average contributors, but it is exceedingly focused on urban areas. It loses the leadership in terms of average contributors in non-urban areas to OSM. In urban areas it has 4 times the contributors of OSM while in general it has almost 20 times as much as OSM contributors. OSM looks more balanced with urban average contributors only 2.4 times the non-urban contributors while Wikipedia and Wikimapia have a relation that goes from 11 times for Wikipedia to 7.5 times for Wikimapia. Wikimapia is lagging behind OSM and Wikipedia in all statistics.

Projects	Non-urban contributors		Urban contributors		All study area contributors	
	Max	Avg.	Max	Avg.	Max	Avg.
Wikipedia	1,018	7.02	2,244	**78.85**	2,244	**19.88**
Wikimapia	11	0.42	34	3.15	34	0.91
OpenStreetMap	42	**8.09**	132	19.17	132	10.07
All Projects	1,046	15.53	2,370	101.16	2,370	30.86

Table 8-4:Contributors productivity in urban grid units, non-urban and all grid units in the study area

8.5 Spatial analysis

We now turn to two levels of analysis. At first, the analysis of the relation between projects on a grid by grid basis ignored the topological relations between grids. In the subsequent level of analysis, we took into consideration also the topological component. Therefore, there was the analysis of the spatial patterns, hotspots and outliers using all combined platforms. The correlation between projects, considering every grid unit separately, is investigated through the development of scatterplot diagrams and rank Spearman's statistics. The scatterplot diagrams have been derived for the overall study area and for non-urban grids while the rank Spearman's test has been derived for the overall study area and for the non-urban grids in section 8.5.1. The global analysis for local clusters and outliers in section 8.5.2 uses the local Moran's I method.

8.5.1Relationships between projects

We used two methods to compare the three projects in a grid by grid basis using the number of contributors to each project as a comparison metric. The design of scatterplot diagrams and the estimation of the Spearman's rank test. The latter is run through couples of projects and gives a measure of the correlation between the two ranked collections ρ and it measures the statistical significance of it (p). The scatterplot method provides a visual impression of the correlation grid by grid. The visual analysis of scatterplots, while it might lead to misleading interpretation due to overlapping points is very helpful to determine outliers. To design the scatterplot, the contributor ratio (Cr) was used. For the calculation of Spearman's rank test, the use of this ratio is not needed since this statistical method ranks the values for every project and compares the ranks but does not consider the relative magnitudes.

All study area	Wikipedia	Wikimapia
Wikimapia	ρ= 0.5229964; p<0.0001	
OpenStreetMap	ρ= 0.4794382; p<0.0001	ρ= 0.5004424; p<0.0001

Table 8-5: Spearman rank test for all grids of incremental projects case study

We see from Table 8-5 that in the study area all projects have a mild positive correlation, with the best correlation between Wikipedia and Wikimapia.

Non-urban	Wikipedia	Wikimapia
Wikimapia	ρ= 0.3465536; p<0.0001	
OpenStreetMap	ρ= 0.313285; p<0.0001	ρ= 0.2911318; p<0.0001

Table 8-6: Spearman rank test for non-urban grids of VGI and Wikipedia case study

When we consider only non-urban areas we have a positive but mild correlation. The strongest correlation is between Wikipedia and Wikimapia as for the overall study area. The low values can also be a consequence of the fact that Wikipedia and Wikimapia are zero inflated mostly in non-urban areas, so the rank test will hardly be able to rank all zeros in a similar way.

Urban grids	Wikipedia	Wikimapia
Wikimapia	ρ= 0. 5054318; p<0.0001	
OpenStreetMap	ρ= 0.445232; p<0.0001	ρ= 0.5847914; p<0.0001

Table 8-7: Spearman rank test for urban grids of VGI and Wikipedia case study

We have seen how urban areas are the ones with larger contributions. In Table 8-7 we see how the ranks of their Cr values are correlated. We have strong to mild positive correlations with the strongest correlation being between Wikimapia and OSM, the two collaborative mapping projects. The other two comparisons show a quite valid correlation, but Wikipedia has overall a better coupling when all the study area is considered. The scatterplot diagrams, given the correlations found through the

Spearman's rank test, are very useful once we focus on the detection and characterisation of outliers.

The relation between Wikipedia's and Wikimapia's contributor ratio in the area got the best Spearman's rank test for this case study. The scatterplot that represents their relation is shown in Figure 8-4. Almost all values are inside a 2% box with Wikimapia values in general smaller. At the upper right corner, a significantly correlated and high popular grid is the ST1776 it is classified as urban. This grid square is located at the centre of Cardiff and contains the centroid of the millennium stadium and the spatial unit referring to the Castle area in Cardiff that are on the top eight places in Wikimapia for the area. In Wikipedia this is the grid unit with more revisions (5,490) and contributors (847) in the study area, the Cardiff Castle, and the Cardiff Arms Park (A rugby stadium) that collected more than a hundred of revisions and contributors.

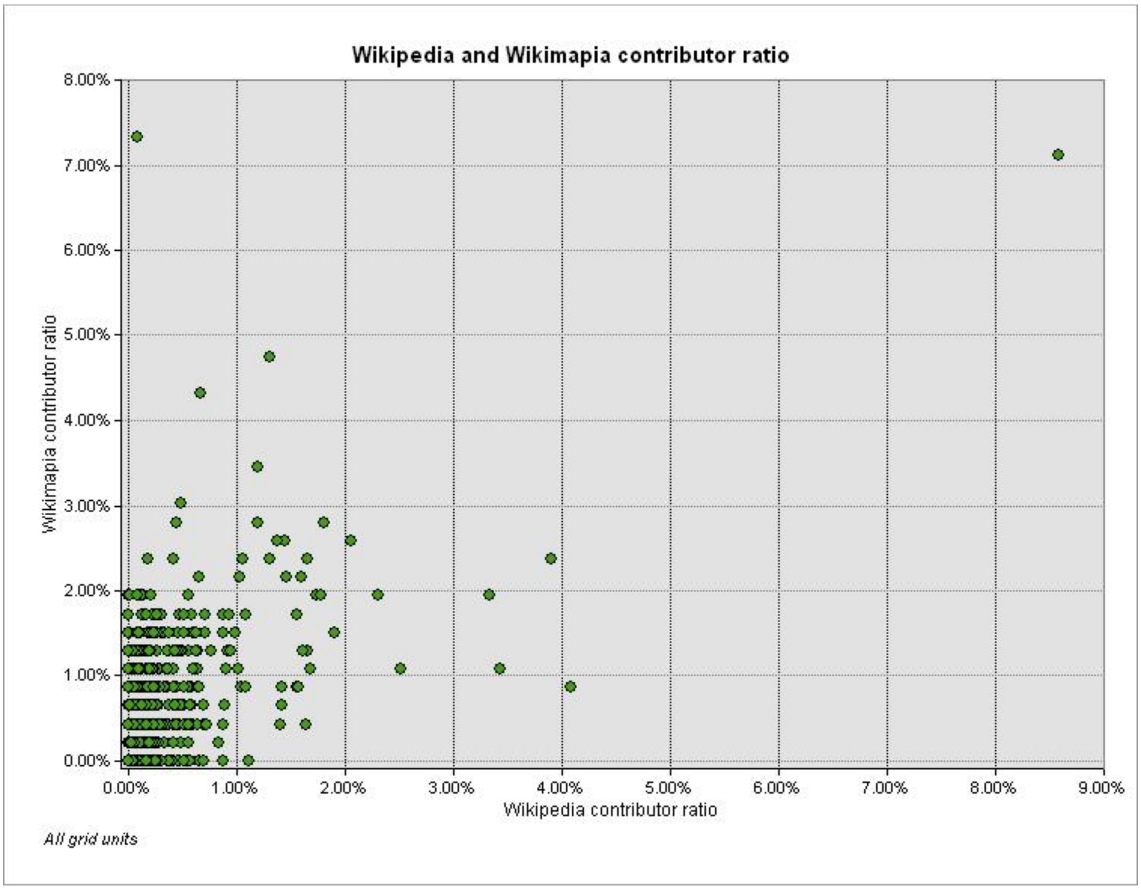

Figure 8-4: Wikipedia and Wikimapia contributor ratio scatterplot for all case study area

Three grid units are disproportionally more popular amongst Wikimapia contributors and they also in urban areas in Cardiff and Newport. This applies also to the grid units disproportionally popular in Wikipedia compared to Wikimapia. The relation between OpenStreetMap's and Wikimapia's contributor ratio for all grid units is in Figure 8-5. The wider spatial distribution of OSM and the relatively scarce popularity of Wikimapia are mutually compensating each other. This is the second strongest correlation amongst incremental projects considering all grid units and the fourth overall correlation amongst incremental projects. Central Cardiff and other urban areas in the surroundings are represented by all points that are out from the main cloud.

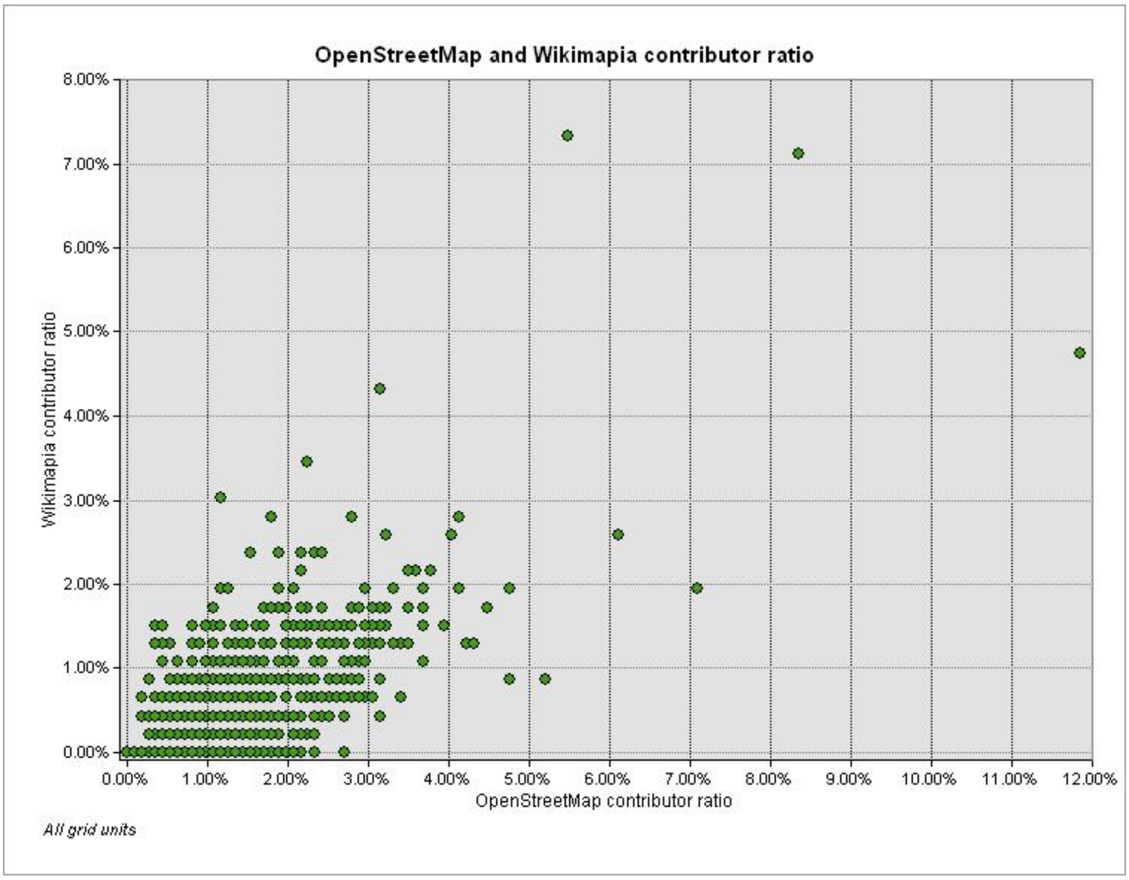

Figure 8-5:OpenStreetMap and Wikimapia contributor ratio scatterplot for all case study area

The scatterplot between OpenStreetMap and Wikipedia for all the study area is in Figure 8-6. We have the most spatial and the most textual project that show a mild correlation. Here we see how the plume is somehow leaning towards OpenStreetMap due to its spatial explicit nature and consequently distribution of objects. Almost all the points far from the main aggregation are representing urban grids. When focusing on points that show a disproportionate success for Wikipedia we find some non-urban areas with the grid covering Tintern Abbey (SO5300) and surrounding areas amongst them.

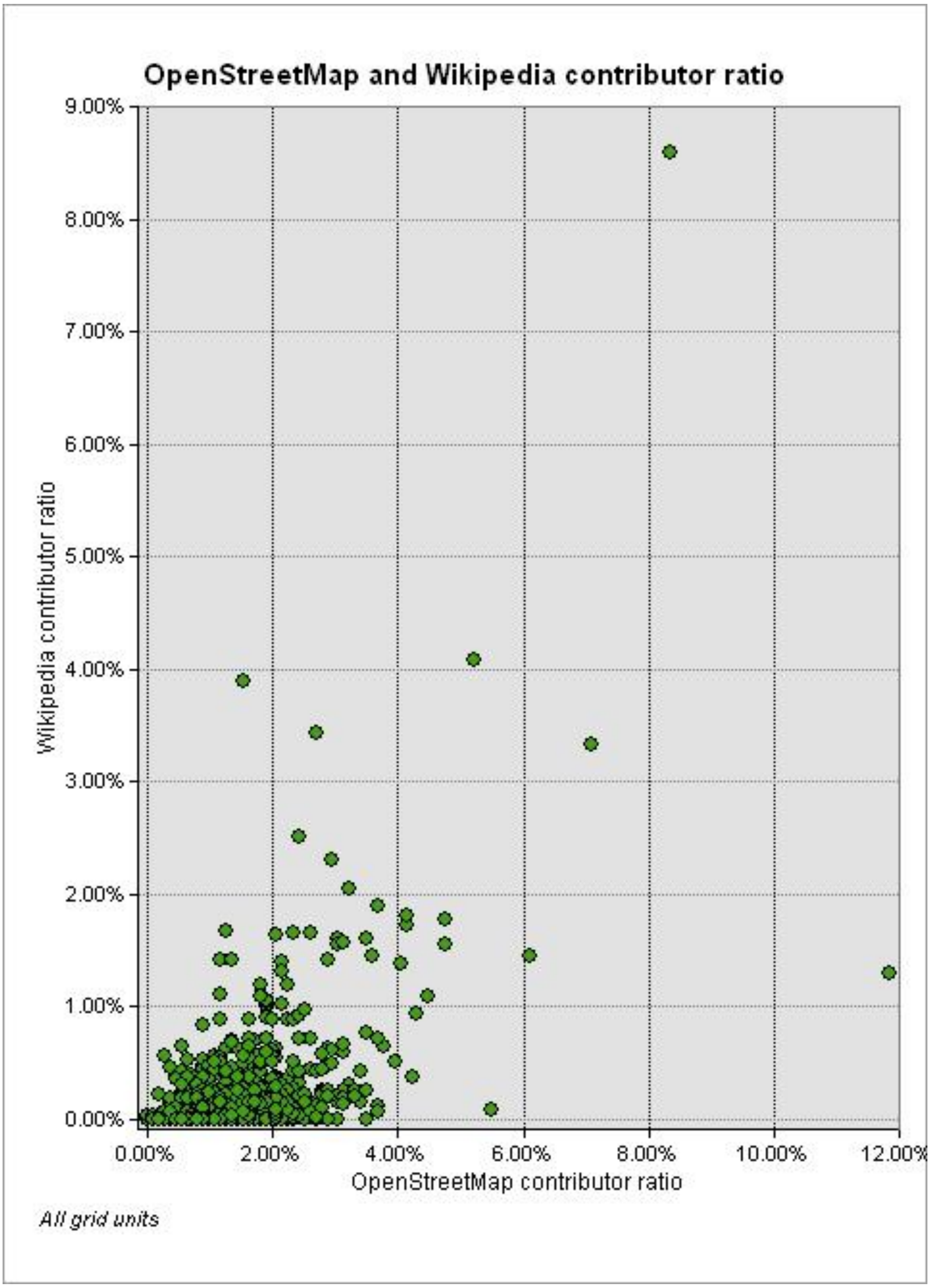

Figure 8-6:OpenStreetMap and Wikipedia contributor ratio scatterplot for all case study area

When we focus only on non-urban areas as in Figure 8-7 the point representing Tintern Abbey and surrounding areas stands out as the leading outlier of OSM. While the leading outlier for Wikipedia is the grid ST0667 containing Cardiff airport and a RAF installation, located in Wikipedia in the exact same point. The airport story raised some controversies in the recent years, this attracted several contributions. The points with exceedingly OSM contributor ratio include areas in the fringe of cities.

Two quite correlated grid units, inside the 1% -2% box, are SO1709 and SO0600. In the first are geolocated three articles regarding Blaenau Gwent as a constituency in both British Parliament and Welsh Assembly. Therefore, those points are in this grid but they refer more widely to the administrative unit and populated areas mostly included in the neighbouring urban grids. They include the three towns that are included in Blaenau Gwent. The two most edited articles in this grid unit are located in open fields with no specific physical elements to refer to but they are loaded with a large amount of electoral results. The absence of any further local specific interest is also witnessed by the absence of any Wikimapia object in the specific grid unit and by the fact that OSM nodes have been changed only once after creation with version 2 as the maximum in the grid unit. The SO0600 grid unit is particularly significant. Here is located a complex cultural element that is related to a problematic relation with the environment, an ecosystem service (coal) whose intensive and not considerate exploitation not only damaged the environment but also induced a tragic disaster.

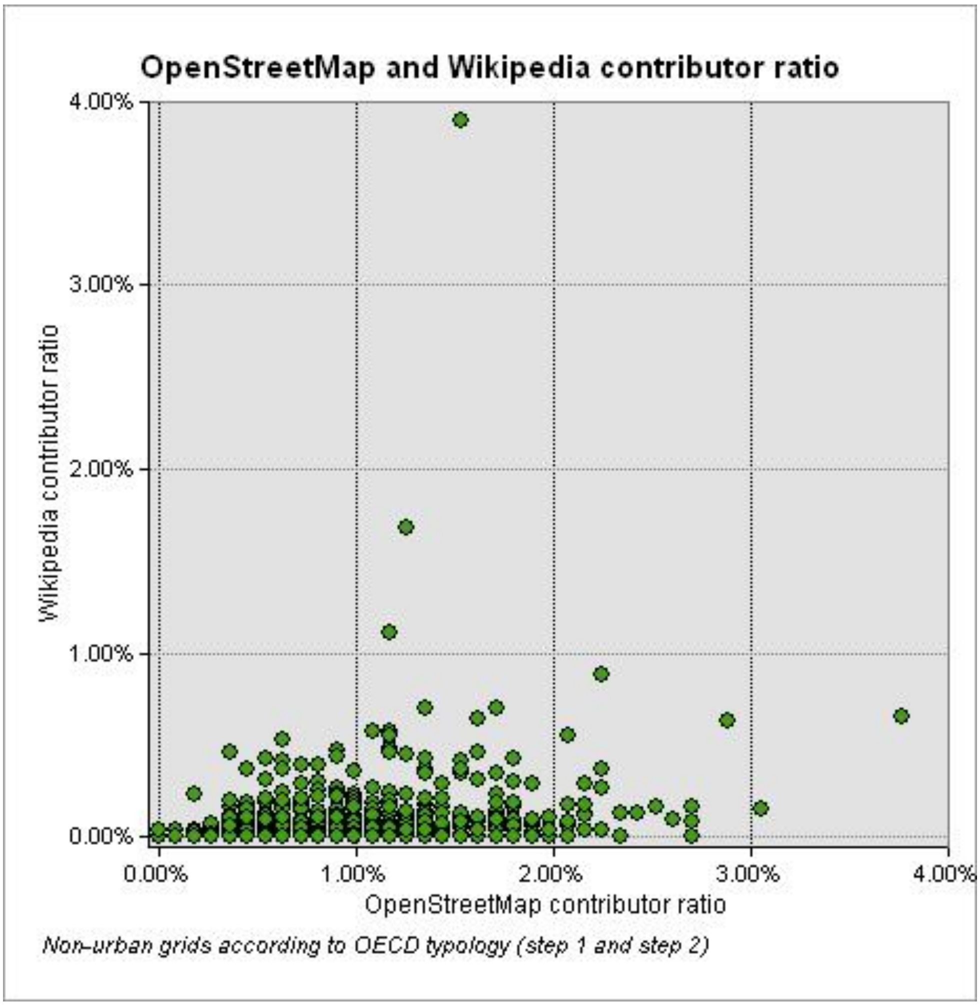

Figure 8-7:OpenStreetMap and Wikipedia contributor ratio scatterplot for Non-urban grids in case study area

The Aberfan disaster article in Wikipedia attracted 225 contributors at the time of data collection[101]. While the elevated number of editors in OSM is mainly due to several corrections on a close highway. The worst tragedy in modern Welsh history happened on 21st October 1966 and claimed the life of 116 children and 28 adults. It was due to the collapse of a colliery spoil tip over the village and it hit directly a junior school killing almost entirely a generation of residents. While the Wikipedia article is very rich and detailed, the physical element reminder of the disaster, the Aberfan disaster Memorial Garden of remembrance is just one POI that had just one contributor in OSM. The gardens have the potential to be designed in more details for mapping and cartographic purposes, being around 80 meters long and 30 meters wide. This lack of representation is maybe derived by the lack of a local person contributing to OSM.

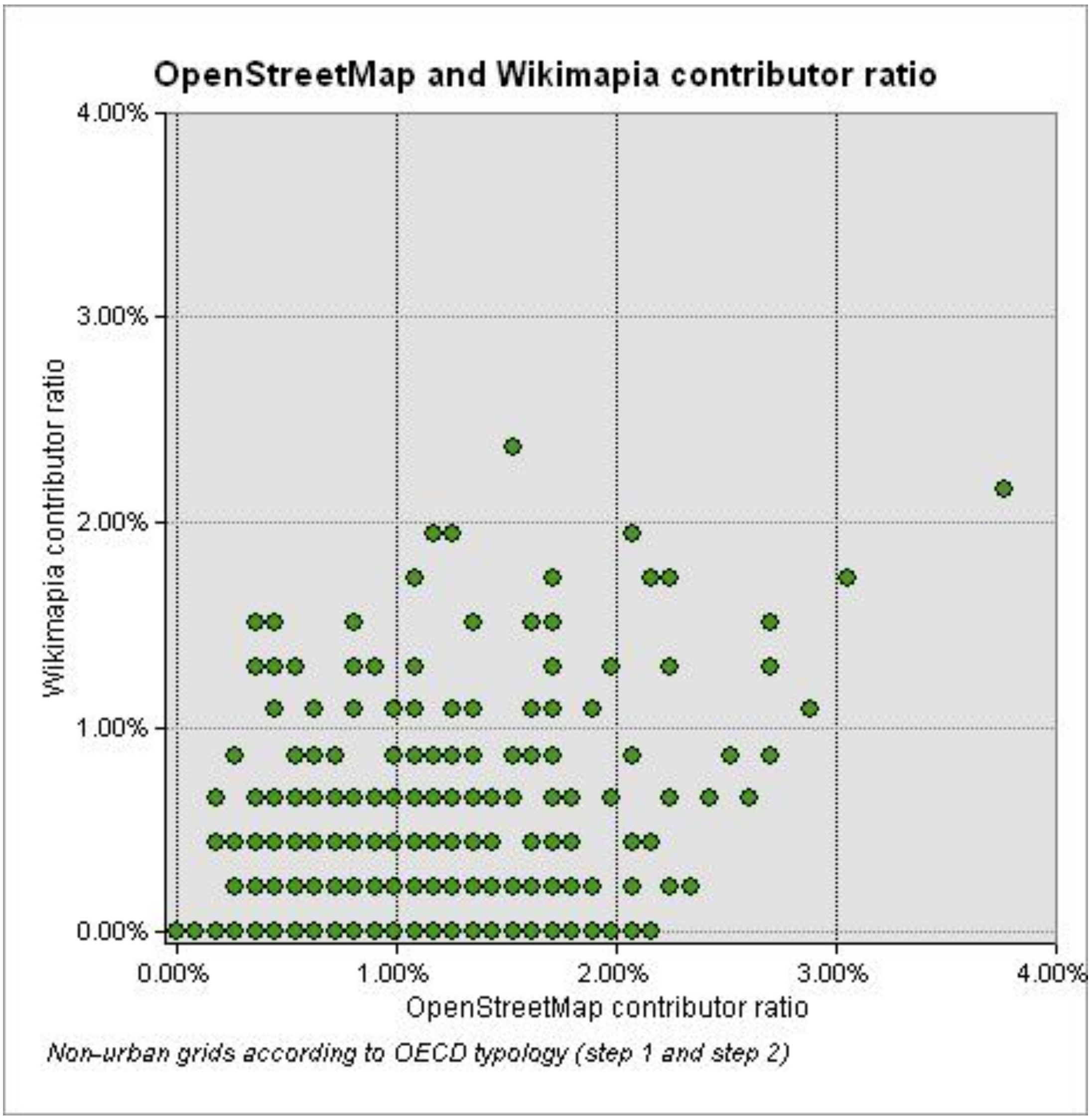

Figure 8-8:OpenStreetMap and Wikimapia contributor ratio scatterplot for Non-urban grids in case study area

This is also witnessed by one tool which allows us to see that there are not really engaged contributors in the area and the ones operating have a quite limited activity

[101] New edits to this Wikipedia article are performed on an almost daily basis

range not centred in Aberfan[102]. The relation between the Cr of OSM and Wikimapia for non-urban BNG grids is in Figure 8-8. This is the overall lowest correlated pair according to the Spearman's rank test but with a ρ= 0.2911318 there is still a mild correlation. The widespread OSM spatial cover makes the cloud leaning towards the OSM axis. The most popular OSM non-urban grid unit which is the second most popular in Wikimapia is SO5300 which we already mentioned. It covers Tintern Abbey and the surrounding area. Other relevant correlations are found in ST1283 and in the above mentioned ST0667 which covers Cardiff's airport. Similarly, ST1283 covers a highly urbanized but spatially limited area with around 2000 inhabitants. All items here intensely mapped are not nature-related. This reinforces our approach towards the preventive exclusion of urban areas but suggests that methodological improvements are still needed. The grid ST0799 focusing on the central part of Aberfan but without including the site of the disaster here emerges in a small group of three grids close to the 2% Cr for both OSM and Wikimapia non-urban grids. The other two grids include the village of Usk (SO3700) and a mostly open area with some shopping centre close to Cardiff urban area (ST1174). When grids show a better success for Wikimapia we have more heritage and natural areas.

Inside the 1-2% correlation box we have the grid ST2264 that includes most of Flat Holm, a small island in the Bristol Channel. This is a place of more natural and cultural interest with the lighthouse as main contributed element in Wikimapia while in OSM the contribution is spread evenly amongst several contributors but most of the edits are focused on the coastline and few physical infrastructures. In ST0866 we have a predominantly natural area close to Barry and the Cardiff international airport. In ST1395 we have features of a park like a landscape carved horse (Sultan the Pit Pony) and an observatory to enjoy the surrounding environment. While those elements are drawn in details in OSM they did not trigger a wide collaboration.

The left-hand side box where Wikimapia is more popular than OSM we have several grids in the area east of Newport characterised by some infrastructure, a large brewery, and a culturally interesting place with the site of Pencoed Castle, a ruined Tudor mansion, also featured in Wikipedia (ST4089). The grid unit ST3691 that covers a golf course who is accurately described in Wikimapia and OSM hosted the Ryder cup in 2010.

---

[102] http://resultmaps.neis-one.org/oooc?zoom=14&lat=51.69527&lon=-3.33752&layers=B0TFFFFFT

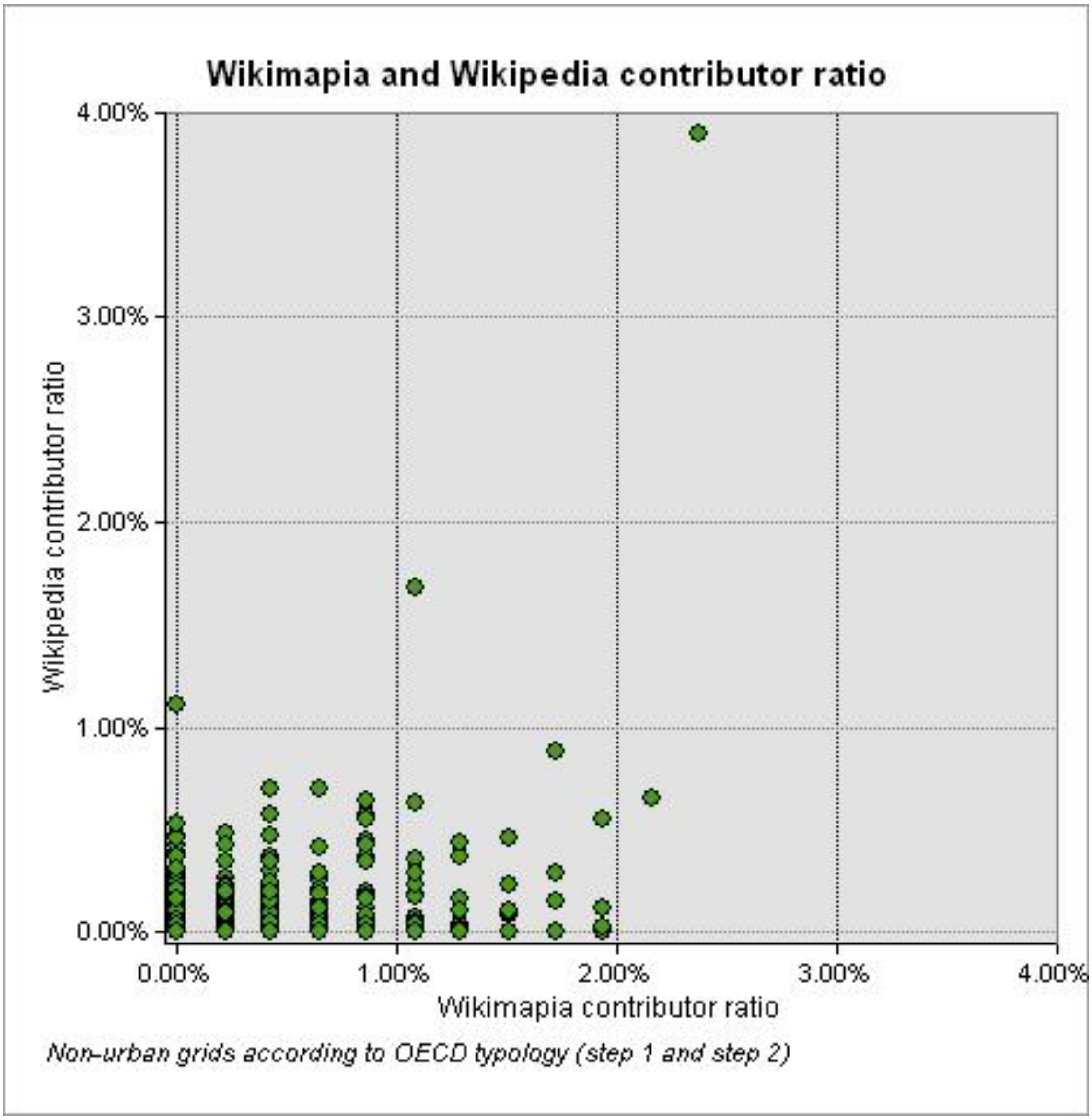

Figure 8-9:Wikimapia and Wikipedia contributor ratio scatterplot for non-urban grids in case study area

The Spearman Rank test found that the best correlation in non-urban grids is between Wikimapia and Wikipedia. Those are also the most *platial* resources we have. In our analysis, they are also referred to space using centroids where all the edits to every individual article are concentrated. In Figure 8-9 we have the relation between their Cr grid by grid for non-urban grids. The highest correlated grid unit is the one covering Cardiff airport (ST0667). Grid SO0600, which covers the Aberfan disaster area, stands out mainly for the contributions to Wikipedia. The third most contributed area in Wikipedia the mentioned Bleanau Gwent grid unit, has no contributors to Wikimapia (SO1709). A group of three grids a bit out of the main cloud and moved towards Wikimapia are SO3700, ST1174, and SO5300. The first covers Usk, a village with history dating back to pre-Roman Britain. Its popularity in Wikipedia is due to its long-lasting history as human settlement. The ST1174 grid includes an administrative unit centroid (Culverhouse Cross) which is the most edited article and a long-lasting transmitting station whose Wikipedia article includes a lot of technical details. The last grid SO5300 includes the mentioned Tintern Abbey. Therefore, we can conclude that Wikimapia, more than Wikipedia, is locally grounded and contextualised.

8.5.2Spatial statistics

This section is devoted to the computation of the global and local statistics that accounts for spatial autocorrelation. The calculation of spatial local autocorrelation helps the identification of grid units that are statistically and spatially relevant. The identification of grid units that individually or in spatial aggregations have values that are in stark contrast with the values of surrounding grids have been identified using the values given by the total amount of contributor ratio to all incremental projects. By analysing the spatial distribution of the ratio of contributors, we observed the focusing of people's interests from urban to non-urban areas when they create or contribute to improve spatial-related information either in form of digital geographic information for mapping and cartography or in terms of spatial knowledge. We then considered the number of contributors to Wikipedia, Wikimapia and OSM, focusing on the distinction between urban and non-urban grid units. We began analysing all the area. We then calculated the Moran's I global spatial autocorrelation index, calculated over the distribution of total Cr for every grid unit using a queen contiguity[103] ($I = 0.267256$, $\rho < 0.001$), reveals that there is < 1% probability that the autocorrelation has been generated by a random distribution. In Figure 8-10 we have the distribution of contributors to incremental sources for the area in south east Wales. There are clusters of grids with elevated number of contributors in both urban and non-urban areas.

[103] Are considered contiguous two polygons when they share an edge and/or a vertex

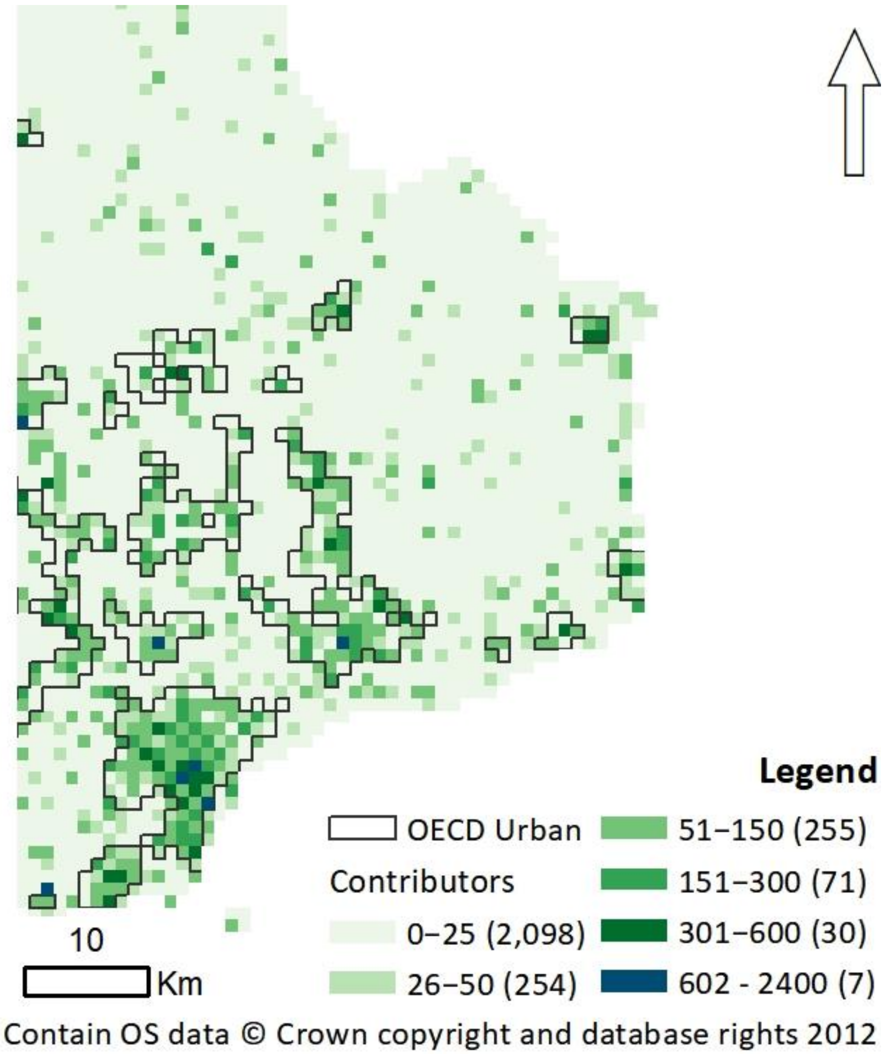

Figure 8-10:Total contributors to incremental projects and classification of BNG grids according to OECD procedure

In general, as can be seen in Figure 8-10, there are several areas where the aggregation of contributors to the projects here analysed is particularly evident. Most clusters of contributors for all three platforms cover urban areas, some other clusters are in the immediate proximities of urban areas, while a smaller number located in faraway non-urban areas, with the vast majority of them in protected areas.

8.6 Spatial trends: clusters and outliers

This study derived clusters and outliers, which will comment on in the following section. Here, we are more interested in the non-urban areas. To detect spatially statistically significant grid units, such as spatial clusters and outliers, this study used Moran's I local indicator of spatial association (LISA) (Anselin 1995) as implemented in ArcGIS 10.2.2 spatial statistic tools. To characterise the connectedness of incremental projects to space we used the sum of the Contributor ratio Cr of the three projects. After several tests, the most significant setting to elicit spatial clusters have been set to a 10km max radius inverse distance. This brought to the finding of both clusters of grid units with high numbers of Cr and grid units that differed markedly from surrounding grid units. The resulting spatial distribution can be found in Figure 8-11, below. There were 988 grid units with significant local autocorrelation. Amongst them 260 are in urban grids and 728 in non-urban grid units. Amongst non-urban grid units 101 are grid units with higher values compared to the surrounding areas while 627 are grid units with lower values compared to the surrounding grids. The cluster and hotspot analysis reflects closely the settlements typologies. The south and west-central areas which are defined by more urban settlements are characterised by large clusters of higher values surrounded by areas of relatively low values.

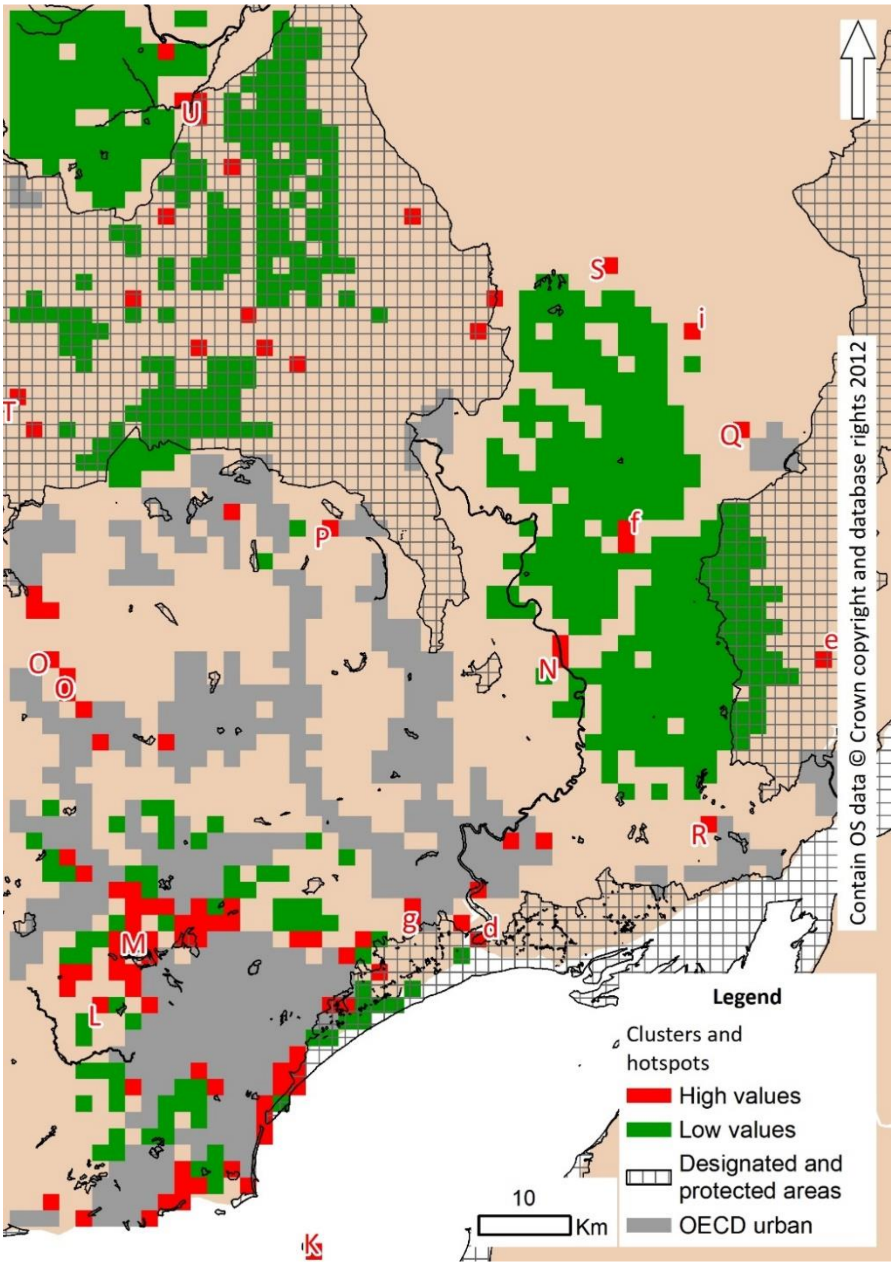

Figure 8-11:Clusters and hotspots, as identified by Moran's I local indicator of spatial association (LISA). Capital letters indicate clusters of grid units discussed in the text, lowercase letters are described in section 6.5.4

The northernmost part, central and easternmost part, largely defined by villages, hamlets and isolated dwellings are statistically characterised by large clusters of low values with scattered individual grids or small clusters of higher values. Focusing only on areas that have not been classified as urban according to OECD we have removed the vast majority of hi values concentrated largely in the south and along The Valleys. Still the vast majority of small clusters of higher values are left in the northern section of the study area. In the following we are going to characterise some clusters pertaining to different classifications and in different settings.

8.7 Visual inspection: Sites of intensive collaboration

Building on the experience of chapter 6 we are firstly comparing the result of the current cluster analysis with the results achieved in chapter 6. Afterwards we are going to characterise some specific outcomes of this chapter.

Comparing Figure 8-11 with Figure 6-12, we see how there are many hi-value clusters that overlap. Almost all the clusters we already characterised in chapter 6 for the area (D, E, F, G, I, J), are all in smaller or larger extent also included in the clusters found in this chapter. To distinguish them from the clusters and hotspot found out in this analysis we added them in Figure 8-11 using lowercase characters. For smaller clusters such as the "new" D and F we have the influence of the geographic reference to centroids used in Wikipedia and Wikimapia versus the fact that pictures are often geolocated in the surrounding area compared to the place of interest. The D identifies the Newport Wetlands SSSI and NNR. The E identifies the Tintern Abbey. The F identifies the Raglan Castle. The G identifies Tredegar House. The I identifies Skenfrith Castle. The only missing cluster is the J whose included several elements like some Neolithic dolmens or burial chambers and National Trust houses. This is maybe caused by the very basic OSM mapping of the area than ignores the gardens of Duffry house so popular amongst photographers and devotes just a POI to the Neolithic elements in the landscape. We also have clusters or individual grids that were identified as statistically significant also in chapter 6 but that we have not characterised and described in detail. We are going to characterise them now assigning them a letter to identify their location in the map in Figure 8-11.

In the southernmost section of the current study area we have an isolated hi value found in the Flat Holm island (Figure 8-11, K). Around Cardiff we have two grids, ST0979 (L) and ST1185. The former includes a service area, West Cardiff, that is drawn in great details in OSM as well as the neighbouring paths and agricultural fields patches of land. The large parking space is popular amongst Flickr users for the meetings and car rallies (SAAB spring rolls in 2014 and 2015). ST1185 (M) in the photosharing case study is a solitary hotspot while in this case study it is included in the large cluster that stretches from central Cardiff and goes northwards along The Valleys. In OSM (M) includes a quite detailed mapping of an industrial estate and of an adjacent river and natural area. In the case of photosharing the highest Cr is from Geograph and is focused on the industrial estate. In the central part of the study area we have SO3700 and SO3701 a two grids hotspot (N) both in the photosharing case study and in this case study. Those two grids include the totality of the village of Usk, already mentioned in the previous section. As said the village also has some interesting monuments that have been subjected to photographic interest. The Castle, the bridge, and the village of Llanbadoc especially contribute to the success of this cluster.

To show the clusters of hi value that are peculiar of this case study we are going firstly to characterise some clusters of hi isolated values in different areas. The ones more interesting are not contiguous to large urban areas. We find here a small cluster of three grids ST0799, ST0798, and SO0600 (O). The latter includes Aberfan, the cluster stretches southwards following the Taff valley. This cluster follows the low densely inhabited valley and the grids are greatly enhanced by OSM infrastructural focus. The grid unit SO2308 (P) is characterised by scenarios deriving from the Welsh heritage of coal mining. In this grid, we have the Big Pit National Coal Museum (https://museum.wales/bigpit/) and, according to Wikimapia sourced information, in this grid we have the set of some shots of a TV series called Coal House whose main setting is in the neighbouring Blaenavon. In Rockfield, Monmouthshire (SO4814), the nice landscape is not the main reason of the high values but the location of Rockfield studio. Considered amongst the best recording studios in the world (Q). The grid ST4690 (R), in the south east of the study area, emerges for the cultural features. It is characterised by the ruins of the roman village of Venta Silurum with the ruins of the basilica and the walls defining the castrum still visible. The ruins are now partially included in the village of Caerwent. The quite isolated grid SO4024 (S) includes the Grosmont village and the ruins of its castle.

Inside the Breacon Beacons national park we have several hot spots. One is the BNG grid SO0416 (T) that features a rail station (Torpantau Station) which is a popular start point for walks in the Brecon Beacons. The rail station has no specific interest, but its value comes from the use of it as a meeting point for nature based sustainable tourism. This area is well mapped in OSM with natural features and trails drawn in details. This grid includes also the Neuadd Waterfalls that while popular amongst photosharing users is not featured in the incremental sources. In all sections of the park the urban and infrastructural focus of OSM makes a hotspot of almost every village too small to be deemed an urban settlement according to the OECD procedure used but that plays a role in hiking trails or that have a cultural feature. In the northern section of the Park there is a cluster of grids (SO1434, SO1534, SO1533) (U) in correspondence of the village of

Telgarth where the articles dedicated to a former hospital, then converted to business park and an ancient castle contributes to its statistical significance.

8.8 Summary

In this chapter we analysed the use of incremental projects to detect environmental spaces. The methodology has been adapted focusing more on the lineage of data. Open and community led projects such as Wikipedia and OpenStreetMap are often enhanced using automated edits and data imports. While there are increasing the quantity of knowledge accumulated, they are not performed by humans in relation with space. Therefore, a considerable amount of work has been devoted to the removal of automated edits but also to the removal of the effect of erroneously designed community generated tools. The projects are mildly correlated and unfold in south east Wales favouring urban areas but also determining non-urban clusters and hotspots. They resulted to be very often overlapping with the clusters found though photosharing project. Rootedness, local history, and contentious events are highlighted by the analysis of incremental projects.

9 Discussion

In this chapter, we discuss the results of our work in the wider context described in the first three chapters, stressing the main aim of our research, the detection of CES using a wide range of CGI resources. We also discuss a first implementation of this process, which we have tested. From the CES perspective in section 2.4.2, we adopted a conceptualisation which draws clear relationships between benefits, practices, and spaces as described in 2016 by (Church et al. 2014) and (Fish et al. 2016b). In Figure 2-5 we provide a graphical representation of this ontological framework. The overall concept is based on the tenets of interactions and mutual influences between physical and cultural components. There are three main interacting blocks that contribute to CES: Environmental spaces, cultural practices, and cultural ecosystem benefits. In this ontological perspective and from studies outlined in the literature it is assumed that an analysis of CES based on habitats or land uses is deemed to fail (see section 2.5.1)[104]. Therefore, environmental spaces cannot be identified solely from geographical contexts such as the use of land use and land cover based analysis. However, they can be identified through the detection of cultural practices and derived benefits. Our approach is based on the digital traces of these cultural practices. From the CGI perspective, we will stress the challenges and solutions that have been adopted to tailor our analysis to CES focused research. We have found that CGI datasets are characterised by various idiosyncratic traits (see 3.3). This is because the category of CGI includes a wide variety of projects with different levels of coordination. Within these projects, the influence of the social component that drives human contribution patterns in space and time is critical. In the following sections, we will also show how the individual traits of CGI projects can contribute to our research agenda. This leads us to discuss the research outcomes of the analytical phase (section 9.1). We then examine the analysis to highlight the main outcomes and address the research objectives (section 9.2) as proposed in section 4.3.2. We end this chapter by recommending how CGI can be integrated and enhance the current mainstream deliberative approach methodologies (section 9.4).

9.1 The CGI resources in space

We followed the digital traces of people interacting with space. This involved activities that required interest expressed through physical presence, information gathering, and information refinement. These spatial practices are either natively digital or have been made digitally discoverable. Our study shows that people tend to prefer specific areas to pursue cultural practices. People take and share pictures more readily when they go to

[104]"explicitly rejecting linear and unidirectional constructions of the contributions ecosystems may make to well-being" (Fish et al. 2016b p. 211)

specific places (Chapter 6), people more readily record environmental information in specific places (Chapter 7), and people show care for specific places by sharing speculative information and cultural knowledge or creating geographic information about them (Chapter 8). This further confirms the findings of previous studies (Girardin et al. 2008, Antoniou et al. 2010, Casalegno et al. 2013, Tenerelli et al. 2016). We found that different families of CGI resources have different nuances of spatial behaviours that can shed light on cultural practices. The digital information component of some of those practices are left when people do participate in the collection and sharing of spatial information. Some key hotspots are shared not only among different applications but also among different CGI families. In south Wales, we found that most of the statistically significant hotspots and outliers for both photosharing and incremental resources overlap. In Greater London, we had a wider variety of surveys but the main hotspots there are also shared amongst a wide range of citizen science biodiversity recordings. In the following sections we will discuss the insights that were gained in the different phases of the research, stressing the challenges and results connecting the different case studies amongst them.

9.1.1 CGI data collection and preparation

The wide range of information that we used had different levels of accessibility. Coding was required to interact with easy to access open APIs, whereas the vast majority of projects had data accessible either through data dumps or through agreements with the data owners. The extraction of information for photosharing and citizen science data was straightforward. However, for the incremental resources, information had been collected in different ways and extraction was more complicated.

The need to work with lineage information is easily implemented with Wikipedia but it is very complicated when dealing with OSM and Wikimapia. To create a full lineage with OSM it is necessary to download and slice two global data dumps. Saving lineage information from Wikimapia requires the design of a bespoke web crawler. The possibility of extracting spatial information in the form of points or areas, and unique identifiers for contributors gave us the minimal elements to apply the analytical methodology (section 4.5.3).

The focus on localised and contextual aware edits required the implementation of a wide range of techniques to filter out uncontextualised, automated edits or data imports. The exclusion of urban areas using the OECD methodology (section 4.5.2) also was set to serve this purpose. This increased the validity of our approach since the excluded urban areas were overrepresented in all projects considered, where large clusters of high values entirely cover the urban areas. The data preparation for citizen science data was more affected by challenges related to the geographic reference of the observations, such as differently sized grids, points, and using different reference systems. The need to combine such a heterogeneous set of references was the main challenge for the citizen science case study. Even after an accurate process of geographic alignment we still found that the geographic reference could be a weak point in the datasets that could influence the outcome of our analysis. We concluded that citizen science biodiversity recordings are more concerned with the accurate identification of the species observed and they can less accurately report the exact location of the sighting. The geographic reference is often inaccurate. From another point of view the abundance of information collected, which often included contextual information, helped to identify those pitfalls. In the datasets, there is often a field devoted to the description of the surrounding environment and sometimes citizen scientists used this to include toponyms. This helped us to identify the inaccurately geo-referred observations. This way we found observations whose geographic reference was indicated as 1km accurate but that were around 50km away from the correct location. This is maybe because people without formal ecological training are more interested in species recognition rather than in the creation of an accurate biodiversity recording which also includes geographic referencing. They are maybe not interested in the correct localisation of observations.

The collection and preparation process was also challenging for incremental projects. Also in this case the main challenge was re-purposing resources intended for different uses (in a way, secondary data analysis). In the case of incremental resources, contributors are more focused on the creation of information in an either spatially explicit or implicit project. The supporting communities here are also comprised by IT enthusiasts who contribute not only by creating information but also by creating automated ways to improve it. Therefore, code has been developed to automatically check specific features of the information. This results in edits that are not expressing localised focus or interest by anyone. The removal of the actions of code on incremental datasets is therefore paramount for the scope of our project. The work on these issues was almost impossible to automate since information about code operating on datasets is not always updated and there is not a strictly enforced code to describe a way to make such edits discoverable in an automated way. We also had to create our own code to remove bugs created by badly designed editing software which was affecting all objects near real edits. This work was extensive, requiring several steps and resources to automate the validation of large amounts of data resulted in different impacts on the different projects with Wikipedia and OSM the most impacted with a 7% of objects and revisions affected. We had to evaluate these results considering two degrees of analysis. Firstly, this global 7% may have a higher local impact, considering that we had imports regarding sea buoys, or considering localised and repeated spelling mistakes. If concentrated in limited areas, this global 7%, could also have a large impact on our evaluations. We may also consider that this 7% is calculated over the richest projects we used. Therefore, in absolute numbers considering our case study area 7% of OSM objects and revisions gives 105,727. This equates to 78% of Wikipedia's objects and revisions but also 27 times all objects and revisions created in Wikimapia. This last consideration not only stresses how the data preparation process was necessary but also shows how we used a dimensionless index that we named Contributor ratio (Cr – 4.5.5.4) when comparing projects of such different dimensions.

The individuality of contributors' identification resulted in an issue that is widespread in citizen science projects but that has also some impact on Wikimapia. One difficulty that we encountered was that while the online recording mechanism of iSpot and iRecord meant that individual volunteers had unique recording IDs, this does not preclude the possibility of individuals registering multiple times. Also, unique recorder IDs are not implemented in the GiGL dataset. Not only the GiGL dataset comes from the merging of 176 surveys organised and maintained by different societies, but also most of the records are generated from the digitisation of paper cards by volunteers. We meet here factors that can contribute to the overestimation or underestimation of the number of contributors if we simply query our databases to compare the strings in the username field. We can envisage an overestimation of the number of volunteers when one individual is recorded in different ways (e.g. As "Albert Person", "A. Person" or "Person"). Underestimation of contributors is more likely since homonymy could merge two distinct individuals with the same username; when more people are recorded as contributors for one observation (e.g. "Albert Person and John Doe"); and when a collective username is used by an organisation ("Wren Group") or even more when no information is provided ("member of the public" anonymous or empty username fields). The wide variety of projects collected by environmental record centres like GiGL deserve a study on their own and this could be a very valid test to elicit CES from these kinds of collections. We need to stress how iSpot information about users was based on anonymous unique identifiers in the dataset delivered to us. We therefore cannot evaluate if there were any collective identities used, or multiple observers. Further engagement with iSpot might lead to a specific study where the incremental information is used and where we can access information pertaining to participants. The new iSpot website launched in 2017 shows individual contributor profiles which include a standardised table for reputation along with some other very specific fields and a free text field devoted to the description of the field of interest[105].

Striving to assess individual and genuine attachment for places, we need to stress how the number of Wikimapia contributors, mostly those at the earliest stages of the projects,

[105]An example of an iSpot contributor profile - https://www.ispotnature.org/view/user/65457

are underestimated. This occurred as the project initially did not require contributors to register, so unregistered contributions were assigned to the GUEST username. This feature encouraged vandalism and Wikimapia gradually introduced user registration after an initial intermediate stage where the GUEST nickname was associated with the IP address of the machine used for the edit. Therefore, 42% of Wikimapia observations are registered under the name Guest or Guest and IP address and these contribute to 25% of the total amount of objects and edits in our study area in south east Wales (section 8.3). The issues underlined here have been considered in the development of the project and their impact mitigated or acknowledged. The difficulty in identifying unique contributors to Wikimapia created two opposing effects. The use of the Guest username aggregated contributors and therefore it reduced their number. The use of Guest+IP address, if not more than one person contributed to Wikimapia from the same IP, overestimates contributions since a user can be connected from different places or the connection can be reassigned a different IP for the same entry point (Dynamic IP). The data preparation phase faced specific challenges for every individual project. Nonetheless the analytical methodology developed allowed us to treat all datasets similarly once the preparation of data had been completed. In the following section, we stress the similarities and relations between families of projects.

9.1.2 Characterisation of families of projects

The different families of projects showed that they are distributed non-randomly in spatial terms. With different magnitude, the analysed projects covered always more than 90% of the study areas. Focusing on the numbers of contributors, the non-spatially explicit projects (Wikipedia and Flickr) are the first and third projects with overall the highest number of contributor per km^2 (Tables 6-1,7-4, and 8-3). These projects are also the most polarised towards urban areas (Table 6-2, and 8-4). In urban settings, they are unlikely to focus on nature related activities. This supports our claim that the long-lasting tradition of citizen science, represented by GiGL in our case, plays a fundamental role in the detection of people attachment for nature in urban settings.

The families of projects all had internal correlations that were generally higher when considering all the grid units, and lower when considering only the non-urban grid units. This is influenced by the fact that urban areas are the recipients of high values and high correlations, while the non-urban behaviour of projects can differ markedly. To measure the correlation between projects we used the Spearman's rank correlation test. We must stress here how the test gives and index of correlation and an evaluation of the statistical significance of the measure. We always had largely statistically significant results. This is due also to the very large number of couples of grids used in our case studies. While the statistical significance is important we must stress how in our case we have zero inflated datasets. Huson (2007) found out that, while not the best option, Spearman's rank test has good results when applied to zero-inflated data. To address the large sample size a better evaluation it is advisable to derive r^2 to avoid having a significance induced simply by the very large sample. In general, we have highest internal correlations measured using the Spearman rank's test for the photosharing resources with a minimum of 0.54 for Panoramio and Geograph in non-urban areas (Table 6-4). Then we have the incremental resources where correlation for the overall study area ranges from 0.48 to 0.52 (section 8.5) to the citizen science resources where values ranges from 0.42 and 0.43 (section 7.4.1).

Focusing on the individuality of projects of the same family we detected that internally, all resources are more likely to be active in urban areas. Even the citizen science surveys have more participants in inner London where we have the most densely inhabited area. Focusing on non-urban areas, the three photosharing platforms, while being the most correlated projects, show common aspects of spatial behaviour. Flickr and Panoramio show almost interchangeable results whereby Flickr stresses more human-made artefacts and social cultural events. Geograph stresses the same places with different intensity. As the policy of Geograph favours the territorial spread of people it is more prone to underscore accessibility, but without implying a specific appreciation for the areas pictured. Knowing the purpose of the originating platform can help in understanding the social-ecological couplings detected.

For citizen science projects, it is important first to underscore how the inaccurate representation of geographic position by some citizen science projects (as noted above), casts a doubt on our overall inclusion of citizen science in the spatially explicit projects as suggested in chapter 5. We also must stress that the citizen science biological recordings used too many ways to attach the geographic reference to the observations, which introduced a level of inconsistency and inaccuracy. There were at least three ways to refer observations: As a point which stands for an exact point on earth's surface; as a point used to identify the centroid of a large area such as a park or a reserve, and lastly, as grid units of the British National Grid. Here the case was made more confusing since there are grids of varying sizes. Following our assessment of the datasets we may assume that the main content of the observations is the identification of the element of the environment while the geographic reference is considered as ancillary information. We found geographically inaccurate environmental observations even in iSpot (section 7.4.1) which provided us with a highly filtered set of data where the stringent qualitative criteria ignored these mistakenly geolocated observations. Some observations were geographically placed as far as 50km from the real observation point. The citizen science case study is more difficult to interpret given that it is a collection of collections, we have three projects of which two are the sum of a plurality of surveys.

The number of contributors per grid unit was spatially correlated across datasets, the same cells being broadly more or less productive in terms of number of records, and number of records per volunteer. The strongest correlations were between iSpot and iRecord. Given the far greater number of GiGL recorders, however, it is not surprising that the correlations between GiGL and the other two datasets were lower. But some elements can help us to understand the results. GiGL represents the long-lasting tradition of amateur naturalists. Its observations mostly come from the digitisation of paper and pencil field observations, whereas iRecord is a modern app based collection of environmental observations. Equally, the iSpot data suggests similarity with iRecord, which is mostly an app based educational and networking tools for citizen scientists. The strong IT supporting component of iSpot and iRecord could be the factor that makes their observations more correlated. The similarity between iRecord and iSpot is also stressed by the engagement analysis that has been performed alongside our analysis in Boakes et al. (2016). They argue that their use of mobile technology and the media exposure of such projects creates large numbers of one-time users of the mobile applications. The most recent app based observations fall in the categories of projects that are subjected to the participation inequality issues that we have already explained in section 3.4.2. But CGI however, contributes to the increased amount of overall environmental observations that are characterised by the low volunteer productivity as can be seen comparing the classes of the legends in Figure 7-4, 7-5, and 7-6. There are however, many different preferences among recording schemes on how they wish to receive data, with some promoting online systems such as iRecord, while others prefer to receive spreadsheets or emails. Volunteers will also have their own individual preferences for collecting data via smartphone or notebooks, for example, which may influence where they submit. Consequently, participants are more encouraged to make casual observations if they can agilely tap on a mobile for an app and submit an observation or take a picture rather than relying on pre-planned activities with the burden of notebooks, pencils, and maybe some quick visual guidance for identification. Traditional observation is probably more accurate but also less casual and restricted to areas where the volunteer is more likely to post many of them. Some taxon focused groups collect records via a network of ‘county recorders’ while others have a national recorder. Local environmental records centres collate data for all taxa for their geographical area. Thus, a given individual recorder may choose, or be steered, to provide data to different collating systems depending on what taxa they record, in which location.

Of the three datasets, GiGL has a much longer history, and many of the contributing volunteers have long lasting relationships with natural history societies and recording groups who share their records with GiGL, such as the London Natural History Society (LNHS). iSpot volunteers have a different motivation from recording societies' – observations are posted in order that the observer can identify a species. iSpot records thus have two key differences to those in GiGL and iRecord: (i) the observer does not need to be able to identify the species, (ii) the observer needs to supply a photograph of the species. The former removes bias caused by knowledge barriers that might arise in GiGL or iRecord but the latter imposes bias that will not be present in the other two datasets – some species are easier to photograph, and thus record than others, for example a bird on open water versus one singing inside a hedge. However, there is no evidence to suggest that these two criteria have influenced data collection. We want to stress here, how the possibility to analyse the original unfiltered iSpot data could lead to different results since. In this case, we may have a better picture of the incremental process that has been removed from the data provided to us. At the same time, the incremental process would be focused not on information about space and place but information about the observation and the taxon, so it may not be useful for the spatial characterisation of CES. However, the inclusion of observations, where identification has been considered below the standard to be provided to us, could provide more information about spatial preferences.

The incremental resources were also characterised by heterogeneity in terms of spatial reference. In terms of geographic detailing of information, we have three different levels. The simplest approach is the one of Wikipedia where geographic information is associated to a point. In Wikimapia there is the possibility to insert areas and the API derives the centroid of such an area. In OSM data there are areas but also every point contributing to the shape delimiting the area. In incremental projects also, the lineage information is referred in a heterogeneous way. In Wikipedia, it mostly focusses on textual information since centroids are almost never moved. In Wikimapia the lineage is associated with the article and the overall shape associated to it. There is no detailed recording of shape changes at the level of individual points. In OSM conversely, the lineage is associated with both the area and with the individual points but there is not a centroid for aerial entities. This heterogeneity has been overcome applying the quadrat count methodology stressing that we are interested in localised contributions. Our analytical method overcame these differences but for such resources that are loosely related to the BNG it may be useful to run sensitivity analysis with alternative grid units.

The incremental resources are more focused on space and place and are mildly correlated showing stronger correlations only in the urban environment. Individually we see how they differ in terms of participation to map urban and non-urban areas. Table 8-3 sums this up showing a disproportionate situation in terms of participation and productivity. While Wikipedia has more than 94% of contributors which covers just 31%

of the grid units, OSM provided 92% of total objects and revisions and Wikimapia had the second largest number of localised articles. Table 8-4 shows how differently those projects behave. The spatial focus of the project is highly influencing the spatial spread of contributions and the number of contributors engaged. Wikipedia, which is primarily focused on textual knowledge, has in urban areas, on average more than 11 times the number of contributors that it has in non-urban areas. Wikimapia has a more spatial aware approach with this relation reduced to 7.5 times while OSM which aims to cover all territory, has 2.4 times the number of contributors in urban areas compared to non-urban.

When we focus our attention to the characterisations of projects of the same family we have that photo-sharing services have been characterised as follows: Flickr is significantly more successful than the other platforms in terms of pictures and contributors. Together with pictures related to the surrounding natural environment, photos are more focused on human environments and activities such as festivals and events. Panoramio's imagery is more evenly distributed than Flickr's and tends to have an evenly spread coverage that also encompasses more natural areas. Flickr and Panoramio reveal very similar photosharing behaviour in non-urban areas. In contrast, Geograph covers almost all the territory and may be a better proxy for accessibility estimations. Nevertheless, for our purposes, some limitations of the Geograph platform may be deriving from the gamified approach that puts emphasis on leaderboards, rewards, and games.

Citizen science projects have been characterised as follows: GiGL surveys reflects the spatial and temporal attitudes of amateur naturalists who have been collecting information for societies. We have both a social engagement and an individual engagement. We have the role of societies but also of long lasting amateur naturalists. Therefore, from one side we have contributors who have individually collected thousands of observations from their gardens. From another side, we perceive the role played by wildlife organisations. We have traces of observation expeditions where organisations play an important role and where the spatial target is a well identified and well known natural area, a park, or a reserve. Here we have shared observations which are anonymised and associated to wildlife organisations (such as LNHS survey 2013), but also initiatives to engage with nature at school level. The app based platforms iSpot and iRecord give us another perspective on the urban engagement for nature. We have here less productive contributors whose distribution is concentrated in more densely inhabited areas. Occasional observers are defined as 'dabblers' in the studies classifying the engagement for projects as in Boakes et al. (2016). This kind of contribution places iSpot and iRecord well in the range of applications used by younger people, less prone to plan in advance, whose reliance on mobile apps is modifying the way services and entertainment are delivered. But this also bears witness to how those apps can target ranges of population that are less likely to participate in participatory mapping initiatives. A better characterisation of those project has been achieved using other data included in the datasets such as the time of the observations and the taxonomic groups of species observed (Boakes et al. 2016).

Incremental resources have several levels of complexity to consider. Firstly, we had to aggregate them based on the way the collaboration between contributors is shaped, because these projects are explicitly and primarily collaborative. This also means that they can entail more conflicts and controversies. For OSM and collaborative mapping in general, so called tag wars have been detected elsewhere (Ballatore 2014, Bittner 2017) but notwithstanding the bilingual nature of Wales and the latently conflicting nationalisms involved, we have not detected such issues. OSM is the most successful collaborative mapping project and very often the internal debate of the community through mailing lists and forums focuses on the creation of standardised road network information. Nonetheless OSM is highly contributing to the creation of spatial information that goes beyond road networks. It is noteworthy how the most popular non-urban grid unit for OSM in the study area is the one containing a mainly natural area with the ruins of Tintern Abbey. This comes above suburban areas which contain more mappable elements, not only in terms of buildings but also in terms of infrastructure. We may argue that contributors more interested in local features may be less interested in the definition of standards and are less visible in the communication channels used by the community.

The conflicting nature of knowledge and its temporal depth and development is more reflected in Wikipedia's articles. Amongst articles with a geographic tag in this area we have mostly geographic entities in terms of physical or administrative units. Time here plays a role in two dimensions. One is about centuries or millennia. In fact, particularly successful are the human settlements with ancient developments, mostly Roman and some pre-Roman settlements, even if they are currently small villages. Time also plays a role in the most recent developments. We found some of the administrative units have high numbers of contributors due to timely development in terms of electoral results included in Wikipedia. Amongst these conventional geographic entities, we also found that the worst environmental disaster of the last 60 years in Wales is ranked the 8th article, with more contributors in the area. The Aberfan disaster is the worst tragedy in modern Welsh history. It happened on 21st October 1966 and claimed the life of 116 children and 28 adults. It was due to the collapse of a colliery spoil tip over the village and which directly hit a junior school, killing almost entirely a generation of residents of the small village (section 8.5.1). Another relevant grid is related to the disputed story of Cardiff's airport. Those elements show how Wikipedia is a way to get information about the rootedness as connection to the historical development of places as well as its contemporary issues and political activism.

Wikimapia is the more place focused resource since its generation comes through the annotation of online maps, its supporting community is not really interested in basic mapping for cartographic and mapping purposes. The community adds information to the basic geographic information already used as a background. Therefore, annotations are usually very significant. We have here also to stress how the API of Wikimapia does not provide road network information that can be created by senior contributors only through the website and is visible overlaid on the online map. On the relevance of the textual information associated to Wikimapia, we need to note how very often Wikimapia's articles include a link to Wikipedia. Therefore, the textual information included is intended as a general introduction while the detailed description is left to Wikipedia. Wikimapia therefore acts as a bridge. It complements the background imagery coming from external sources with geographic, place specific information that is then linked to more detailed textual information in Wikipedia. But its map based development stresses locality and often underlines cultural and natural features of interest that do not often emerge with the other platforms.

9.1.3 Clusters and hotspots

The overall statistical analysis was aimed at the characterisation of projects to evaluate their potential as tools to find environmental spaces with specific characteristics.

The spatial statistical analysis aimed at characterising the potential of the families of projects as tools to detect environmental spaces. The statistical deduction of spatial clusters and hotspots helps in finding aggregations of grids or individual grids whose value of Cr differs markedly from the surrounding areas flagging the grid or a group of them.

Focusing firstly on clusters and hotspots from photosharing we noticed that while clusters and hotspots are located both in urban and in non-urban environments, we directed our attention to the importance of elements related to open spaces and natural environment. When photosharing is involved the combination of four factors influences the trends. One element particularly evident when we analyse pictures shared by people, is the presence of accessible views over specific landscapes (peaks and beaches); this forms the basis of larger clusters, with high numbers of contributors. The second element is the presence of historic and picturesque human artefacts such as ruins, castles, and heritage houses, which are relevant when we consider not only photosharing but also incremental resources. A third factor is the one that identifies places organised to stress the role of biodiversity, such as the National Botanic Garden of Wales and the Newport Wetlands visitor centre. They are characterised by the concentration of formal biological recordings, and by the concentration of pictures that can be used to enhance biological recordings as in Barve (2014). Those places have also been highly popular amongst citizen scientists. The fourth factor we recognised is the presence of surface water. The importance of blue areas has been stressed in all our analysis conducted in Greater London. Using photos in section 6.7 or analysing citizen scientists in urban areas (section 7.5.2). Very often, we identified a combination of these five elements where events are organised. This can be assumed as the fifth factor. For example, in Flickr mainly, there was a wide range of pictures taken in the areas around Tredegar House and the Newport Wetlands visitor centre, including distant objects and landscapes as well as close-up shots of flowers and insects.

When focusing on volunteer-led biodiversity recordings, we found clear recording hotspots (section 7.5.2). Here we have five recurring elements that influence the aggregation of volunteers. Hotspots may be attributable to sites particularly rich in biodiversity, sites easily accessible to volunteers, sites which organise surveys with members of the public, sites regularly visited by recording societies, or to the influence of outstanding recorders (two cell had > 30.000 records from a single recorder each). The important role and the statistical significance of the London Wetland Centre and Rainham Marshes Nature Reserve as well as the main trend identified which follows the river Thames and the green areas overseeing its flow up to Battersea, stresses the important role of the association between green and blue areas. The two environmental centres have high species diversity, observation hides and visitor amenities and thus are likely to attract volunteers with a wide range of interests and experience. The strong preference for open waters has been explained by Boakes et al. (2016) because the vast majority of observations are targeting birds. Bird watchers seem to favour sites with open water, perhaps because of higher visibility and species richness. Recording societies considerably influence the spatial spread of volunteers favouring specific parks and reserves. Organisational factors emerge as relevant also in the citizen science case study. Noteworthy our analysis found out how also urban areas without specific nature focused infrastructure but with urban private gardens and low-density developments can favour the connection with nature. In some cases, we had very active individuals collecting thousands of observations in their gardens in low density neighbourhoods, or areas where several volunteers recorded in their private gardens sustaining the creation of a hotspot.

We analysed incremental projects excluding urban areas in south east Wales and we found a remarkable similarity with hotspots and outliers found in the same area using photosharing projects (section 8.7 and 6.5.4). Therefore, the same factors identified from photosharing platforms are also highlighted by incremental projects but with different degrees of influence. The visual element is here less important. But all the remaining factors are persistent and combine natural, social, and organisational elements. With incremental projects, we have found that open space meeting points, places where people meet and socialise waiting to engage in a main activity elsewhere can lead to localised production of information. Two points especially were both focus of pictures and of detailed geographic information in south east Wales, a large parking in a service area (West Cardiff) where car rallies meet. A similar point has also been raised in the photosharing sub project (section 6.6) when amongst the top grid units for photosharing we noted several rail stations even when they do not bear iconic value. Apart from these underlining factors, incremental projects can be used to witness rootedness, and contested knowledge such as in the case of the Aberfan disaster article or the political activities of local authorities. Therefore, focusing more on natural elements, we can underline how the interest of the population in natural features and open spaces is not confined to natural reserves and other designated areas. Some natural elements are more likely in the vicinity of most intensely contributed areas. We underline the role of open views and blue areas, or the combination of both. The aggregation of people and the traces of their engagement for the surrounding areas is supported by the accessibility and organisation of cultural practices. A similar trend characterises human-made artefacts, the intrinsic nature of which is characterised by a higher degree of accessibility.

9.2 Addressing the research objectives

Having placed our research in the correct disciplinary environment and after having gone through a review of relevant literature both in CES and CGI, in Chapter 4 we have already addressed some research questions (section 4.3.1) which arose at the end of the Introduction chapter (section 1.6). In chapter 4 then, we proposed the main research objectives of this work as follows:

- *Demonstrate that a wide range of crowdsourced geographic information can be used to detect people's attachment for places.*
- *Understand and identify which components of Cultural Ecosystem Services are more likely targeted with the use of a wider range of crowdsourced geographic information*
- *Develop analysis methodology and research approach appropriate for the operationalisation of CES investigation using CGI*

In the following sections, we address them referring to our analytical work as it has been produced in chapters 6 to 8. We are also placing them in the overall domains as analysed in chapters 2 and 3. The implication of those results will be also projected in the overall picture of the relation between society and nature as depicted in chapter one.

9.2.1 Research Objective 1: Demonstrate that a wide range of crowdsourced geographic information can be used to detect people's attachment for places

During our literature review we have seen how CGI is already used in various ways to acknowledge place based knowledge from participants. In our case we explicitly targeted how CGI has been used to elicit qualitative value of places (sections 2.5.3 and 3.5). The picture that came out was unbalanced, in favour of photographic collections (sections 2.6.3 and 3.5.7). These are to date, the most used CGI sources for CES investigation and best practices from them have also been achieved so that our methodology (section 4.5) owes to the successful implementations and limitations of those early adopters. We have not only built on top of the methodologies used for photosharing, but we have also embraced biodiversity recordings contributed by volunteers that previously had almost no adopters, and we have pioneered the use of lineage information from sources which collaboratively and incrementally create knowledge. The latter to date have not been used to elicit social-territorial bonds. Through the analysis of CGI, we found that statistically people show their spatial bonds using CGI. The statistical analysis showed how the nine different CGI sources we used have all their individual say on the unfolding of people's attachment for places (sections 6.5.1, 7.4.1, and 8.6.1). Grouping into families of projects here, we also had a considerable amount of overlap in terms of spatial clusters when considering some of them (sections 6.5.3, 8.6 and 6.7.1.2 - 7). Using only one specific project pertaining to a specific family gives interesting insights but it could lead to an incomplete picture of the cultural practices and spaces to aim at. Therefore, the use of CGI for CES analysis should be very inclusive. Another factor that could lead to better results when embracing different CGI resources comes from the fact that different families of projects and even different projects can have different success rates in space, time, and demographic terms (sections 3.4.2 to 3.4.5). Some resources may also be subject to disappearance such as Panoramio, whose use was featured in a considerable amount of studies targeting CES and in our first case study. In our analysis, we also found that spatial biases may be related to the technology used to record information, as in the case of citizen science where the two app based projects showed similar patterns, that were weakly related to the long-lasting tradition of volunteered non-digital biodiversity recordings as collected by GiGL. This can be seen when comparing the patterns of Figure 7-1 with Figures 7-2 and 7-3 and considering the correlation analyses 7.4.1. These weakly related spatial biases may reflect the fact that the different ways of engaging with nature are targeting different ranges of populations, which also show different attitudes towards the use of technology and to engagement with nature. Therefore, there is not only a combination of platforms covering a wide range of experiences and practices that connect humans to nature, but the variety of motivations and engagement types involved can help to overcome cultural and age based participation biases that affect mainstream participatory and deliberative approaches.

9.2.2 Research Objective 2: Understand and identify which components of Cultural Ecosystem Services are more likely targeted with the use of a wider range of crowdsourced geographic information

Referring to the CES dimensions in the conceptual framework we adopted in 2.4.2, we first identified them in CGI literature that was not explicitly targeting CES (section 3.1). To visualise to which of the components depicted by Church et al. (2014) have been found through our use of three families of CGI, in **Error! Reference source not found.** we developed a visualisation where we overlaid to the diagram developed by Fish a colour-coded representation of the three CGI families. We assigned the blue colour to the photosharing projects, the green colour to the citizen science biodiversity recordings and the grey colour for the purely incremental projects. In the following we list the practices, benefits, and environmental spaces that we found during our analysis. People express their cultural preferences by moving toward specific places and by sharing their experience, or caring about the natural environment, or creating information about them. Therefore, they are playing the double role of receptors as well as reporters of CES. The action of taking and sharing pictures has been assimilated as a cultural practice in various studies (section 2.5.3) but it is a cultural practice which aims at all human interest and not specifically to ecosystem generated benefits. As in the literature, we detected clusters and hotspots both in urban and non-urban environments (sections 6.5.2 and 8.5). In our work, we excluded urban areas and focused on the open spaces since most studies have stressed how the urban environment is usually the primary target of the contribution in photosharing. Therefore, in photosharing and in the case of incremental resources, we assumed that urban environments have high "distracting" potential and so, filtered out urban areas (section 6.4 and 8.4). While in the urban environment we analysed citizen science activities which are cultural practices aiming only at the natural environment.

We have found that some key hotspots are shared amongst different families of applications and amongst applications. It is striking how photosharing applications and incremental resources had the same high value hotspots found in south east Wales (sections 8.6 and 8.7). The vast majority of non-urban clusters and hotspots are focused on cultural features, either human-made or natural. As we expected, social-ecological couplings occur also outside of urban and designated areas, stressing here also the role of cultural heritage elements. Although there are different CGI services and communities involved, some areas have been more popular, and arguably, this is an indication that those places are “worth sharing socially” through pictures when visiting them or caring about them through information. We have already identified the factors that influence the trends (9.1.3). These factors are detected more selectively in some CGI resources than others.

Photosharing can be considered a cultural practice when it is used to share and spread knowledge of a place for its specific features. More precisely, we can stress the role of environmental spaces such the landscapes that are included in popular panoramic views, as bearers of significant benefits such as **experiences**, as well as being the potential repository of **identities**. Even more, we can consider those landscapes bearing **rootedness** and **sense of place** and **belonging** when we consider the large numbers of heritage monuments such as castles, ruins, and heritage mansions or houses that are featured in the hotspots and clusters elicited through our analysis. Kenter (2016c) stresses how sense of place is interwoven with cultural heritage and the stories related to places and artefacts, this involved shared and collective values which are interpreted through and with historical depth. These cultural elements are found in open landscapes sometimes with gardens, sometimes in proximity of scenic areas. Human culture, **rootedness** and **identity** are also reflected in the temporal depth that has been unveiled by the analysis of incremental sources. Both long term historical **belonging**, and more short term political activism can be linked to detailed description in encyclopaedic articles, as well as the creation of detailed data for maps, or the augmentation of maps with circumstantial information about events or other cultural elements not visible in aerial photography or topographical maps.

In urban areas we used environmental citizen science which is already a cultural practice. It allowed us to identify environmental spaces like **blue spaces** predominantly associated with surrounding **green spaces** and the main trend in Greater London that connects them (section 7.5.1). These settings are particularly favoured by volunteers observing birds (also called birders), according to the taxonomic analysis section of Boakes et al (2016). This could be due to the co-presence of a bird-friendly environment with open views. The role played as environmental spaces by places actively organised to focus on the biodiversity emerges from the urban analysis of citizen science biodiversity recordings. Citizen science activities are eminently related to the development of capabilities such as **knowledge**, **dexterity,** and **judgement**. In a very similar way also incremental resources are creating knowledge about places. They are capable of intercepting benefits that contribute to local identities digging more in depth the historical development of settlements and communities. Their peculiar contribution stays in intercepting narratives which are linked to cultural aspects who have very limited, invisible, or marginally visible physical components. Human **capabilities** and **experiences** are greatly enhanced in these places, as we found in our study areas in south Wales, the National **Botanic Garden** of Wales, and the Newport **Wetlands visitor**

centre. Here, close-up photographs of natural elements resemble the kind of images used to identify and correctly classify nature in biodiversity recordings. Barve (2014) used the images from photosharing platforms to enhance biodiversity recordings. Kenter (2016c) stressed how biodiversity can be related to tourism and recreation. In fact, there is a degree of overlap between citizen science and photosharing when dealing with nature. Biodiversity recordings performed by volunteers are already listed among cultural practices related to **producing and caring** (Fish et al. 2016b).

We detected three trends in the analysis of citizen science data. One is characterised by the traditional amateur naturalists often connected through conservation organisations. In this case more social and shared activities introduce a spatial bias towards areas deemed to allow more fulfilling excursions. Some organisations run surveys explicitly and exclusively targeting well defined parks or conservation areas thus enhancing observations in that specific domain and reinforcing the *fortress* or *island* approach (Palomo et al. 2013) to conservation rather than unveiling the daily benefits we get from nature. Conversely the second trend is evident from the spatial distribution of mobile app based observations. Those observations are less planned, arguably more solitary, and spatially tend to be less disperse and concentrated in more densely inhabited areas. This way citizen science is arguably embracing a wider and more diverse or complementary range of population. The third trend is characterised by garden originated environmental observations. These can comprise in some cases, many observations from specific individuals or they can be a more widespread activity. They have been detected in **low dense urban areas** which can also be considered a sort of environmental spaces.

When analysing hotspots originated with the leading contribution of conservation organisations we have a predominant benefit in enhancing **capabilities** to learn from the environment, mostly while in the field. This is enjoyed in the Welsh locations mentioned above as well as in London's main hotspots in the **Wetland Centre** and **Rainham Marshes Nature Reserve,** but also in lesser known but rather central natural parks such as the **Carmley Street Natural Park**. When we think about the app based less planned observations, we are targeting a more immediate **experiential** relation with nature. It may be enhanced towards learning about the environment, but this process may be deferred to the social interaction mediated through the web or if known through iSpot or similar resources. The third trend may be the sign of a stronger **sense of belonging**. We can stress how in almost all situations benefits are bundled and coexist where practices are organised. Therefore, the interest of the population in natural features and open spaces is not confined to natural reserves and other designated areas, but also in those places it is supported by the accessibility and the organisation of cultural activities. This is favoured by the presence of human-made artefacts. In their vicinity then we have the successful organisation of events that may lead to the production of cultural goods.

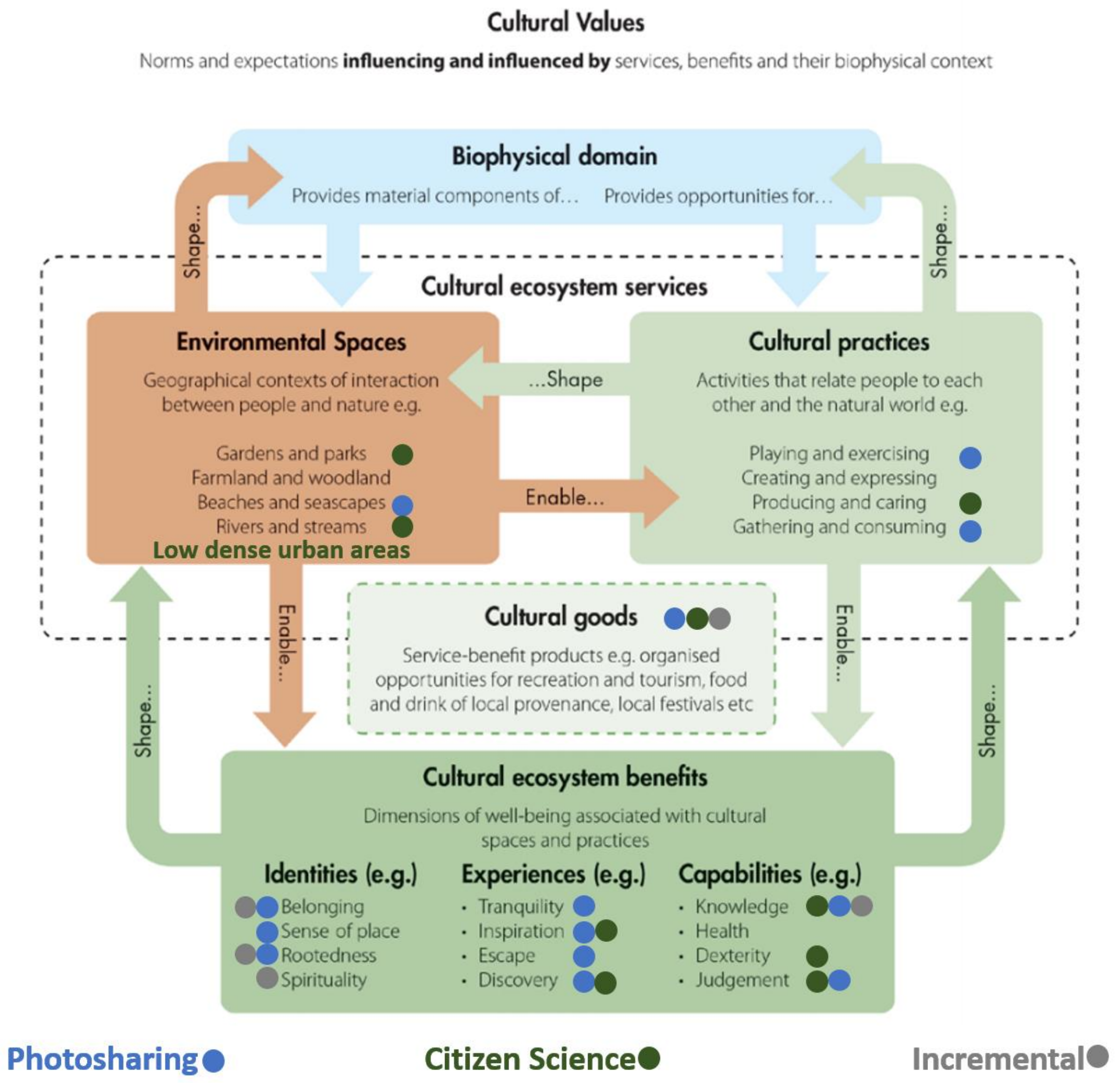

Figure 9-1: CES components elicited in this work identified in the framework for CES designed by Church et al. (2014)

9.2.3 Research objective 3: Develop analysis methodology and research approach appropriate for the operationalisation of CES investigation using CGI

Our analytical work stemmed from the analysis of methodological information in literature. In chapter 2 we focused on the works explicitly aiming at the detection of CES while in chapter 3 we analysed the methodologies of studies using CGI to detect social-spatial bonds without framing their work in the ES approach. Even the combination of the best practices needed improvements to adapt our in-depth knowledge of CGI to support the detection of environmental spaces. Our resulting methodology stands out for the improvements it includes, and because it derives from an in-depth knowledge of the modalities of CGI data production. Our approach has been different from most studies that used CGI since we did not directly assume that the data was fit for the purpose, but we engaged with communities generating data and then we analysed it. We tried to understand not only the semantic behind the data structure, that the organisation coordinating volunteers put in place, but also the internal functioning of communities of practices. Our aim is to track human perceptions, cultural practices, and benefits beyond technological affordances. We have several recommendations that derive from this study. They are all aiming at the operationalisation of the use of CGI sources to improve CES detection.

Recommendation 0: **The projects characteristics and data production practices must be known**. Data has not to be taken for granted, and then used. We have examples of evident misuse of CGI data in literature since researchers did not know the projects. There are examples of works that used data where its quality has never been assessed as in (Seresinhe et al. 2015, 2017, Upton et al. 2015, van Zanten et al. 2016). To cite only one of the improper uses of CGI we have the use of data on building derived from OpenStreetMap. Such data has been used to check if pictures contributed to a photosharing projects were taken indoors. There are no studies on the completeness of building coverage on OSM and even a very brief check of the online map of OSM gives the immediate awareness that the provision of buildings is extremely scarce and for buildings OSM cannot be used a replacement for authoritative data.

Recommendation 1: **The communities that collect data must be known**. Not only the organisation behind the database design but also eventually the technological asset of the project. In fact, in section 3.3 we suggested to broaden the components of CGI to include also the community support services. To collect and correctly interpret information, we have not relied solely on academic publications but have been engaged in mailing lists, and read and contributed to the creation of information. Very often CGI contributors have conflicting ideas about data quality assurance methods. There may be a group of volunteers who considers strict standardisation and centralisation of ontological and linguistic structures a primary target, while others are keener on allowing the free flow of local original and peculiar information. Knowing how opposing instances in a community unfold can help understand the relevance of data and the processes that are needed to make it operational.

Recommendation 2: It is of paramount importance to know and evaluate **fitness for purpose of data** in several ways. This requires recommendation 1 to be implemented. We have subjected our datasets to extensive data cleaning and filtering processes. But even with this careful preparation, data wrongly geolocated has influenced our statistical analysis as we have seen in the case of iSpot. This derived from the fact that locational information was not deemed essential from participants as the main scope of the project was to learn how to accurately classify environmental observations.

Recommendation 3**: Bespoke data preparation.** We did not only rely on statistical and geostatistical methods to validate the significance of our results. Willing to assess our work thoroughly, we went to analyse the factors that lead to the statistical significance of specific areas. We often found that the data preparation phase and the filtering procedures, while widely shareable, must be designed and tested individually for every platform as every platform and community using it has individual technological and behavioural similarities and differences. We even have seen that the app based biodiversity recordings show a similar pattern, which is dissimilar to the one shown by traditional amateur naturalists. Moreover, we have seen that platforms and groups of platforms need bespoke data preparation processes. As an example, only some citizen science biodiversity recordings can contain collective usernames (X recording society volunteer) or even multiple names associated to an individual observation ("Tom, Dick, Harry, etc"), while in environmental citizen science we have not unearthed automated contributions that are characteristic of incremental sources only.

Recommendation 4: **Ethics** in data from crowds. Using the data, we have been able to follow people in time and space. Most of the openly public APIs allow downloads of data where people are identified through nicknames or name-surname couple. The combination of several resources may lead to the creation of personal profiles towards a kind of *web demographics* (Chow 2013). This entails considerable possibilities but also considerable risks. This issue has been labelled geoprivacy[106]. The adoption of this term is not widespread and it still ambiguous since it is found associated with the need to obscure the position of rare or endangered species in biodiversity recordings as in iNaturalist (https://www.inaturalist.org/pages/geoprivacy). Governments are implementing regulations to ensure the fair use of data and the implementation of CGI for CES must be considerate in the use of details that may emerge analysing data.

Recommendation 5: **Project continuation and sustainability**. Some of the projects we used are subject to market rules, therefore they may be soon fold or be downgraded. This work has already been affected by the volatility of the market with the suppression of Panoramio, one of our sources. Most of its social features have been obliterated by Google, that used it simply as a source of images to enhance Google maps' experience. Amongst the nine sources remaining two are eminently commercial (Flickr, Wikimapia). Two are subsidised by geographical or environmental public bodies, (GiGL, Geograph). Two are subsidised by research communities (iSpot, iRecord) and the two remaining are open source community that rely on sustainment by volunteers and through sponsorship (Wikipedia and OSM).

This issue can be tackled working with interested parties and ensuring the adoption of CGI resources by government agencies to ensure project continuation and sustainability (Haklay et al. 2014). whilst also considering the other forms of benefits that government

[106]http://stko.geog.ucsb.edu/geoprivacy2015/

can derive from CGI. Government involvement can also support and enforce the implementation of open licenses and data systems.

9.3 Limitations

This study was an exploration that wanted primarily to expand the horizon of the quantitative analysis of CGI for CES research. Before our study, academic focus only targeted photosharing websites. They have been used as the first methodological attempts to help shape our methodology. There are several assumptions that we made that have therefore limited the range of our accomplishments. In the following we describe the limitations of our approach from the quantitative and the qualitative point of view.

One main limitation has been that we have used secondary data without reaching out more in depth to the contributors who created it or the communities in which it originated. We have not analysed our samples of population, delivering for instance an online survey. We expected that we had different sets of population engaged in different projects, also in consideration of the IT literacy required. We know from previous studies and literature that users of some specific social media fall into certain age ranges, And the population sample can show biases in spatial, gender and temporal participation, (section 2.4.2 to 2.4.5). We have not investigated with people their reasons and motives to choose specific places. We have also relied on literature in this case but did not pursue the possibility to contact contributors on those projects where the possibility existed. However, the way in which most projects are designed, census like information is shielded from public consumption, and we would not have known if we were considering the same individuals contributing to multiple projects.

Our work enhanced specific components of the territory and several management practices that may be influential for the development or the enjoyment of cultural practice. We decided to explore the breadth of several kinds of CGI platforms rather than go in depth, developing a regression analysis or machine learning based techniques. These have provided some interesting results when used to analyse human behaviour and preferences in natural reserves when analysing Flickr pictures (Richards and Friess 2015, Walden-Schreiner et al. 2018). For this purpose, we used a limited variety of secondary data. There are several sources of geographic and qualitative information that can be combined to create a regression analysis. In the UK for example, there are several surveys which deal with cultural practices and tourism (Natural England 2014). These are seen as complementary to the most known and widely used MENE survey (Natural England 2016). Moreover, National Trust provides GIS data concerning their holdings and open or limited access (by foot) to open land (http://uk-nationaltrust.opendata.arcgis.com/). We could use 3D GIS data as already used in some studies (Tenerelli et al. 2017, Yoshimura and Hiura 2017) to apply viewshed analysis for the detection of favourable points for panoramic views. The usefulness of a spatial based regression analysis stands in the insights that can be derived on non-spatial factors influencing people's attitudes towards the environment, and to underline the role of different management practices, and access management. It could lead to an enhanced multicriteria evaluation combining spatial and non-spatial factors.

Another limitation we can highlight here is because the approach towards the use of Incremental data has been quite basic. Complex tagging schemas are associated to the information. Especially for OSM there is also the possibility to analyse the other data primitives used, such as lines and relations. But given that the results discussed in chapter 8 are already highly significant, this may only lead to limited improvements.

9.4 Integrating CGI analysis in CES research, unleashing the online deliberative process engagement

CES focused research is strongly geared towards the elicitation of values and benefits through multi stages and multiple methods of engagement in what is mostly a deliberative process. The drawbacks of those methods have often been found in the limited regional areas for which these studies can be coherently developed, the time and resources needed to implement all the phases, and in the difficulty to reach out of "usual" populations samples, usually voluntarily samples. Rather than replacing those methods, a preliminary analysis of CGI, the engagement with CGI communities, and the production of maps with areas or elements highlighted for the participatory process can enhance CES research in a variety of ways.

In the following we speculate on the many ways and stages in which CGI desk based preparatory studies can enhance CES research. Using our methods to manage large datasets could lead to the definition of maps at national level highlighting hotspots and clusters. By using aerial imagery and images taken from the CGI platforms, several ways to reach out towards the public can be created. After this, large scale national implementation of the methods and approaches could produce various visualisations and we can envisage several streams of development. These rely on the two parallel ways of engaging with populations, online and by co-presence.

Citizen science can teach several lessons on the way to hold parallel and complementary ways to engage with people. The body of knowledge in this rising discipline is a growing asset not only in the use of IT, but in community management and the management of people. One of these lesson is that to recruit a community of people interested in a specific project it is wise to engage with communities already interested in similar issues. Geograph, iSpot and some collateral projects related to OSM are directly intended to enhance participation promoting leaderboards and nice visualisations of data from contributors. Enhancing these approaches and connecting them enables the recruitment of online local "champions". Online engagement can be developed alongside the more traditional approaches used in the deliberative processes already described by Kenter (2016c) and Fish et al. (2016a) and workshops could evolve as platforms where the two different communities, nurtured online and offline, can meet and discuss.

We encourage this approach since we expect that CGI sources are mostly used by ranges of populations usually not engaged in participatory deliberative processes. Therefore, engaging with them and allowing them to know that their contribution matters may lead to overcoming the social barriers of self-selected population samples. We have also seen how the technologies used to collect information in the case of citizen science influence the spatial and social footprint of engagement, and more closely match the needs of those sections of society geared towards immediate and instant access and immediate sharing of experiences. Another advantage of the use of CGI relies in the fact that the CGI platforms are not selective between locals and tourists, while very often the other methodologies choose to engage either with tourists or not. This way the deliberative process chooses to disregard the cultural benefits that one of these populations has from the territory. Engaging with CGI communities will also help to improve the data preparation procedure which resulted in being the most cumbersome part of our work. We expect that a coarser data preparation process would be implemented at national level, with a more detailed and tailored fine-tuned process to match local singularities that can always occur when the focus is geared from the national scale to the local environments.

From a wider perspective, we see how the cultural and local knowledge aspects inform all ES. Therefore, as advocated by Costanza et al. (2017), we could argue that the adoption of deliberative practices might support research and policy formation involving all ES. Hence deliberative practices, strengthened with CGI data and communities, and integrated with citizen science engagement and retention practices might became mainstream approaches in research and subsequent policy formation for a wellbeing centred development of economy where the role of nature is enhanced. In the following chapter we summarise the main conclusions and expand on the wider perspective hinted above.

10 Conclusion and future developments

In this chapter we conclude by summarising the development of the thesis in section (10.1). The main results of the work are shown in a specific section (10.2) while the potential of the approach developed is in section 10.3. Section 10.4 is a reflection on the ethical implications of this work and of the overall approach. We conclude with section 10.5 which envisages future developments and opportunities for the overall ES approach.

10.1 Summary of the thesis

The work this thesis presents has demonstrated the potential contribution of Crowdsourced Geographic Information (CGI) to ecological research. It addresses the need to support research aimed at the detection of areas which contribute non-material benefits to human wellbeing. Such benefits are called Cultural Ecosystem Services (CES) once we embrace the Ecosystem Service (ES) approach.

Chapter 1 shows the specific traits of social and technological developments that led to CGI on the one hand, and on the other, the development of the environmental discourse in the last decades that led to the ES approach. Both domains have developed towards a more inclusive role given to human society and culture. Technology has had a dual effect on society. Transportation enhances accessibility through increased mobility while IT enables and supports user interactivity with other humans, but also with online platforms allowing people to contribute to the collection and refinement of information. Technology has then enhanced mobility, accessibility, communication; all elements that are creating new social and spatial morphologies. This has led to the definition of new digital spaces for social interaction and new social communities with shared understanding and values. New social-technical systems rather than simply user generated data have been created.

The environmental discourse needed to overcome the ever-growing dichotomy between environmental preservation and health on the one hand, and economic development on the other. This also considering the rising health and environmental fallouts of short-sighted and profit-focused investments in economic development. In doing so the environmental discourse started mimicking the language of economics, defining the stocks of natural resources as Natural Capital (NC). Effective policy design and implementation needs the effective engagement of people in open democratic societies. Thus, the definition and the advancement of the concept of sustainable development. It is the development characterised by environmental responsibility, social equity, economic viability, and cultural vitality. More public engagement in the ES approach highlights the connection between human wellbeing and NC. There has been an increase in efficacy and policy implementation as a result of the environmental discourse including society and stressing the role of people engagement.

Meanwhile, participatory research practices such as citizen science, have been rising in importance and receiving governmental support. The detection of cultural aspects and their links to human wellbeing is therefore of paramount importance. However, a gap has been identified in the detection of CES due to the non-materiality that characterises cultural benefits. The limited attention this has received in research is in sharp contrast to its importance, since cultural choices, knowledge systems and traditions are behind every use of all the natural resources that makes NC. Choices transform capital in wealth, choices transform NC in ES. But new technological affordances together with transdisciplinary approaches have the potential to bridge those gaps. To envisage better integration between policy making and new social and technological affordances we focused our attention on the ES approach and CES research, while also delving into the social-technical systems that are CGI, and the expertise they offer.

Chapter 2 investigates the current discourse on ES stressing the concurring streams of research and conceptualisations. This chapter underlines some criticism, and explores the methodologies employed to analyse the spatial extent of ES. Several studies refer to a cascade model which assumes that ES are spawn of NC. This shapes the main research methodologies which use geographic data on natural resources to map, evaluate, and create models on all ES. Other conceptualisations are critical of the cascade model and we have described some of them. When we focus on CES, we reveal that several studies conducted from a NC centric perspective acknowledge their limited correspondence to locally perceived CES. They are also characterised by a distinctive spatial distribution since the carrier of the benefit is information that has not material place in maps.

Those idiosyncratic elements of CES led to the definition of a new bespoke ontological framework during the development of the UK National Ecosystem Assessment (UK NEA). In this framework it is clearly indicated how CES are non-material benefits that spawn from interactions between NC and humans. CES research is therefore pointing towards the detection of the interaction of two elements and not focusing on NC alone. This is in accordance with some conceptualisations critical of the cascade model. To overcome the limits of the above cited NC centric research on CES, two approaches appear in the literature. One led to the use of statistics on tourism and the design of qualitative data acquired through large scale surveys. This methodology still keeps researchers apart from communities that are simply interrogated but are not involved in the value identification process. There is a unidirectional flow of information. The second approach implies that value identification and even formation are generated when researchers and communities engage in participatory processes where a two-way communication informs research, and a deliberative approach supported by cartographic representations is taken. This methodology is more effective and very useful lessons have been learnt from its implementation. This section of CES focused research is at the forefront of transdisciplinarity with collaborations with landscape and social focused research. The most effective deliberative methods are the ones using maps. These two approaches are applied to large scale country-wide analytical works or to local\regional scale studies. The two approaches to CES detection are inversely related. Scale, precision, and locality are inversely related while the need to invest large amounts of time and funds is a challenge to both methodologies. The most effective deliberative methodologies are affected also by difficulties in identifying and selecting appropriately the population sample to engage. It has been proven that they tend to engage consistently with a limited range of population, characterised by older, retired, middle-upper class, educated white people. To overcome such limitations the use of CGI data has often been implemented in a mismatched way. The main commonalities between these implementations are: the use of CGI as secondary data, the interpretative approach of implicitly expressed cultural preferences and, the focus almost exclusively on photosharing projects. These examples created valuable precedents for our work but still, the full potential of CGI for CES research has been missed since CGI has been simply used as a source of data.

Chapter 3 delves into CGI starting from the distinctive characteristics of the overarching phenomenon of user generated content which is linked to the rise of, amongst others, crowdsourcing and citizen science. The first section of this chapter includes the description of the main components of CGI and the depiction of terminological issues, which involves the widely used term VGI and ontological clarifications. The design of a very simple two-dimensional classification is derived for all CGI projects. This in light of our use of CGI. The two dimensions are constituted by two dichotomies. The first distinguishes projects based on their spatial scope. The second distinguishes projects based on their collaboration pattern. CGI projects are the subject of several studies highlighting specific traits as well as limitations, when they are used as secondary data. The main critical points are from two different perspectives. Data quality and reliability are the main concern from an information perspective, while, from a participation point of view, biases in gender, social, spatial, and temporal terms are always detected. Studies on the use of CGI as a source of qualitative value are rising in numbers and quality. The large amount of information included in the data collections enabled analysis over several domains and using a wide range of technologies and methodologies.

The chapter then turns to scrutinise the studies concerning the three families of CGI projects that are used in the analytical sections. These are photo collections, citizen science environmental citizen science and especially biological recordings, and explicitly collaborative knowledge creation projects that we named *incremental* projects for brevity. Citizen science surveys when including digitised records from the past constitute the oldest example of CGI, while photo collections have been widely analysed as a source of secondary information. The use of citizen science surveys and incremental sources to derive social/spatial bounds is almost unprecedented.

Chapter 4 identifies research gaps, objectives, and methodologies. The gaps and issues related to the studies aiming at the detection of CES have previously been underlined. The use of census or protected area spatial units for CES detection has also been deemed inappropriate while recurrent. This also extends to all classifications and indicators who used census data and census units such as the rural-urban classification. The studies using CGI for the detection of social/spatial bonds have been affected by recurring limitations. CGI has been used as a source of data without checks on its fitness for the purpose. The biases have been often ignored as well as the social aspects that characterises CGI. Very often one only source of CGI has been used relying heavily and blindly on its biases. Few studies have analysed more than one CGI project.

This chapter defines the research issues identified here: the need to demonstrate that various CGI resources can contribute to CES research, the components of CES that can be elicited from CGI resources, and the identification of recommendations to operationalise the use of CGI for CES research. Chapter four also describes the chosen analytical methodology. Not only it has been derived from best practices, but also by reflecting on specific issues that arise when CGI resources are used. The analytical methodology therefore, considers the individual projects and their specific features. An extensive amount of work has been dedicated to the data preparation. It has been developed engaging with the CGI projects and communities, improving best practices, and considering methods developed in other disciplines. Amongst the preparation methods we present a procedure to identify urban spatial grids based on advanced studies on census modelling and territorial analysis. We also described when we needed to develop and implement coded procedures to extract, select, and remove spuriously or automatically created data. The analysis methodology then relies on spatial units that are not based in either census, land use or protection status. The evaluation of people interest for a spatial unit is assumed as the number of people contributing to information creation and refinement in the specific unit. The statistical relevance of such measurement is assessed through correlation and autocorrelation analysis. To account for the different sizes of projects a dimensionless index has been designed.

Chapter 5 introduces the study areas and the CGI resources that are the objective of the analytical work. South east Wales and Greater London have been chosen. The former was included to represent a wide range of environments from the coastline to coastal urban settlements and the linear settlements developed along valleys that connect the coastline to a National Park. This area also includes a wide range of areas listed for their scenic beauty. The Greater London area is a cosmopolitan metropolis where millions of inhabitants and millions of tourists share an urban environment characterised by the presence of iconic landmarks, large parks, green and blue areas both in the most densely populated inner area and in the less dense outer area. The second part of this chapter shows the 9 CGI resources that have been involved in the analysis. Their genesis, similarities and idiosyncratic characteristics are analysed also in light of the dichotomies identified in Chapter three. The first family of projects (Flickr, Panoramio, and Geograph) has technological arrangements that similarly support and promote the collection and online sharing of photos taken by members of the public. The second family of projects (GiGL, iSpot, iRecord) includes projects whose aim is the collection, diffusion and promotion of volunteer led biological recordings. We focused especially on volunteer collected information which is the totality of iSpot and iRecord, and a relevant part of GiGL that also digitizes past collections of environmental information providing us with a deeper temporal depth and a wide variety of volunteer arrangements. The third family of projects includes incremental or explicitly collaborative knowledge creation projects. Wikipedia is the most successful example of user generated content project. In Wikipedia, this work analysed articles containing a spatial tag. Wikimapia and OpenStreetMap are very popular collaborative mapping projects which engage people in the creation of geographic information or in the annotation of aerial maps accessible online.

Chapters 6-7-8 are the thesis analytical phase. The data preparation processes and the wide range of approaches are explained and the outcome of this phase is highlighted.

Chapter 6 is devoted to photosharing projects. It highlighted how relevant similarities between projects allow nonetheless to characterise them. While Flickr has a declination towards social events and gatherings, Geograph stresses accessibility, while Panoramio favours open views and scenic areas. At the end of the chapter a side projects briefly analyses the spatial distribution of pictures shared on Panoramio and Geograph in Greater London stressing how the spatial spread of pictures is different and the role played by blue areas.

Chapter 7 reports on milder correlations between the citizen science biological recordings projects. But areas with open views on blue and green areas in metropolitan London are consistently favoured. Interesting hotspots are also detected in low density areas and the different ways in which citizen scientists engage with nature are highlighted in this chapter.

Chapter 8 presents the most difficult case of data preparation since the widely open projects such as Wikipedia and OSM are characterised by very active communities where IT experts play a pivotal role in providing automated ways to enhance information.

This led to the need to remove both automated and unsupervised edits, but also erroneous edits inserted by badly designed editing applications. The study area is a subsection of the study area of chapter 6. Notwithstanding the hugely different IT medium and scope of the projects of these two chapters we found the same statistically significant spatial units. In the case of incremental sources, the collaboration patterns identified some specific areas where historical depth or contentious events attracted more contributors.

10.2 Main research results

The analysis has allowed us to identify CES in south Wales and Greater London. In **chapter 9** we discuss the main findings of this work. The analysis provided relevant outcomes for CES characterisation, for the contribute of CGI to CES research, and some comments for the development of CGI projects. For CES characterisation we found that environmental spaces where social cultural interactions are detected are, according to the most recent conceptualisations on the ES approach (Mace 2016, Costanza et al. 2017), places where interactions between NC and human, social, and built capital are detected. Benefits are not merely derived in a unidirectional cascade process from NC. But they are the spawn of NC and human, social, and built capital. Views needs to be accessible, social gatherings and visits must be organised and set in specifically designated places. The localisation of cultural heritage is very often an indication of a pre-existing valued natural setting or has induced the creation and maintenance of "natural" beauty in the surrounding areas. Societies and social activities that come with them shape the spatial preferences of people. Preferences are very often framed considering physical affordances.

Technology and the new ways in which people interact with space (Castells et al. 2007) are also influencing the interactions between humans and nature. In particular, rising numbers of casual recorders or dabblers associated with technologically enabled, volunteer led, biological recordings projects (Boakes et al. 2016), are inserting daily unusual *spottings* in the debate on nature and residency*. Nature on the doorstep* is a new, unfolding perspective on the relationship between human settlements and nature, that is nicely reflected in London by Daniel Raven-Ellison's initiative to create a Greater London National Park City (http://www.nationalparkcity.london/). This supports concepts such as conservation without borders (Palomo et al. 2014).

CES come in bundles with heritage values especially when the natural environment is the setting for iconic and symbolic buildings that contribute to the shaping of local identity and rootedness. In the development of CES research we can affirm that engaging with CGI projects can be a fruitful investment since CGI engages a range of participants usually not involved in deliberative projects. CGI data is produced by local residents and tourists[107] which are usually excluded from household-bound surveys. Combining mobile enabled recordings with digitised records research is not bound to only one set of technological affordances, or cultural and social background. CGI communities and data have embedded biases that are indications of cultural preferences. The engagement with CGI communities is functional to their recognition. CGI is also characterised by ranges of populations and communities belonging to the *mobile society*, where engagement with people and space is more spontaneous and less planned. For CGI projects we can reaffirm the need to create a more solid ground for their longevity (Haklay et al. 2014). However, to ensure the value of CGI projects is available for multiple uses as secondary data we need first and foremost a deeper understanding of the demographic profiles of people involved. Another debatable issue is the use of automated systems to enhance the perceived quality of data. There are different kinds of quality according to the different uses. This thesis has gained insights from biases, but for other uses the same biases might be evaluated as detrimental because they can diminish data quality. Therefore, we recommend that data should be preserved in the most possible untouched status so that users may then apply their own individual quality assessment and improvement procedures.

10.3 Further developments – the latent potential of CGI

Future developments and improvements to this work can be pursued in several directions by expanding the pool of CGI resources, complementing it with a qualitative analysis of CGI communities, and deepening the analytical methodology using other components of CGI projects. The use of the same methodology as other CGI sources which have been excluded in this work, is the simplest development. Several CGI projects have already been used to analyse socio/spatial relations in generic social media (Facebook) (Wilson et al. 2012), microblogging (Twitter) (Adnan et al. 2013) or reviews on leisure services (Foursquare, TripAdvisor) (Kisilevich et al. 2010b, Floris and Campagna 2014, Quercia and Saez 2014). For most of them our methodology to exclude urban areas will be very useful. We have also projects that track people's activities through GPS and collect their spatial behaviour while riding bikes, exercising in the outdoors using also fitness trackers who register physical parameters (Strava, Ride with GPS, Routeyou) (Baker and Verstockt 2017) some of them are supported by public bodies such as Ordnance Survey who collects leisure routes. Open air alternative reality games such as geocaching (Mendes et al. 2013, Cord et al. 2015) or the more advanced products of Niantic labs such as *Ingress* or the more advanced *Pokémon go* which uses augmented reality technologies contain also user generated elements in space that can lead to spatial preferences (Juhász and Hochmair 2017).

[107] Some studies tried to distinguish among them measuring the time frame of contribution to CGI projects

A more effective development for better balanced analysis between space and society might be in the deeper engagement with CGI communities and their demographic profiling. This could be done either through qualitative research methods engaging with CGI communities, or matching their spatial footprints with census data and eliciting sociodemographic knowledge from name/surname couples, when available. The traditional way to investigate the social aspects of volunteers would be to initiate an investigation using established social science research methods and engage directly with contributors' communities, and organisations coordinating the data collection efforts. An alternative analysis could be conducted on the demographic profiling of contributors to find patterns and trends related to variables such as gender, social class, and age in relation to their area of residence (known as geodemographic analysis). Similar geodemographic analysis in the USA has demonstrated the importance of analysing the target population of every form of social media, as stressed by Duggan and Brenner (2013) and Li et al. (2013). We have mentioned similar studies where Twitter and OSM have been used (Adnan and Longley 2015, Mashhadi et al. 2015). Because cultural and social differences strongly influence CES-related preferences (Wood et al. 2013, López-Santiago et al. 2014, UK National Ecosystem Assessment 2014), and engagement with open space and nature is ethnically bounded (Agyeman 1990), we would expect such analysis to yield useful results. When the name and surname of participants are available it is possible to use methods, and tools developed for geography, such as Onomap (http://www.onomap.org/) and the works of Mateos (2014), that can derive social and ethnographic profiles from name/surname couples. But these possibilities may be deterred by ethical considerations as suggested in the following section 10.4.

The third direction for development of this work is the expansion of the current set of analytical tools used. This work explores the potential of the lowest common denominator of the CGI information components as described in section 3.3. They are the position and unique identification of contributors. From a broader perspective, CGI projects are multifaceted sources of social and spatial information that can be very relevant to CES research and practice. But to overcome the limitation that considers CGI projects as just a source of data, we might consider analysing the internal functioning of the communities by mining information from mailing list registries and Wikipages, which include also lineage information.

But even focusing only on the information bundle provided by the organisations through APIs or data dumps, there are several components that have not been used in this work. Very often in CGI projects there is also a descriptive section where contributors describe the surrounding environment or add subjective comments. Websites and data dumps contain several social features such as interest groups and the facility to add "likes", comments, tags, and ratings. Contributors in some cases, can improve the information inserted by others. In this case, the collaboration pattern of the project is incremental. Therefore, several improvements to our study, as well as the identification of complementary streams of research, could flow from this abundant swathe of information.

The social component of CGI sources can be further analysed by considering spatial or spatiotemporal co-occurrences, not only on the same platform but also across platforms. It could involve also social network analysis (Cranshaw et al. 2010), when mutual interactions such as comments, “likes,” and ratings are part of the platform. The influence of social-networking tools on photosharing websites has already been analysed by Nov et al. (2010) while social network derived analytical methodologies have been applied to OSM (Keßler et al. 2011, Mooney and Corcoran 2012a, 2014, Mooney et al. 2013, Stein et al. 2015). A deeper analysis would improve the understanding of the behaviour of contributors in gamified crowdsourcing platforms such as Geograph. To facilitate the application of the set of analysis of this work, the automation of the data preparation phase could be implemented using data mining techniques used to detect carto-vandalism and improve CGI quality (Ballatore 2014, Senaratne et al. 2016). Some studies take advantage of the textual component included in comments and tags attached to pictures; studies of conceptualisations from crowdsourced information aim to do this. Other studies have derived ontologies and folksonomies (Garduño Freeman 2010) from Flickr tags, “likes”, and groups (Hollenstein and Purves 2010). The technological setting where data is described and linked semantically is one of the potential options to operationalise and expand our approach. The cultural links we establish with the environment have the potential of being unravelled through semantic links. Similar approaches have already been implemented for quality assessment and query expansion of CGI data semantically enabled and linked with linguistic resources (Gliozzo 2013, Ballatore 2016, Bucher et al. 2016, Lemmens et al. 2016).

10.4 Ethical implications

The use of crowdsourced data is problematic and the legal and ethical issue related to it have been investigated widely. As emphasised before, CGI and citizen science share most of this aspect and relevant studies for citizen science such as Scassa et al. (2015) widely applies to the CGI resources we have used in this work. To address more explicitly this problem Mooney et al. (2017) underscored how the spatial element of CGI adds a very powerful tool that is able to disclose people life in space and time. This can lead to the violation of privacy.

Privacy can involve human and non-human subjects. In this work we already faced three approaches to the protection of privacy of human and non-human subjects. Most of them have been implemented in the biodiversity recordings projects. The first approach is the one implemented specifically by iSpot and Panoramio. It consists in the anonymization of users. Panoramio and iSpot provided us just a unique identifier with no username or name/surname couples. The second approach consists in reducing the precision of the geographic reference. In all biodiversity recordings datasets rare and vulnerable species are recorded with a spatial precision that is kept low intentionally to avoid disturbances to the non-human subjects. The third approach is the one that iRecord and GiGL used. They provided us data through the signature of a specific agreement that stated the modalities of the use of their data. Such data licenses included also the prohibition to distribute the original data as it has been shared with us. This is a sensible approach considering that their data included identifiable information about contributors.

The review of the data protection act (United Kingdom. 2018) following the EU directive on General Data Protection Regulation (European Union 2016) states rigid procedures to store and use personal data electronically. Therefore, privacy can be under threat if CGI data holders do not comply with the enhanced protection guaranteed by the new regulations. From another perspective we suggested how our approach can be enhanced if we implement a more detailed socio-demographic analysis of the contributors. The provision of the mentioned laws and regulations allow the use of personal data for research purposes. This means that while biodiversity recordings data and marginally incremental sources can be already considered as compliant, our intended use makes also the other datasets compliant to the provisions of the new regulations. A strict implementation of the regulations is shaped similarly to the first approach we describe. The storage of personal information and their release through APIs from platforms can be more problematic. Since there must be no connection between spatial and personal data the use of usernames is a favourable factor. But very often people like to use name/surname couples. The ways in which companies will shape their services accessible through APIs following the introduction of the new regulations is a potential challenge to our approach.

10.5 Reflections on the ES conceptualisation

From a wider perspective this work contributes to the advancement of the ES field in a twofold way. It provides a critical perspective on the shaping of its conceptualisations and it leads to the development of wider ways to engage and enhance the discourse around it.

10.5.1 Conceptual frameworks

The analysis has shown how the ES perspective misses a real unifying conceptualisation that embraces holistically the reality.

The two most recent reviews and criticism on the ES perspective propose the revision or substitution of the overarching ontological order (Costanza et al. 2017, Díaz et al. 2018).

Interestingly they agree on a broader role assigned to CES or non-material benefits. They are not merely one category, but they are underlying most of other services. But still these conceptualisations miss the opportunity to more precisely target a deeper understanding of the unfolding of the intricacies between human, natural built, and social cultural aggregations.

There is missing an underlying temporal dimension. This is witnessed by the large amount of natural/heritage areas in our analysis. Human decisions and preferences are often the result of a combination of components some of which are related to long-lasting accrue of a heritage which has several dimensions. The current setting of reality is just the timely upshot of an evolution where the previous heritage is continuously evaluated and subjected to inertia, more or less invasive development, maintenance, or restoration. But ignoring the previous upshots and balances between social, built, and natural environment equates to ignoring the value that all of its components played in the past, and the latent potential they have for development.

Therefore, this drives not only on a lacking analytical perspective but misses some opportunities.

10.5.2 Including time

We suggest the development or the adoption of a conceptual framework which includes the historical unfolding of the relationship between social, built, and natural environment. One conceptualisation that can sustain it is based on the concept of territory and the related territoriality. We refer to the second meaning as reported in the Oxford geography dictionary for territory a really interesting perspective is the one developed by the French-swiss geographer Claude Raffestin (2012). Very similar conceptualisations are also developed around the concept of place which in some conceptualisations has a social, ecological, and historical depth.

The inclusion of the temporal depth is also strongly advocated when we see how the connection with nature is related to the sense of place and identity. In some disciplinary context this is called *genius loci*, referring to the roman concept of spirits of the place.

10.5.3 Engaging audiences

Local stakeholders can have a more engaging experience when preliminary studies can delve on a deeper perspective explaining the unfolding of the relation and the mutual construction between the social, built, and natural worlds. Our use of data created by communities shed a different light on people's attitude towards space and it can have implications for several actors.

Engaging visualisations can generate public debate on the use of resources. We already experimented some in the form of maps. The maps used in this work are intended as a tool to support the analytical phase and bear limited value for engagement. But we already designed some maps that literally enhance the value of the data used and stress some of the main outcomes of our work. We called them the spotlight maps. Their simplicity and design, together with the absence of cartographic must-be such as scale factors is intentional. They are intended for people who know the area, so additional references are unnecessary. Those maps have not been drawn but they are derived from our data. They have been derived using the values of contributors for every grid square. We created a surface which connects every centre point of the grids. Centre points have associated as value the number of contributors for each grid unit. The resulting surface has been represented using a graded scale which turns from white for the highest values to black to the lowest values. We than set a high value of transparency for this layer and overlaid it to expectation surfaces. This way we obtain the visual effect of spotlight aimed at specific locations.
We obtained in London (Figure 10-1, below) the map showing how not all green and blue areas host biodiversity recording activities proportionally to their extent. This can lead to reflection on how several factors contribute to the attractiveness for citizen science of the green and blue areas. The coloured circles overlaid over the map tend to stress the main idea behind the map design.

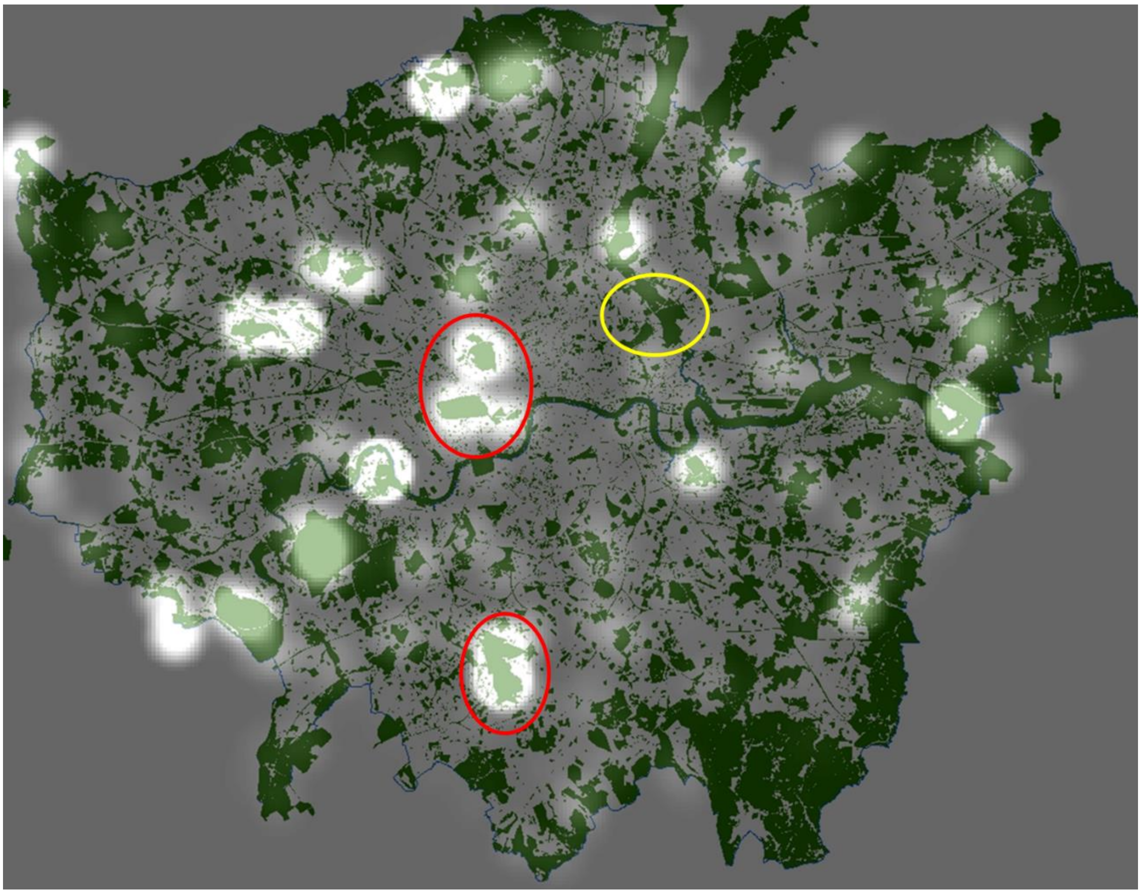

Figure 10-1:Spotlight map of citizen science and green/blue areas in Greater London

In Wales (Figure 10-2) we drawn a spotlight map showing the disconnection between urban areas and places where people more likely share images on photographic collections online. As already explained, we know from literature and we also have found in our study that usually urban areas play a major role in photographic collections, but this map, enhanced with the use of the coloured circles shows how you have natural areas that attract contribution to photosharing platforms but also urban areas who do not appear under the spotlight. This can suggest several approaches to analyse the factors that contribute to the success of natural areas or to the unattractiveness of urban ones. This to enhance urban and natural environments to deliver more wellbeing.

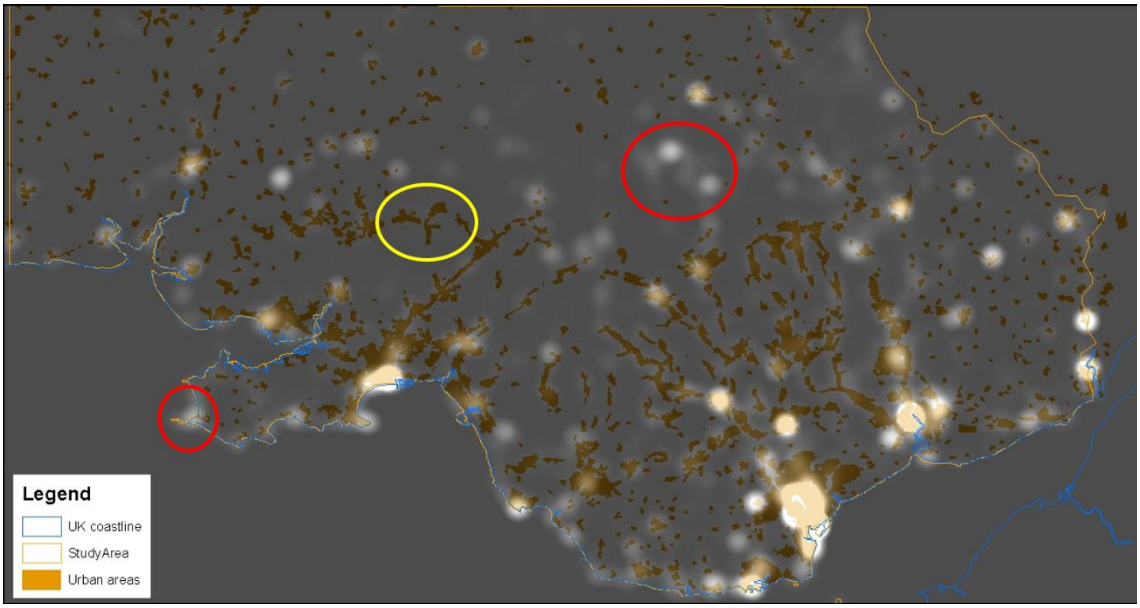

Figure 10-2: Spotlight map of photosharing and urban areas in South Wales

Both those maps do not follow standard cartographic rules since they are designed to engage rather than to inform. In the version shown in fact we reported the circles we used to stress how the density of contributors is not matching the expectation surface. In London you can notice how several green areas are not highlighted such as the Hackney marshes and adjacent Queen Elisabeth Olympic Park, while other areas got an intensity of citizen science contribution that is covering areas way larger than parks alone.

On a wider area perspective, planners can benefit from our approach. They can see which areas are "underperforming" in terms of the delivery of engagement, flagging out the missed opportunities and underused resources. From another perspective once they embrace the historical depth of an enhanced conceptualisation they can see how to reignite local based economic development rediscovering and redesigning typical local productions.

Similarly, resource managers can understand which activities, access policies and supporting infrastructures can enhance visitation and revenue from non-consumptive use of nature and cultural resources towards a really sustainable development which is able to gain economic development which is able to combine the preservation of nature, culture and is able to enhance the role of communities and social groups.

References

Abson, D. J., H. von Wehrden, S. Baumgärtner, J. Fischer, J. Hanspach, W. Härdtle, H. Heinrichs, A. M. Klein, D. J. Lang, P. Martens, and D. Walmsley. 2014. Ecosystem services as a boundary object for sustainability. Ecological Economics 103:29–37.

Adams, B., and G. McKenzie. 2013. Inferring Thematic Places from Spatially Referenced Natural Language Descriptions. Pages 201–221*in* D. Sui, S. Elwood, and M. Goodchild, editors.Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice. Springer Netherlands, Dordrecht.

Adnan, M., G. Lansley, and P. A. Longley. 2013. A geodemographic analysis of the ethnicity and identity of Twitter users in Greater London. Proceedings of the 21st Conference on GIS Research UK:1–6.

Adnan, M., and P. A. Longley. 2013. Analysis of Twitter Usage in London, Paris, and New York City. Lueven.

Adnan, M., and P. A. Longley. 2015. Geo-temporal Twitter demographics. International Journal of Geographical Information Science:1–21.

Agyeman, J. 1990. Black People in a White Landscape: Social and Environmental Justice. Built Environment 16:232–236.

Ambrose-Oji, B., and T. F. Pagella. 2012. Spatial Analysis and Prioritisation of Cultural Ecosystem Services : A Review of Methods. Research Report, Forest Research. Alice Holt Lodge Farnham, Surrey.

Amitay, E., N. Har'El, R. Sivan, and A. Soffer. 2004. Web-a-Where: Geotagging Web Content. Pages 273–280SIGIR '04: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval.

Anderson, C. 2006. The Long Tail. Wired:1–8.

Anselin, L. 1995. Local indicators of spatial association — LISA. Geographical Analysis 27:93–115.

Antoniou, V. 2016. Volunteered Geographic Information measuring quality, understanding the value. GEOmedia 20.

Antoniou, V., J. Morley, and M. Haklay. 2010. Web 2.0 Geotagged Photos: assessing the spatial dimension of the phenomenon. Geomatica 64:99–110.

Arsanjani, J. J., and M. Bakillah. 2014. Understanding the potential relationship between the socio-economic variables and contributions to OpenStreetMap. International Journal of Digital Earth 8:1–16.

Arsanjani, J. J., C. Barron, M. Bakillah, and M. Helbich. 2013. Assessing the Quality of

OpenStreetMap Contributors together with their Contributions. 16th AGILE International Conference on Geographic Information Science.:14.-17.

Arsanjani, J. J., P. Mooney, M. Helbich, and A. Zipf. 2015a. An Exploration of Future Patterns of the Contributions to OpenStreetMap and Development of a Contribution Index. Transactions in GIS 19:896–914.

Arsanjani, J. J., P. Mooney, A. Zipf, and A. Schauss. 2015b. Quality Assessment of the Contributed Land Use Information from OpenStreetMap Versus Authoritative Datasets. Pages 37–58*in* J. J. Arsanjani, P. Mooney, A. Zipf, and M. Helbich, editors.OpenStreetMap in GIScience. Notes in G. Springer, Heidelberg, New York, Dordrecht, London.

Arsanjani, J. J., A. Zipf, P. Mooney, and M. Helbich, editors. 2015c. OpenStreetMap in GIScience. Lecture No. Springer, Heidelberg, New York, Dordrecht, London.

Auer, S., C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives. 2008. Dbpedia: A nucleus for a web of open data. The Semantic Web:722–735.

Badura, T., I. J. Bateman, M. Agarwala, and A. Binner. 2016. Valuing preferences for ecosystem related goods and services. Page *in* R. Haines-Young, M. Potschin, R. Fish, and K. Turner, editors. Routledge Handbook of Ecosystem Services. Routledge, London and New York.

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

Bagstad, K. J., D. J. Semmens, S. Waage, and R. Winthrop. 2013b. A comparative assessment of decision-support tools for ecosystem services quantification and valuation. Ecosystem Services 5:27–39.

Baker, K., and S. Verstockt. 2017. Cultural heritage routing: A recreational navigation-based approach in exploring cultural heritage. Journal on Computing and Cultural Heritage 10:1–20.

Ballatore, A. 2014. Defacing the map : Cartographic vandalism in the digital commons. The Cartographic Journal 51:214–224.

Ballatore, A. 2016. Semantic Challenges for Volunteered Geographic Information. Pages 87–95*in* C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors.European Handbook of Crowdsourced Geographic Information. Ubiquity Press License: CC-BY 4.0., London, UK.

Ballatore, A., M. Bertolotto, and D. C. Wilson. 2012. Geographic Knowledge Extraction and Semantic Similarity in OpenStreetMap. Knowledge and Information Systems 1:1–21.

Barnaud, C., and M. Antona. 2014. Deconstructing ecosystem services: Uncertainties and controversies around a socially constructed concept. Geoforum 56:113–123.

Barve, V. 2014. Discovering and developing primary biodiversity data from social networking sites: A novel approach. Ecological Informatics 24:194–199.

Bateman, I. J., D. J. Abson, N. Beaumont, A. Darnell, C. Fezzi, N. Hanleys, A. Kontoleon, D. Maddison, P. Morling, J. Morris, S. Mourato, U. Pascual, G. Perino, A. Sen, D. Tinch, K. Turner, G. Valatin, B. Andrews, V. Asara, T. Askew, U. Aslam, G. Atkinson, N. Beharry-Borg, K. Bolt, M. Cole, M. Collins, E. Comerford, E. Coombes, A. Crowe, S. Dugdale, H. Dunn, J. Foden, S. Gibbons, R. H. Haines-Young, C. Hattam, M. Hulme, M. Ishwaran, A. a. Lovett, T. Luisetti, G. MacKerron, S. Mangi, D. Moran, P. Munday, J. Paterson, G. Resende, G. Siriwardena, J. Skea, D. van Soest, and M. Termansen. 2011. Economic values from ecosystems. Pages 1067–1152*in* UK National Ecosystem Assessment, editor.The UK National Ecosystem Assessment Technical Report. UNEP-WCMC, Cambridge UK.

Bateman, I. J., B. H. Day, S. Georgiou, and I. Lake. 2006. The aggregation of environmental benefit values: Welfare measures, distance decay and total WTP. Ecological Economics 60:450–460.

Bateman, I. J., B. Day, M. Agarwala, P. Bacon, T. Badura, A. Binner, A. De-Gol, B. Ditchburn, S. Dugdale, B. Emmett, S. Ferrini, C. Carlo Fezzi, A. Harwood, J. Hillier, K. Hiscock, P. Snowdon, G. Sunnenberg, S. Vetter, and S. Vinjili. 2014. UK National Ecosystem Assessment Follow-on. Work Package Report 3: Economic value of ecosystem services:1–246

Bateman, I. J., A. R. Harwood, G. M. Mace, R. T. Watson, D. J. Abson, B. Andrews, A. Binner, A. Crowe, B. H. Day, S. Dugdale, C. Fezzi, J. Foden, D. Hadley, R. H. Haines-Young, M. Hulme, A. Kontoleon, A. A. Lovett, P. Munday, U. Pascual, J. Paterson, G. Perino, A. Sen, G. Siriwardena, D. van Soest, and M. Termansen. 2013a. Bringing Ecosystem Services into Economic Decision-Making: Land Use in the United Kingdom. Science 341:45–50.

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

Bennett, E. M., W. Cramer, A. Begossi, G. Cundill, S. Díaz, B. N. Egoh, I. R. Geijzendorffer, C. B. Krug, S. Lavorel, E. Lazos, L. Lebel, B. Martín-López, P. Meyfroidt, H. A. Mooney, J. L. Nel, U. Pascual, K. Payet, N. P. Harguindeguy, G. D. Peterson, A. H. Prieur-Richard, B. Reyers, P. Roebeling, R. Seppelt, M. Solan, P. Tschakert, T. Tscharntke, B. L. Turner, P. H. Verburg, E. F. Viglizzo, P. C. L. White, and G. Woodward. 2015. Linking biodiversity, ecosystem services, and human well-being: three challenges for designing research for sustainability. Current Opinion in Environmental Sustainability 14:76–85.

Bieling, C., and T. Plieninger. 2012. Recording Manifestations of Cultural Ecosystem Services in the Landscape. Landscape Research 38:649–667.

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

Bishr, M., and L. Mantelas. 2008. A trust and reputation model for filtering and classifying knowledge about urban growth. GeoJournal 72:229–237.

Bittner, C. 2017. Diversity in volunteered geographic information: comparing OpenStreetMap and Wikimapia in Jerusalem. GeoJournal 82:887–906.

Bizer, C., J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann. 2009. DBpedia-A crystallization point for the Web of Data. Web Semantics: Science, Services and Agents on the World Wide Web.

Blumenstock, J. E. 2008. Size matters: word count as a measure of quality on wikipedia. Pages 1095–1096Proceedings of the 17th ACM International Conference on the World Wide Web (WWW ’08).

Boakes, E. H., G. Gliozzo, V. Seymour, M. Harvey, C. Smith, D. B. Roy, and M. Haklay. 2016. Patterns of contribution to citizen science biodiversity projects increase understanding of volunteers’ recording behaviour. Scientific Reports 6:33051.

Bonney, R., H. Ballard, R. C. Jordan, E. McCallie, T. Phillips, J. Shirk, and C. C. Wilderman. 2009a. Public Participation in Scientific Research: Defining the Field and Assessing Its Potential for Informal Science Education. Page Science Education.

Bonney, R., C. B. Cooper, J. Dickinson, S. Kelling, T. Phillips, K. V. Rosenberg, and J. Shirk. 2009b. Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. BioScience 59:977–984.

Bonter, D. N., and C. B. Cooper. 2012. Data validation in citizen science : a case study from Project FeederWatch. Frontiers in Ecology and the Environment:305–307.

Boyd, J., and S. Banzhaf. 2007. What are ecosystem services? The need for standardized environmental accounting units. Ecological Economics 63:616–626.

Britannica Encyclopedia. 2006. Fatally Flawed. Encyclopædia Britannica, Inc.

Brown, G., and N. Fagerholm. 2015. Empirical PPGIS/PGIS mapping of ecosystem services: A review and evaluation. Ecosystem Services 13:119–133.

Bruns, A. 2007. Produsage : Towards a Broader Framework for User-Led Content Creation. Pages 99–106C&C ’07 Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition. ACM, New York, NY, USA,.

Bucher, B., F. Gilles, C. Metral, and R. Lemmens. 2016. Enhancing the management of quality of VGI: contributions from context and task modelling. Pages 131–142*in* C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors.European Handbook of Crowdsourced Geographic Information. Ubiquity Press License: CC-BY 4.0., London, UK.

Burgess, N. D., S. Darrah, S. Knight, and F. S. Danks. 2016. Approaches To Mapping Ecosystem Services. UNEP World Conservation Monitoring Centre (UNEP-WCMC), Cambridge CB3 0DL, UK.

Calvet-Mir, L., H. March, D. Corbacho-Monné, E. Gómez-Baggethun, and V. Reyes-García. 2016. Home Garden Ecosystem Services Valuation through a Gender Lens: A Case Study in the Catalan Pyrenees. Sustainability 8:718.

Campagna, M. 2016. Social Media Geographic Information : Why social is special when it goes spatial ? Pages 45–54*in* C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors.European Handbook of Crowdsourced Geographic Information. Ubiquity Press License: CC-BY 4.0, London, UK.

Camps-Calvet, M., J. Langemeyer, L. Calvet-Mir, and E. Gómez-Baggethun. 2015. Ecosystem services provided by urban gardens in Barcelona, Spain: Insights for policy and planning. Environmental Science and Policy 62:14–23.

Capineri, C. 2016. The Nature of Volunteered Geographic Information. Pages 15–33*in* C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors.European Handbook of Crowdsourced Geographic Information. Ubiquity Press License: CC-BY 4.0., London, UK.

Capineri, C., M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors. 2016. European Handbook of Crowdsourced Geographic Information. Ubiquity Press License: CC-BY 4.0, London, UK.

Carey, P. D., S. Wallis, P. M. Chamberlain, A. Cooper, B. A. Emmett, L. C. Maskell, T. McCann, J. Murphy, L. R. Norton, B. Reynolds, W. A. Scott, I. C. Simpson, S. M. Smart, and J. . Ullyett. 2009. Countryside Survey Methodology. Page Countryside Survey: UK results from 2007. CEH Centre for Ecology & Hydrology.

Carey, P. D., S. M. Wallis, B. Emmett, L. C. Maskell, J. Murphy, L. R. Norton, I. C. Simpson, and S. M. Smart. 2008. Countryside Survey: UK Headline Messages from 2007. Oxford, UK.

Carpenter, S. R., H. a Mooney, J. Agard, D. Capistrano, R. S. Defries, S. Díaz, T. Dietz, A. K. Duraiappah, A. Oteng-Yeboah, H. M. Pereira, C. Perrings, W. V Reid, J. Sarukhan, R. J. Scholes, and A. Whyte. 2009. Science for managing ecosystem services: Beyond the Millennium Ecosystem Assessment. Proceedings of the

National Academy of Sciences of the United States of America 106:1305–1312.

Casalegno, S., R. Inger, C. DeSilvey, and K. J. Gaston. 2013. Spatial Covariance between Aesthetic Value & Other Ecosystem Services. PLoS ONE 8:e68437.

Castells, M. 2010. The Rise of the Network Society. second. Blackwell, Cambridge, MA; Oxford, UK.

Castells, M., M. Fernandez-Ardevol, J. L. Qiu, and A. Sey. 2004. Mobile Communication and Society. A cross-cultural analysis of available evidence on the social uses of wireless communication technology. Pages 0–392*in* M. Castells, M. Fernandez-Ardevol, J. L. Qiu, and A. Sey, editors.International Workshop on Wireless Communication Policies and Prospects A Global Perspective. Annenberg Research Network on International Communication., Anneberg.

Chan, K. M. A., J. Goldstein, T. Satterfi, N. Hannahs, K. Kikiloi, R. Naidoo, N. Vadeboncoeur, and U. Woodside. 2011. Cultural services and non-use values. Pages 206–228*in* P. Kareiva, H. Tallis, T. H. Ricketts, G. C. Daily, and S. Polasky, editors.Natural Capital: Theory and Practice of Mapping Ecosystem Services. Oxford University Press, Oxford, UK.

Chan, K. M. A., and T. Satterfield. 2016. Managing Cultural Ecosystem Services for Sustainability. Pages 343–358*in* M. B. Potschin, R. Haines-Young, M. Potschin, R. Fish, and K. Turner, editors.Routledge Handbook of Ecosystem Services. RoutledgE, London and New York.

Chesnokova, O., M. Nowak, and R. S. Purves. 2017. A crowdsourced model of landscape preference. 13th International Conference on Spatial Information Theory (COSIT 2017):1–13.

Chow, T. E. 2013. “We know who you are and we know where you live”: A research agenda for web demographics. Pages 265–285*in* D. Sui, S. Elwood, and M. Goodchild, editors.Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice. Dordrecht.

Church, A., J. Burgess, N. Ravenscroft, W. Bird, K. Blackstock, E. Brady, M. Crang, R. Fish, P. Gruffudd, S. Mourato, J. Pretty, D. Tolia-kelly, K. Turner, and M. Winter. 2011. Cultural Services. Pages 633–692UK National Ecosystem Assessment: Technical Report. UNEP-WCMC, Cambridge.

Church, A., R. Fish, S. Mourato, J. Tratalos, L. Stapleton, C. Willis, P. Coates, S. Gibbons, C. Leyshon, M. B. Potschin, N. Ravenscroft, R. Sanchis-Guarner, M. Winter, and J. O. Kenter. 2014. UK National Ecosystem Assessment Follow-on. Work Package Report 5: Cultural ecosystem services and indicators. LWEC, UK.

CICES V4.3. 2012. Common International Classification of Ecosystem Services (CICES , Version 4 . 3). Report to the European Environment Agency:1–17.

Clarke, L., and J. Agyeman. 2011. Shifting the Balance in Environmental Governance:

Ethnicity, Environmental Citizenship and Discourses of Responsibility. Antipode 43:1773–1800.

Coates, P., E. Brady, A. Church, B. Cowell, S. Daniels, C. DeSilvey, R. Fish, V. Holyoak, D. Horrell, S. Mackey, R. Pite, A. Stibbe, R. Waters, M. Dudley, and J. O. Kenter. 2014. Arts & Humanities Perspectives on Cultural Ecosystem Services (CES) - Arts & Humanities Working Group (AHWG): Final report.

Cockings, S., A. Harfoot, D. Martin, and D. Hornby. 2011. Maintaining existing zoning systems using automated zone-design techniques: methods for creating the 2011 Census output geographies for England and Wales. Environment and Planning A 43:2399–2418.

Cohn, J. P. . 2008. Citizen Science : Can Volunteers Do Real Research ? BioScience 58:192–197.

Comber, A., P. Mooney, R. S. Purves, D. Rocchini, and A. Walz. 2015. Comparing national differences in what people perceive to be there: mapping variations in crowd sourced land cover. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-3/W3:71–75.

Comber, A., P. Mooney, R. S. Purves, D. Rocchini, and A. Walz. 2016. Crowdsourcing: It Matters Who the Crowd Are. The Impacts of between Group Variations in Recording Land Cover. Plos One 11:e0158329.

Cooper, N., E. Brady, H. Steen, and R. Bryce. 2016. Aesthetic and spiritual values of ecosystems: recognising the ontological and axiological plurality of cultural ecosystem “services.” Ecosystem Services 21:218–229.

Cord, A. F., F. Roeßiger, and N. Schwarz. 2015. Geocaching data as an indicator for recreational ecosystem services in urban areas: Exploring spatial gradients, preferences and motivations. Landscape and Urban Planning 144:151–162.

Costanza, R. 2008. Ecosystem services : Multiple classification systems are needed. Biological Conservation:350–352.

Costanza, R. 2016. Ecosystem Services in theory and practice. Pages 15–25*in* M. Potschin, R. Haines-Young, R. Fish, and R. K. Turner, editors.Routledge Handbook of Ecosystem Services. Routledge, London and New York.

Costanza, R., R. Arge, R. De Groot, S. Farberk, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R. V. O. Neill, J. Paruelo, R. G. Raskin, and P. Suttonkk. 1997. The value of the world’ s ecosystem services and natural capital 387:253–260.

Costanza, R., and H. E. Daly. 1992. Natural Capital and Sustainable Development. Conservation biology 6:37–46.

Costanza, R., L. Fioramonti, and I. Kubiszewski. 2016. The UN Sustainable Development Goals and the dynamics of well-being. Frontiers in Ecology and the

Environment 14:59.

Costanza, R., R. de Groot, L. Braat, I. Kubiszewski, L. Fioramonti, P. Sutton, S. Farber, and M. Grasso. 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.

Costello, M. J. 2009. Motivating Online Publication of Data. BioScience 59:418–427.

Council of Europe. 2000. European Landscape Convention. Report and Convention Florence ETS No. 17:8.

Courter, J. R., R. J. Johnson, C. M. Stuyck, B. a Lang, and E. W. Kaiser. 2012. Weekend bias in Citizen Science data reporting: implications for phenology studies. International journal of biometeorology.

Cowling, R. M., B. Egoh, A. T. Knight, P. J. O'Farrell, B. Reyers, M. Rouget'll, D. J. Roux, A. Welz, and A. Wilhelm-Rechman. 2008. An operational model for mainstreaming ecosystem services for implementation. Proceedings of the National Academy of Sciences of the United States of America 105:9483–9488.

Cranshaw, J., E. Toch, J. Hong, A. Kittur, and N. Sadeh. 2010. Bridging the Gap Between Physical Location and Online Social Networks. Human Factors:119–128.

Daily, G. C. 1997. Introduction: What are ecosystem services? Pages 1–10*in* G. C. Daily, editor.Nature's Services: Societal Dependence on Natural Ecosystems. Island Press, Washington, DC.

Daily, G. C., S. Alexander, P. R. Ehrlich, L. Goulder, J. Lubchenco, P. A. Matson, H. A. Mooney, S. Postel, S. H. Schneider, D. Tilman, and G. M. Woodwell. 1997. Ecosystem Services: Benefits Supplied to Human Societies by Natural Ecosystems. Issues in Ecology.

Dales, N., N. Brown, and J. Lusardi. 2014. Assessing the potential for mapping ecosystem services in England based on existing habitats.

Daw, T., K. Brown, S. Rosendo, and R. Pomeroy. 2011. Applying the ecosystem services concept to poverty alleviation: the need to disaggregate human well-being. Environmental Conservation 38:370–379.

Day, B., and G. Smith. 2016. Outdoor Recreation Valuation (ORVal) Data Set Construction. Exeter.

Demeritt, D. 2002. What is the "social construction of nature"? A typology and sympathetic critique. Progress in Human Geography 26:767–790.

Department for Environment Food and Rural Affairs. 2013. The 2011 Rural-Urban Classification For Small Area Geographies : A User Guide and Frequently Asked Questions (v1 . 0):1–22.

Derungs, C., and R. S. Purves. 2016. Characterising Landscape Variation through Spatial Folksonomies. Applied Geography 75:1–24.

Díaz, S., U. Pascual, M. Stenseke, B. Martín-López, R. T. Watson, Z. Molnár, R. Hill, K.

M. A. Chan, I. A. Baste, K. A. Brauman, S. Polasky, A. Church, M. Lonsdale, A. Larigauderie, P. W. Leadley, A. P. E. van Oudenhoven, F. van der Plaat, M. Schröter, S. Lavorel, Y. Aumeeruddy-Thomas, E. Bukvareva, K. Davies, S. Demissew, G. Erpul, P. Failler, C. A. Guerra, C. L. Hewitt, H. Keune, S. Lindley, and Y. Shirayama. 2018. Assessing nature's contributions to people. Science 359:270 LP-272.

Diebold, F. X. 2012. A Personal Perspective on the Origin (s) and Development of " Big Data ": The Phenomenon , the Term , and the Discipline. Social Science Research network.

Dijkstra, L., and H. Poelman. 2014. A harmonised definition of cities and rural areas: the new degree of urbanisation. Brusselles, Belgium.

Duggan, M., and J. Brenner. 2013. The Demographics of Social Media Users — 2012. Washington, DC.

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

ECSA. 2015. Ten principles of citizen science. http://ecsa.citizen-science.net/sites/default/files/ecsa_ten_principles_of_citizen_science.pdf.

Eitzel, M. V, J. L. Cappadonna, C. Santos-Lang, R. E. Duerr, A. Virapongse, S. E. West, C. C. M. Kyba, A. Bowser, C. B. Cooper, A. Sforzi, A. N. Metcalfe, E. S. Harris, M. Thiel, M. Haklay, L. Ponciano, J. Roche, L. Ceccaroni, F. M. Shilling, D. Dörler, F. Heigl, T. Kiessling, B. Y. Davis, and Q. Jiang. 2017a. Citizen Science Terminology Matters: Exploring Key Terms - CORRECTION. Citizen Science: Theory and Practice 2:1.

Eitzel, M. V, J. L. Cappadonna, C. Santos-Lang, R. E. Duerr, A. Virapongse, S. E. West, C. C. M. Kyba, A. Bowser, C. B. Cooper, A. Sforzi, A. N. Metcalfe, E. S. Harris, M. Thiel, M. Haklay, L. Ponciano, J. Roche, L. Ceccaroni, F. M. Shilling, D. Dörler, F. Heigl, T. Kiessling, B. Y. Davis, and Q. Jiang. 2017b. Citizen Science Terminology Matters: Exploring Key Terms. Citizen Science: Theory and Practice 2:1.

Elwood, S., M. F. Goodchild, and D. Sui. 2013. Prospects for VGI research and the emerging fourth paradigm. Pages 361–375*in* S. Elwood, M. F. Goodchild, and D. Sui, editors.Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice. Springer Netherlands, Dordrecht.

EPA, U. S. 2015. National Ecosystem Services Classification System (NESCS): Framework Design and Policy Application. EPA-800-R-15-002. United States Environmental Protection Agency.

Estima, J., and M. Painho. 2015. Investigating the Potential of OpenStreetMap for Land Use/Land Cover Production: A Case Study for Continental Portugal. Pages 273–

294*in* J. J. Arsanjani, P. Mooney, A. Zipf, and M. Helbich, editors.OpenStreetMap in GIScience. Lecture No. Springer, Heidelberg, New York, Dordrecht, London.

European Communities. 2008. The economics of ecosystems & biodiversity.

European Union. 2016. Regulation 2016/679 of the European parliament and the Council of the European Union. Official Journal of the European Communities 2014:1–88.

Fagerholm, N., N. Käyhkö, F. Ndumbaro, and M. Khamis. 2012. Community stakeholders' knowledge in landscape assessments - Mapping indicators for landscape services. Ecological Indicators 18:421–433.

Fischer, A., and A. Eastwood. 2016. Coproduction of ecosystem services as human-nature interactions-An analytical framework. Land Use Policy 52:41–50.

Fischer, F. 2012. VGI as big data. A new but delicate geographic data-source. GeoInformatics April/May:46–47.

Fish, R. 2011. Environmental decision making and an ecosystems approach: Some challenges from the perspective of social science. Progress in Physical Geography 35:671–680.

Fish, R., J. Burgess, A. Footitt, and K. Turner. 2011. Participatory and deliberative techniques to support the monetary and non- monetary valuation of ecosystem services : An introductory guide. Defra Project Code: NR0124.

Fish, R., and A. Church. 2014. Cultural ecosystem services: stretching out the concept. Environmental Scientist 23:31–36.

Fish, R., A. Church, C. Willis, M. Winter, J. A. Tratalos, R. H. Haines-Young, and M. B. Potschin. 2016a. Making space for cultural ecosystem services: Insights from a study of the UK nature improvement initiative. Ecosystem Services 21:329–343.

Fish, R., A. Church, and M. Winter. 2016b. Conceptualising cultural ecosystem services: a novel framework for research and critical engagement. Ecosystem Services 21:208–217.

Fish, R., M. B. Potschin, R. K. Turner, and R. H. Haines-Young. 2016c. Ecosystem services: Never waste the opportunity offered by a good crisis. Pages 1689–1699*in* M. Potschin, R. Haines-Young, R. Fish, and K. Turner, editors.Routledge Handbook of Ecosystem Services. Routledge, London and New York.

Fish, R., and E. Saratsi. 2015. Naturally Speaking... A Public Dialogue on the UK National Ecosystem Assessment. CRPR, University of Exeter, Exeter.

Fish, R., E. Saratsi, M. Reed, and H. Keune. 2016d. Stakeholder participation in ecosystem services decision-making. Pages 256–270*in* M. B. Potschin, R. Haines-Young, R. Fish, and R. K. Turner, editors.Routledge Handbook of Ecosystem Services. Routledge. Routledge, London and New York.

Fisher, B., K. Turner, M. Zylstra, R. Brouwer, R. de Groot, S. Farber, P. Ferraro, R. Green, D. Hadley, J. Harlow, P. Jefferiss, C. Kirkby, P. Morling, S. Mowatt, R.

Naidoo, J. Paavola, B. Strassburg, D. Yu, and A. Balmford. 2008. Ecosystem services and economic theory: integration for policy-relevant research. Ecological applications : a publication of the Ecological Society of America 18:2050–67.

Fisher, B., R. K. Turner, N. D. Burgess, R. D. Swetnam, J. Green, R. E. Green, G. Kajembe, K. Kulindwa, S. L. Lewis, R. Marchant, a. R. Marshall, S. Madoffe, P. K. T. Munishi, S. Morse-Jones, S. Mwakalila, J. Paavola, R. Naidoo, T. Ricketts, M. Rouget, S. Willcock, S. White, and A. Balmford. 2011. Measuring, modeling and mapping ecosystem services in the Eastern Arc Mountains of Tanzania. Progress in Physical Geography 35:595–611.

Fisher, B., R. Turner, and P. Morling. 2009. Defining and classifying ecosystem services for decision making. Ecological Economics 68:643–653.

Floris, R., and M. Campagna. 2014. Social Media Geographic Information in Tourism Planning. Pages 417–430*in* R. Papa, A. Nuzzolo, E. Calderon, L. Bertolini, R. Fistola, A. Galderisi, C. Gargiulo, and G. Mazzeo, editors.Smart city:planning for energy, transportation and sustainability of the urban system.

Frank, S., C. Fürst, L. Koschke, and F. Makeschin. 2012. A contribution towards a transfer of the ecosystem service concept to landscape planning using landscape metrics. Ecological Indicators 21:30–38.

Frank, S., and U. Walz. 2017. Modelling Ecosystem Services. Pages 83–88*in* B. Burkhard and J. Maes, editors.Mapping Ecosystem Services. Pensoft Publishers.

Fritz, S., I. McCallum, C. Schill, C. Perger, R. Grillmayer, F. Achard, F. Kraxner, and M. Obersteiner. 2009. Geo-Wiki.Org: The Use of Crowdsourcing to Improve Global Land Cover. Remote Sensing 1:345–354.

Fritz, S., I. McCallum, C. Schill, C. Perger, L. See, D. Schepaschenko, M. van der Velde, F. Kraxner, and M. Obersteiner. 2012. Geo-Wiki: An online platform for improving global land cover. Environmental Modelling and Software 31:110–123.

Gallego, F. J. 2010. A population density grid of the European Union. Population and Environment 31:460–473.

Gardiner, M. M., L. L. Allee, P. M. J. Brown, J. E. Losey, H. E. Roy, and R. R. Smyth. 2012. Lessons from lady beetles: accuracy of monitoring data from US and UK citizen-science programs. Frontiers in Ecology and the Environment 10:471–476.

Garduño Freeman, C. 2010. Photosharing on Flickr: intangible heritage and emergent publics. International Journal of Heritage Studies 16:352–368.

Gee, K., and B. Burkhard. 2010. Cultural ecosystem services in the context of offshore wind farming: A case study from the west coast of Schleswig-Holstein. Ecological Complexity 7:349–358.

Geoghegan, H., A. Dyke, R. Pateman, and G. Everett. 2016. Understanding Motivations for Citizen Science. Final report on behalf of UKEOF. niversity of Reading, Stockholm Environment Institute (University of York) and University of the West of England.

Ghermandi, P. A. L. D. Nunes, and F. Eni. 2009. Recreational, Cultural and Aesthetic Services from Estuarine and Coastal Ecosystems. Fondazione Eni Enrico Mattei, Milan.

GiGL. 2016. Key London Figures. http://www.gigl.org.uk/our-data-holdings/keyfigures/.

Giles, J. 2005. Internet encyclopaedias go head to head. Nature 438:900–901.

Girardin, F., F. D. Fiore, C. Ratti, and J. Blat. 2008. Leveraging explicitly disclosed location information to understand tourist dynamics: a case study. Journal of Location Based Services 2:41–56.

Girres, J.-F., and G. Touya. 2010. Quality Assessment of the French OpenStreetMap Dataset. Transactions in GIS 14:435–459.

Glaser, M., B. Ratter, G. Krause, and M. Welp. 2012. New Approaches to the Analysis of Human-Nature Relations. Page 231*in* M. Glaser, B. Ratter, G. Krause, and M. Welp, editors.Human-nature interactions in the anthropocene : potentials of social-ecological systems analysis. Routledge.

Glasze, G., and C. Perkins. 2015. Social and Political Dimensions of the OpenStreetMap Project: Towards a Critical Geographical Research Agenda. Pages 143–166*in* J. J. Arsanjani, P. Mooney, A. Zipf, and M. Helbich, editors.OpenStreetMap in GIScience. Lecture No. Springer, Heidelberg, New York, Dordrecht, London.

Gliozzo, G. 2013. Crowdsourced geodata and the semantic web : a use case. Page Proceedings of GIS Research UK 2013. Liverpool.

Gómez-Baggethun, E., D. N. Barton, P. Berry, R. Dunford, and P. A. Harrison. 2016. Concepts and Methods in Ecosystem Services Valuation. Pages 99–111*in* M. B. Potschin, R. Haines-Young, R. Fish, and R. K. Turner, editors.Routledge Handbook of Ecosystem Services. Routledge, London and New York.

Gómez-Baggethun, E., R. de Groot, P. L. Lomas, and C. Montes. 2010. The history of ecosystem services in economic theory and practice: From early notions to markets and payment schemes. Ecological Economics 69:1209–1218.

Gomez-Baggethun, E., and M. Ruiz-Perez. 2011. Economic valuation and the commodification of ecosystem services. Progress in Physical Geography 35:613–628.

Goodchild, M. F. 2007a. Citizens as sensors: the world of volunteered geography. GeoJournal 69:211–221.

Goodchild, M. F. 2007b. Editorial : Citizens as Voluntary Sensors : Spatial Data Infrastructure in the World of Web 2 . 0. International Journal of Spatial Data Infrastructures Research 2:24–32.

Graham, M., B. Hogan, R. K. Straumann, and A. Medhat. 2014. Uneven Geographies of User-Generated Information: Patterns of Increasing Informational Poverty. Annals of the Association of American Geographers 104:746–764.

Graham, M., and S. De Sabbata. 2015. Mapping information wealth and poverty: the geography of gazetteers. Environment and Planning A 47:1254–1264.

Graham, M., M. Stephens, and S. Hale. 2013. Featured graphic. mapping the geoweb: A geography of Twitter. Environment and Planning A 45:100–102.

Gregory, I., A. Baron, D. Cooper, A. Hardie, P. Murrieta-Flores, and P. Rayson. 2014. Crossing Boundaries: Using GIS in Literary Studies, History and Beyond. Keys for architectural history research in the digital era (« Actes de colloques »).

de Groot, R., B. Fisher, M. Christie, J. Aronson, L. Braat, J. Gowdy, R. H. Haines-Young, E. Maltby, A. Neuville, S. Polasky, R. Portela, and I. Ring. 2010. Integrating the ecological and economic dimensions in biodiversity and ecosystem service valuation. Pages 1–40*in* TEEB, editor.The Economics of Ecosystems and Biodiversity: Ecological and economic foundation. Earthscan, Cambridge.

Haines-Young, R. H., and M. B. Potschin. 2009a. The links between biodiversity, ecosystem services and human well-being. Pages 110–139*in* D. Raffaelli and C. L. J. Frid, editors.Ecosystem ecology: a new synthesis. Cambridge University Press, Cambridge.

Haines-Young, R. H., and M. B. Potschin. 2009b. Methodologies for defining and assessing ecosystem services. Final Report, JNCC, Project Code C08-0170-0062.

Haines-Young, R. H., and M. B. Potschin. 2011. Common International Classification of Ecosystem Services (CICES): 2011 Update European Environment Agency. Centre for Environmental Management, University of Nottingham, UK, Nottingham.

Haines-Young, R. H., and M. B. Potschin. 2016. Categorisation systems : The classification challenge. Nottingham.

Haklay, M. 2010a. Geographical Citizen Science – clash of cultures and new opportunities. Position paper for GIScience workshop on the role of VGI in advancing science:1–6.

Haklay, M. 2010b. How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets. Environment and Planning B: Planning and Design 37:682–703.

Haklay, M. 2013. Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. Pages 105–122*in* D. Sui, S. Elwood, and M. F. Goodchild, editors.Crowdsourcing Geographic Knowledge. Springer Netherlands, Dordrecht.

Haklay, M. 2016a. Why is participation inequality so important? Pages 35–44*in* C.

Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors.European Handbook of Crowdsourced Geographic Information. Ubiquity P. London, UK.

Haklay, M. 2016b. Why is participation inequality important? Pages 35–44*in* C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors.European Handbook of Crowdsourced Geographic Information. Ubiquity Press Licence CC-BY 4.0, London, UK, UK.

Haklay, M., V. Antoniou, S. Basiouka, R. Solden, and P. Mooney. 2014. Crowdsourced geographic information use in government. World Bank - Global Facility for Disaster Reduction and Recovery, London.

Haklay, M., S. Basiouka, V. Antoniou, and A. Ather. 2010. How many volunteers does it take to map an area well? The validity of linus' law to Volunteered Geographic Information. The Cartographic Journal 47:315–322.

Haklay, M., A. Singleton, and C. Parker. 2008. Web Mapping 2 . 0 : The Neogeography of the GeoWeb. Geography Compass 2:2011–2039.

Haklay, M., and P. Weber. 2008. OpenStreet map: User-generated street maps. IEEE Pervasive Computing 7:12–18.

Hampton, S. E., C. A. Strasser, J. J. Tewksbury, W. K. Gram, A. E. Budden, A. L. Batcheller, C. S. Duke, and J. H. Porter. 2013. Big data and the future of ecology. Frontiers in Ecology and the Environment 11:156–162.

Hand, E. 2010. Citizen science: People power. Nature 466:685–687.

Harvey, F. 2013. To volunteer or to contribute locational information? Towards truth in labeling for crowdsourced geographic information. Pages 31–42*in* D. Sui, S. Elwood, and M. Goodchild, editors.Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice. Springer Netherlands, Dordrecht.

Hawkes, J. 2001. The Fourth Pillar of Sustainability: Culture's essential role in public planning. Page Common Ground Publishing Pty Ltd.

Haywood, B. K. 2014. A "Sense of Place" in Public Participation in Scientific Research. Science Education 98:64–83.

Hecht, B. J., and D. Gergle. 2010. On the "localness" of user-generated content. Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10:229–232.

Hecht, B., M. Rohs, J. Schöning, and A. Krüger. 2007. WikEye – Using Magic Lenses to Explore Georeferenced Wikipedia Content. Proceedings of the 3rd International Workshop on Pervasive Mobile Interaction Devices PERMID (2007):6–10.

Hecht, B., and M. Stephens. 2014. A Tale of Cities : Urban Biases in Volunteered Geographic Information. Pages 197–205Proceedings of the Eighth International

AAAI Conference on Weblogs and Social Media. Association for the Advancement of Artificial Intelligence (www.aaai.org).

Herbert, S., and M. Brobyn. 2011. Spatial Report 2009 - 2011. Page Monitor of Engagement with the Natural Environment. Natural England.

Hernández-Morcillo, M., T. Plieninger, and C. Bieling. 2013. An empirical review of cultural ecosystem service indicators. Ecological Indicators 29:434–444.

Hey, T. 2009. The fourth paradigm : Data-intensive scientific discovery. Page (T. Hey, S. Tansley, and K. Tolle, Eds.) E-Science and Information Management: Third International Symposium on Information Management in a Changing World, IMCW 2012, Ankara, Turkey, September 19-21, 2012. Proceedings. Microsoft Research, Redmond, Washington.

Hinterberger, F., F. Luks, and F. Schmidt-Bleek. 1997. Material flows vs. “natural capital”: What makes an economy sustainable? Ecological Economics 23:1–14.

Hobbs, S. J., and P. C. L. L. White. 2012. Motivations and barriers in relation to community participation in biodiversity recording. Journal for Nature Conservation 20:364–373.

Hochmair, H. H. 2010. Spatial association of geotagged photos with scenic locations. Pages 91–100*in* G. Griesebner and J. Strobl, editors.In Proceedings of the Geoinformatics Forum Salzburg. Wichmann, Heidelberg.

Höflich, J. R. 2005. A Certain Sense of Place Mobile comunication and Local orientation. Pages 159–168*in* K. Nyiri, editor.A Sense of Place: The Global and the Local in Mobile Communication. Passagen, Wien.

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

Howard, B., J. Neumann, and R. O’ Riodan. 2016. Tool Assessor Supporting practical assessment of natural capital in land-use decision making. Peterborough.

Howe, J. 2006. The rise of crowdsourcing. Wired Magazine 14:1–4.

Hristova, D., G. Quattrone, A. Mashhadi, and L. Capra. 2013. The life of the Party: Impact of Social Mapping in OpenStreetMap. Pages 234–243Proceedings of ICWSM ’13. AAAI.

Huson, L. W. 2007. Performance of Some Correlation Coefficients When Applied to Zero-Clustered Data. Journal of Modern Applied Statistical Methods 6:530–536.

Irvine, K. N., L. O’Brien, N. Ravenscroft, N. Cooper, M. Everard, I. Fazey, M. S. Reed, and J. O. Kenter. 2016. Ecosystem services and the idea of shared values. Ecosystem Services 21:184–193.

Jackson, C., C. Østerlund, K. Crowston, G. Mugar, and K. D. Hassman. 2015. Motivations for Sustained Participation in Citizen Science: Case Studies on the Role of Talk. Pages 1624–1634System Sciences (HICSS), 2015 48th Hawaii International Conference on. IEEE Computer Society.

Jeawak, S. S., C. B. Jones, and S. Schockaert. 2017. Using Flickr for characterizing the environment : an exploratory analysis. Page 21:1-21:13 *in* E. Clementini, M. Donnelly, M. Yuan, C. Kray, P. Fogliaroni, and A. Ballatore, editors. 13th International Conference on Spatial Information Theory (COSIT 2017). Schloss Dagstuhl--Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany.

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

Jiang, B. 2016. Head/tail Breaks for Visualization of City Structure and Dynamics. Pages 169–183*in* C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors.European Handbook of Crowdsourced Geographic Information. Ubiquity Press License: CC-BY 4.0., London, UK.

Juhász, L., and H. H. Hochmair. 2017. Where to catch ‘ em all ? – A geographic analysis of Pokémon Go locations. Geo-spatial Information Science 5020:1–3.

Kachkaev, A., and J. Wood. 2013. Crowd-sourced Photographic Content for Urban Recreational Route Planning. Page University Transport Study Group UK.

Kachkaev, A., and J. Wood. 2014. Automated planning of leisure walks based on crowd-sourced photographic content. Pages 227–246Paper presented at the 46th Annual Universities' Transport Study Group Conference, 06-01-2014 - 08-01-2014, Newcastle, UK.

Kaplan, R. D. 2015. The Revenge of Geography. Orbis 59:479–490.

Katz, J. E., and M. Aakhus. 2002. Perpetual Contact: Mobile Communication, Private Talk, Public Performance. Page (J. E. Katz and M. A. Aakhus, Eds.) Perpetual Contact Mobile Communication Private Talk Public Performance. Cambridge University Press.

Kenter, J. O. 2016a. Deliberative and non-monetary valuation. Pages 271–287*in* M. Potschin, R. Haines-Young, R. Fish, and R. K. Turner, editors.Routledge handbook of Ecosystem Services. Routledge, London and New York.

Kenter, J. O. 2016b. Editorial: Shared, plural and cultural values. Ecosystem Services 21:175–183.

Kenter, J. O. 2016c. Integrating deliberative monetary valuation, systems modelling and participatory mapping to assess shared values of coastal ecosystem services. Ecosystem Services 21:291–307.

Kenter, J. O., L. O'Brien, N. Hockley, N. Ravenscroft, I. Fazey, K. N. Irvine, M. S. Reed, M. Christie, E. Brady, R. Bryce, A. Church, N. Cooper, A. Davies, A. Evely, M. Everard, R. Fish, J. A. Fisher, N. Jobstvogt, C. Molloy, J. Orchard-Webb, S. Ranger, M. Ryan, V. Watson, and S. Williams. 2015. What are shared and social values of ecosystems? Ecological Economics 111:86–99.

Kenter, J. O., M. S. Reed, and I. Fazey. 2014a. The Deliberative Value Formation model. Ecosystem Services 21:194–207.

Kenter, J. O., M. S. Reed, K. N. Irvine, E. O'Brien, E. Brady, R. Bryce, M. Christie, A. Church, N. Cooper, A. Davies, N. Hockley, I. Fazey, N. Jobstvogt, C. Molloy, N. Orchard-Webb, J., Ravenscroft, M. Ryan, and V. Watson. 2014b. UK National Ecosystem AssessmentFollow-on. Work Package Report 6: Shared, plural and cultural values of ecosystems. Page UK National Ecosystem Assessment Follow-on. UNEP-WCMC, LWEC, UK.

Kenter, J. O., M. S. Reed, K. N. Irvine, E. O'Brien, R. Bryce, M. Christie, N. Cooper, N. Hockley, I. Fazey, C. . Orchard-Webb, J., Ravenscroft, N., Raymond, P. Tett, and V. Watson. 2016. Shared values and deliberative valuation: Future directions. Ecosystem Services 21:358–371.

Keßler, C., and R. T. A. de Groot. 2013. Trust as a Proxy Measure for the Quality of Volunteered Geographic Information in the Case of OpenStreetMap. Pages 225–245Geographic Information Science at the Heart of Europe. Lecture Notes in Geoinformation and Cartography.

Keßler, C., J. Trame, and T. Kauppinen. 2011. Tracking editing processes in volunteered geographic information: The case of OpenStreetMap. Identifying objects, processes and events in spatio-temporally distributed data (IOPE), workshop at conference on spatial information theory 12:6–8.

Kienast, F., and J. Holfenstein. 2016. Modelling Ecosystem Services. Pages 144–155*in* M. B. Potschin, R. Haines-Young, M. Potschin, R. Fish, and K. Turner, editors.Routledge Handbook of Ecosystem Services. Routledge, London and New York.

Kisilevich, S., M. Krstajic, D. Keim, N. Andrienko, and G. Andrienko. 2010a. Event-Based Analysis of People's Activities and Behavior Using Flickr and Panoramio Geotagged Photo Collections. 2010 14th International Conference Information Visualisation:289–296.

Kisilevich, S., F. Mansmann, P. Bak, D. Keim, and A. Tchaikin. 2010b. Where Would You Go on Your Next Vacation? A Framework for Visual Exploration of Attractive Places. 2010 Second International Conference on Advanced Geographic Information Systems, Applications, and Services:21–26.

Kitchin, R. 2013. Big data and human geography : Opportunities , challenges and risks. Dialogues in Human Geography 3:262–267.

Kittur, A., and R. E. Kraut. 2008. Harnessing the wisdom of crowds in wikipedia: quality through coordination. Proceedings of the 2008 ACM conference on Computer supported cooperative work:37–46.

Kuemmerle, T., C. Levers, M. Baumann, K. Tieskens, and S. Nynke. 2014. Indicator database and report describing the input data, methodology, and data quality for each indicator. Page HERCULES Sustainable futures for Europe's HERitage in CULtural landscapES: Tools for understanding, managing, and protecting landscape functions and values. HERCULES.

Lambio, C., and T. Lakes. 2017. Placing of photos on the internet: Critical analysis of biases on the depictions of France and Afghanistan on FLICKR. Geoforum 82:21–25.

Landers, D. H., and A. M. Nahlik. 2013. Final Ecosystem Goods and Services Classification System (FEGS-CS). EPA/600/R-13/ORD-004914., U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C.

Landers, D. H., A. M. Nahlik, and C. R. Rhodes. 2016. The Beneficiary Perspective - Benefits and Beyond. Pages 74–87*in* M. Potchin, R. H. Haines-Young, R. Fish, and R. K. Turner, editors.Routledge Handbook on Ecosystem Services. Routledge, London and New York.

Laney, D. 2001. 3-D Data Management: Controlling Data Volume, Velocity and Variety. Application Delivery Strategies 949:4.

Lanter, D. P. 1990. LINEAGE IN GIS: THE PROBLEM AND A SOLUTION. Page NCGIA Technical Paper. National Center for Geographic Information and Analysis, Santa Barbara, CA 93106.

Lemmens, R., and D. Deng. 2008. Web 2.0 and Semantic Web: Clarifying the meaning of spatial features. Pages 1–14Workshop “Semantic Web meets Geospatial Applications”, held in conjunction with AGILE 2008.

Lemmens, R., F. Gilles, S. De Sabbata, B. Jiang, and B. Bucher. 2016. Querying VGI by semantic enrichment. Pages 185–194*in* C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors.European Handbook of Crowdsourced Geographic Information. Ubiquity Press License: CC-BY 4.0., London, UK.

Lesser, E., D. Ransom, R. Shah, and B. Pulver. 2012. Collective Intelligence: Capitalizing on the crowd. Page Business & Information Systems Engineering.

Li, L., M. F. Goodchild, and B. Xu. 2013. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. Cartography and Geographic Information Science 40:61–77.

Lin, W. 2015. Revealing the making of OpenStreetMap: A limited account. The Canadian Geographer / Le Géographe canadien 59:69–81.

Lin, Y.-P., W.-C. Lin, H.-Y. Li, Y.-C. Wang, C.-C. Hsu, W.-Y. Lien, J. Anthony, and J. R.

Petway. 2017a. Integrating Social Values and Ecosystem Services in Systematic Conservation Planning : A Case Study in Datuan Watershed. Sustainability 9:22.

Lin, Y.-P., W.-C. Lin, H.-Y. Li, Y.-C. Wang, C.-C. Hsu, W.-Y. Lien, J. Anthony, and J. R. Petway. 2017b. Supplementary material to:Integrating Social Values and Ecosystem Services in Systematic Conservation Planning: A Case Study in Datuan Watershed.

Liu, Z., D. R. Scannell, M. B. Eisen, and R. Tjian. 2006. Encyclopaedia Britannica and Nature: a response 23 March 2006 In our issue of 15 December 2005 we published a news article that compared the Internet offerings of. Nature 438:720–731.

Loai Ali, A. 2016. Tackling the thematic accuracy of areal features in OpenStreetMap. Pages 113–129*in* C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors.European Handbook of Crowdsourced Geographic Information. Ubiquity Press License: CC-BY 4.0., London, UK.

Local Records Centres. 2014. Working together for biodiversity in London and South East England.

Loft, L., A. Lux, and T. Jahn. 2016. A social-ecological perspective on ecosystem services. Pages 88–93The routledge handbook on ecosystem services. Routledge, London and New York.

Longley, P. A., and M. Adnan. 2016. Geo-temporal Twitter demographics. International Journal of Geographical Information Science 30:369–389.

López-Santiago, C. A., E. Oteros-rozas, B. Martín-lópez, T. Plieninger, E. G. Martín, and A. José. 2014. Using visual stimuli to explore the social perceptions of ecosystem services in cultural landscapes : the case of transhumance in Mediterranean Spain. Ecology And Society 19:16.

Luck, G. W., G. C. Daily, and P. R. Ehrlich. 2003. Population diversity and ecosystem services. Trends in Ecology and Evolution 18:331–336.

Ludwig, I., A. Voss, and M. Krause-Traudes. 2010. How Good is OSM ? - Method and Results for Germany Categories and Subject Descriptors. Page Sixth International Conference on Geographic Information Science 2010 Zurich, Switzerland 14-17 Sep 2010. Zurich, Switzerland.

Mace, G. M. 2016. Ecosystem Services: Where is the Discipline Heading? Pages 602–606*in* M. B. Potschin, R. Haines-Young, M. Potschin, R. Fish, and K. Turner, editors.Routledge Handbook of Ecosystem Services. Routledge, London and New York.

Maes, J., N. D. Crossman, and B. Burkhard. 2016. Mapping Ecosystem Services. Pages 188–204*in* M. B. Potschin, R. Haines-Young, M. Potschin, R. Fish, and K. Turner, editors.Routledge Handbook of Ecosystem Services. Routledge, London and New

York.

Maes, J., B. Egoh, L. Willemen, C. Liquete, P. Vihervaara, J. P. Schägner, B. Grizzetti, E. G. Drakou, A. La Notte, G. Zulian, F. Bouraoui, M. Luisa Paracchini, L. Braat, and G. Bidoglio. 2012. Mapping ecosystem services for policy support and decision making in the European Union. Ecosystem Services 1:31–39.

Markoff, J. 2009. A Deluge of Data Shapes a New Era in Computing. The New York Times.

Martin, D., S. Cockings, and S. Leung. 2015. Developing a Flexible Framework for Spatiotemporal Population Modeling. Annals of the Association of American Geographers 105:754–772.

Martinez, D. 2006. Protected Areas, Indigenous Peoples, and the Western Idea of Nature. Pages 215–218*in* D. Harmon, editor.Proceedings of the 2005 George Wright Society Conference on Parks, Protected Areas, and Cultural Sites. The George Wright Society, Hancock, Michigan.

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

Martinotti, G. 1993. Metropoli. La nuova morfologia sociale della città. Page icon. Il Mulino, Bologna.

Martinotti, G. 1996. Four Populations : Human Settlements and Social Morphology in the Contemporary Metropolis. European Review 4:3–23.

Mashhadi, A., G. Quattrone, and L. Capra. 2013. Putting ubiquitous crowd-sourcing into context. Proceedings of the 2013 conference on Computer supported cooperative work - CSCW ’13 1:611.

Mashhadi, A., G. Quattrone, and L. Capra. 2015. The Impact of Society on Volunteered Geographic Information: The Case of OpenStreetMap. Pages 125–141*in* J. Jokar Arsanjani, A. Zipf, P. Mooney, and M. Helbich, editors.OpenStreetMap in GIScience: Experiences, Research, and Applications. Notes in G. Springer International Publishing, Heidelberg, New York, Dordrecht, London.

Mateos, P. 2014. Names, Ethnicity and Populations. Springer, Heidelberg, New York, Dordrecht, London.

Max-Neef, M. A. 1991. Human scale development: conception, application and further reflections. Apex press.

Meadows, D. ., D. Meadows, J. Randers, and W. Behrens. 1972. The Limits to Growth. Page Chelsea. New York.

Medcalf, K. ., N. Small, C. . Finch, and J. . Parker. 2012. Spatial framework for assessing evidence needs for operational ecosystem approaches. Page JNCC Report No 469. Peterborough.

Mendes, R. N., T. Rodrigues, and A. M. Rodrigues. 2013. Urban geocaching. What happened in Lisbon during the last decade? International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 29th Urban Data Management Symposium, 29 – 31 May, 2013, London, United Kingdom XL:29–31.

Milcu, A. I., J. Hanspach, D. J. Abson, and J. Fischer. 2013. Cultural Ecosystem Services : A Literature Review and Prospects for future research. Ecology and Society 18:44.

Millennium Ecosystem Assessment. 2003. Ecosystems and human well-being: a framework for assessment. Island Press, Washington, D.C.

Millennium Ecosystem Assessment. 2005. Ecosystems and Human Well-Being: Synthesis. Island Press, Washington, DC.

Mitchell, W. 1996. City of bits. Page City of Bits. MIT Press, Cambridge, MA.

Mitchell, W. J. 2001. Where I'm @. Journal of the American Planning Association 67:144.

Mitchell, W. J. 2005. The Revenge of Place. Pages 45–53*in* K. Rattenbury, editor.This is Not Architecture: Media Constructions. Routledge, London.

Mooney, P., and P. Corcoran. 2012a. Characteristics of Heavily Edited Objects in OpenStreetMap. Future Internet 4:285–305.

Mooney, P., and P. Corcoran. 2012b. How social is OpenStreetMap. Proceedings of the 15th Association of AGILE 2012 International Conference on Geographic Information Science:24–27.

Mooney, P., and P. Corcoran. 2014. Analysis of interaction and co-editing patterns amongst openstreetmap contributors. Transactions in GIS 18.

Mooney, P., P. Corcoran, and A. Winstanley. 2010a. A study of data representation of natural features in OpenStreetMap. Proceedings of the 6th GIScience International Conference on Geographic Information Science:150.

Mooney, P., P. Corcoran, and A. C. Winstanley. 2010b. Towards quality metrics for OpenStreetMap. Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems GIS 10:514–517.

Mooney, P., A.-M. Olteanu-Raimond, G. Touya, N. Juul, S. Alvanides, and N. Kerle. 2017. Considerations of Privacy , Ethics and Legal Issues in Volunteered Geographic Information. Pages 119–135 in G. Foody, L. See, S. Fritz, P. Mooney, A.-M. Olteanu-Raimond, C. Fonte, and V. Antoniou, editors. Mapping and the Citizen Sensor. Ubiquity Press Licence: CC-BY 4.0, London.

Mooney, P., K. Rehrl, and H. Hochmair. 2013. Action and interaction in volunteered geographic information: a workshop review. Journal of Location Based Services 7:291–311.

Mourato, S., G. Atkinson, M. Collins, S. Gibbons, G. Mackerron, and G. Resende. 2010. Economic Analysis of Cultural Services. Page UK NEA Economic Analysis Report. UK NEA Economic Analysis Report, London, WC2A 2AE.

Murrieta-Flores, P., C. Donaldson, C. J. Rupp, D. Cooper, and I. Gregory. 2013. Digital Literary Geographies : A Spatial Analysis of Lake District Landscape Writings:283850.

Napolitano, M., and P. Mooney. 2011. Mvp osm: A tool to identify areas of high quality contributor activity in openstreetmap. Bulletin of the Society of Cartographers 45:31–36.

Natural England. 2009. Experiencing landscapes: capturing the cultural services and experiential qualities of landscape: Annex B, copies of the research topic guides used. Natural England Commissioned Report NECR024:1–19.

Natural England. 2011. Monitor of Engagement with the Natural Environment Spatial Report 2009 - 2011.

Natural England. 2014. Note on surveys undertaken in the outdoor recreation, leisure, sport, heritage and related areas.

NaturalEngland. 2016. Monitor of Engagement with the Natural Environment: The national survey on people and the natural environment –Annual Report from the 2010–11 Survey. Natural England Commissioned Report NECR083.

Neis, P., and A. Zipf. 2012. Analyzing the Contributor Activity of a Volunteered Geographic Information Project — The Case of OpenStreetMap. ISPRS International Journal of Geo-Information 1:146–165.

Nielsen, J. 2006. The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities. https://www.nngroup.com/articles/participation-inequality/.

Nov, O., O. Arazy, and D. Anderson. 2011. Technology-Mediated Citizen Science Participation : A Motivational Model. Artificial Intelligence.

Nov, O., M. Naaman, and C. Ye. 2010. Analysis of Participation in an Online Photo-Sharing Community : A Multidimensional Perspective. Journal of the American Society for Information Science and Technology 61:555–566.

O' Reilly, T. 2005. What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html.

Openshaw, S., and P. J. Taylor. 1979. A Million or so Correlation Coefficients: Three Experiments on the Modifiable Areal Unit Problem. Pages 127–144*in* N. Wringley, editor.Statistical Applications in the Spatial Sciences. Pion, London.

OSM contributors. 2015. Wiki main page. http://wiki.openstreetmap.org/.

Ostermann, F. O., Huang, N. Andrienko, C. Capineri, K. Farkas, and R. S. Purves. 2015. Extracting and Comparing Places Using Geo-Social Media. Page The 9th

International Symposium on Spatial Data Quality (ISSDQ 2015). La Grande Motte, France.
Oteros-Rozas, E., B. Martín-López, N. Fagerholm, C. Bieling, and T. Plieninger. 2017. Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. Ecological Indicators.
Pagella, T. F., and F. L. Sinclair. 2014. Development and use of a typology of mapping tools to assess their fitness for supporting management of ecosystem service provision. Landscape Ecology 29:383–399.
Palomo, I., B. Martín-López, M. B. Potschin, R. H. Haines-Young, and C. Montes. 2013. National Parks, buffer zones and surrounding lands: Mapping ecosystem service flows. Ecosystem Services 4:104–116.
Palomo, I., C. Montes, B. Martín-López, J. a. González, M. García-Llorente, P. Alcorlo, and M. R. G. Mora. 2014. Incorporating the social-ecological approach in protected areas in the anthropocene. BioScience 64:181–191.
Pascual, U., R. Muradian, L. Brander, E. Gómez-baggethun, B. Martín-lópez, M. Verma, P. Armsworth, M. Christie, H. Cornelissen, F. Eppink, J. Farley, L. Pearson, C. Perrings, S. Polasky, J. Mcneely, R. Norgaard, R. Siddiqui, R. D. Simpson, and R. K. Turner. 2010. The economics of valuing ecosystem services and biodiversity. Page The Economics of Ecosystems and Biodiversity.
Pasley, R., P. Clough, R. S. Purves, and F. a. Twaroch. 2008. Mapping geographic coverage of the web. Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems:1–9.
Paulos, E. 2012. You amateur! Interactions 19:52.
Paulos, E., R. J. Honicky, and B. Hooker. 2009. Citizen Science - Enabling Participatory Urbanism. Page 470*in* M. Foth, editor.Handbook of research on urban informatics : the practice and promise of the real-time city. Idea Group Inc (IGI).
Perkins, C. 2014. Plotting practices and politics: (Im)mutable narratives in OpenStreetMap. Transactions of the Institute of British Geographers 39:304–317.
Perrings, C. 2006. Ecological Economics after the Millennium Assessment Charles. International Journal of Ecological Economics & Statistics (IJEES) 6:8–22.
Peterson, M. J., D. M. Hall, A. M. Feldpausch-Parker, and T. R. Peterson. 2010. Obscuring ecosystem function with application of the ecosystem services concept: Essay. Conservation Biology 24:113–119.
Pettorelli, N., H. Schulte to Bühne, A. Tulloch, G. Dubois, C. Macinnis-Ng, A. M. Queirós, D. A. Keith, M. Wegmann, F. Schrodt, M. Stellmes, R. Sonnenschein, G. N. Geller, S. Roy, B. Somers, N. Murray, L. Bland, I. Geijzendorffer, J. T. Kerr, S. Broszeit, P. J. Leitão, C. Duncan, G. El Serafy, K. S. He, J. L. Blanchard, R. Lucas, P. Mairota,

T. J. Webb, and E. Nicholson. 2017. Satellite remote sensing of ecosystem functions: opportunities, challenges and way forward. Remote Sensing in Ecology and Conservation:1–23.

Phillips, T., R. Bonney, and J. L. Shirk. 2012. What Is Our Impact? Toward a Unified Framework for Evaluating Outcomes of Citizen Science Participation. Page 302*in* J. L. Dickinson and R. Bonney, editors.Citizen Science: Public Participation in Environmental Research. Cornell University Press.

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

Pocock, M. J. O., H. E. Roy, C. D. Preston, and D. B. Roy. 2015. The Biological Records Centre: A pioneer of citizen science. Biological Journal of the Linnean Society 115:475–493.

Polous, K., P. Mooney, J. M. Krisp, and L. Meng. 2013. Mining Event-Related Knowledge from OpenStreetMap. Pages 109–124*in* J. M. Krisp, editor.Progress in Location-Based Services. Springer, Heidelberg, New York, Dordrecht, London.

Potschin, M. B., and R. H. Haines-Young. 2011. Ecosystem services: Exploring a geographical perspective. Progress in Physical Geography 35:575–594.

Potschin, M. B., and R. H. Haines-Young. 2013. Landscapes, sustainability and the place-based analysis of ecosystem services. Landscape Ecology 28:1053–1065.

Potschin, M. B., and R. H. Haines-Young. 2016a. Defining and measuring ecosystem services. Pages 1–18*in* M. Potschin, R. Haines-Young, R. Fish, and R. K. Turner, editors.Routledge Handbook of Ecosystem Services. Routledge, London and New York.

Potschin, M. B., and R. H. Haines-Young. 2016b. Frameworks for ecosystem assessments. Pages 128–139*in* M. B. Potschin, R. Haines-Young, M. Potschin, R. Fish, and K. Turner, editors.Routledge Handbook of Ecosystem Services. Routledge, London and New York.

Potschin, M. B., R. H. Haines-Young, R. Fish, and K. Turner. 2016a. Ecosystem Services in the Twenty-First Century. Pages 1–9*in* M. Potschin, R. Haines-Young, R. Fish, and K. Turner, editors.Routledge Handbook of Ecosystem Services. Routledge, London and New York.

Potschin, M. B., R. H. Haines-Young, R. Fish, and R. K. Turner. 2016b. Routledge Handbook of Ecosystem Services.

Potschin, M. B., E. Primmer, E. Furman, and R. H. Haines-Young. 2016c. Have Ecosystem Services Been Oversold? A Response to Silvertown. Trends in Ecology and Evolution 31:334–335.

Powney, G. D., and N. J. B. Isaac. 2015. Beyond maps: A review of the applications of

biological records. Biological Journal of the Linnean Society 115:532–542.
Priess, J. A., and L. Kopperoinen. 2017. Citizen Science. Pages 223–226*in* B. Burkhard and J. Maes, editors.Mapping Ecosystem Services. Pensoft Publishers, Sofia.
Purves, R. S., and C. Derungs. 2015. From Space to Place: Place-Based Explorations of Text. International Journal of Humanities and Arts Computing 9:74–94.
Purves, R. S., A. Edwardes, and J. Wood. 2011. Describing places through user generated content. First Monday 16:1–10.
Purves, R. S., and W. A. Mackaness. 2016. A methodological toolbox for exploring collections of textually annotated georeferenced photographs. Pages 145–156*in* C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors.European Handbook of Crowdsourced Geographic Information. Ubiquity Press License: CC-BY 4.0., London, UK.
Quattrone, G., L. Capra, and P. De Meo. 2015. There's No Such Thing as the Perfect Map: Quantifying Bias in Spatial Crowd-sourcing Datasets. Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15:1021–1032.
Quattrone, G., A. Mashhadi, and L. Capra. 2014a. Mind the Map : The Impact of Culture and Economic Affluence on Crowd-Mapping Behaviours. Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing - CSCW '14:934–944.
Quattrone, G., D. Quercia, C. Smith-clarke, A. Mashhadi, D. Quercia, C. Smith-clarke, and L. Capra. 2014b. Modelling Growth of Urban Crowd-Sourced Information. Pages 563–572Proceedings of the 7th ACM international conference on Web search and data mining - WSDM '14.
Queiroz, C., M. Meacham, K. Rich, A. V Norström, J. Norberg, and G. D. Peterson. 2015a. Electronic Supplementary Material for Mapping bundles of ecosystem services reveals distinct types of multifunctionality within a Swedish landscape. Ambio 44:13.
Queiroz, C., M. Meacham, K. Richter, A. V. Norström, E. Andersson, J. Norberg, and G. D. Peterson. 2015b. Mapping bundles of ecosystem services reveals distinct types of multifunctionality within a Swedish landscape. Ambio 44:89–101.
Quercia, D., and D. Saez. 2014. Mining Urban Deprivation from Foursquare: Implicit Crowdsourcing of City Land Use Location. PERVASIVE computing:30–36.
Raffaelli, D., J. M. Bullock, S. Cinderby, I. Durance, B. Emmett, J. Harris, K. Hicks, T. H. Oliver, D. Paterson, and P. C. L. White. 2014. Big Data and Ecosystem Research Programmes. Pages 41–77*in* G. Woodward, A. J. Dumbrell, D. J. Baird, and M. Hajibabaei, editors.Advances in Ecological Research. Academic Press.

Raffestin, C. 2012. Space, territory, and territoriality. Environment and Planning D: Society and Space 30:121–141

Raudsepp-Hearne, C., G. D. Peterson, and E. M. Bennett. 2010. Ecosystem service bundles for analyzing tradeoffs in diverse landscapes. Proceedings of the National Academy of Sciences of the United States of America 107:5242–7.

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

Roll, U., J. C. Mittermeier, G. I. Diaz, M. Novosolov, A. Feldman, Y. Itescu, S. Meiri, and R. Grenyer. 2016. Using Wikipedia page views to explore the cultural importance of global reptiles. Biological Conservation 204:42–50.

Russell, R., A. D. Guerry, P. Balvanera, R. K. Gould, X. Basurto, K. M. A. Chan, S. Klain, J. Levine, and J. Tam. 2013. Humans and Nature: How Knowing and Experiencing Nature Affect Well-Being. Annual Review of Environment and Resources 38:473–502.

Russell, S., T. Blackstock, M. Christie, M. Clarke, K. Davies, C. Duigan, I. Durance, R. Elliot, H. Evans, C. Falzon, P. Frost, S. Ginley, N. Hockley, S. Hourahane, L. Jones, J. Korn, P. Ogden, S. Pagella, T. F. Pagella, B. Pawson, B. Reynolds, D. Robinson, B. Sanderson, J. Sherry, J. Skates, E. Small, B. Spence, and C. Thomas. 2011. Status and Changes in the UK ' s Ecosystems and their Services to Society : Wales. Pages 979–1044 The UK National Ecosystem Assessment Technical Report.

Scassa, T., and H. Chung. 2015. Typology of Citizen Science Projects from and Intellectual Property Perspective: Invention and Authorship Between Researchers and Participants. Policy Memo Series 5:17.

Schaich, H., C. Bieling, and T. Plieninger. 2010. Linking Ecosystem Services with Cultural Landscape Research. Gaia Ecological Perspectives For Science And Society 19:269–277.

Schmidt, L., J. G. Nave, and J. Guerra. 2006. Who's afraid of Local Agenda 21? Top-down and bottom-up perspectives on local sustainability. International Journal of Environment and Sustainable Development 5:181–198.

Schmidt, M., and S. Klettner. 2013. Gender and Experience-Related Motivators for Contributing to OpenStreetMap. Pages 13–18*in* P. Mooney and K. Rehrl, editors.International Workshop on Action and Interaction in Volunteered Geographic Information (ACTIVITY). Leuven, May 14-17, 2013.

Schmidt, M., S. Klettner, and R. Steinmann. 2013. Barriers for contributing to VGI projects. Proc. ICC '13.

Schröter, M., and A. P. E. van Oudenhoven. 2016. Ecosystem Services Go Beyond Money and Markets: Reply to Silvertown. Trends in Ecology and Evolution 31:333–

334.

Schröter, M., E. H. van der Zanden, A. P. E. van Oudenhoven, R. P. Remme, H. M. Serna-Chavez, R. S. de Groot, and P. Opdam. 2014. Ecosystem Services as a Contested Concept: A Synthesis of Critique and Counter-Arguments. Conservation Letters 7:514–523.

Scott, A. 2002. Assessing Public Perception of Landscape: The LANDMAP experience. Landscape Research 27:271–295.

See, L., A. Comber, C. Salk, S. Fritz, M. van der Velde, C. Perger, C. Schill, I. McCallum, F. Kraxner, and M. Obersteiner. 2013. Comparing the quality of crowdsourced data contributed by expert and non-experts. PloS one 8:e69958.

See, L., P. Mooney, G. Foody, L. Bastin, A. Comber, J. Estima, S. Fritz, N. Kerle, B. Jiang, M. Laakso, H.-Y. Liu, G. Milčinski, M. Nikšič, M. Painho, A. Pődör, A.-M. Olteanu-Raimond, and M. Rutzinger. 2016. Crowdsourcing, Citizen Science or Volunteered Geographic Information? The Current State of Crowdsourced Geographic Information. ISPRS International Journal of Geo-Information 5:55.

Sen, A., A. Darnell, A. Crowe, I. Bateman, and P. Munday. 2011. Economic Assessment of the Recreational Value of Ecosystems in Great Britain - Report to the Economics Team of the UK National Ecosystem Assessment. Page UK NEA Economic Analysis Report. University of East Anglia.

Sen, A., A. R. Harwood, I. J. Bateman, P. Munday, A. Crowe, L. Brander, J. Raychaudhuri, A. A. Lovett, J. Foden, and A. Provins. 2014. Economic Assessment of the Recreational Value of Ecosystems: Methodological Development and National and Local Application. Environmental and Resource Economics 57:233–249.

Sen, S. W., H. Ford, D. R. Musicant, M. Graham, O. S. B. Keyes, and B. Hecht. 2015. Barriers to the Localness of Volunteered Geographic Information. Chi:197–206.

Senaratne, H., A. Bröring, and T. Schreck. 2013. Using reverse viewshed analysis to assess the location correctness of visually generated VGI. Transactions in GIS 17:369–386.

Senaratne, H., A. Mobasheri, A. L. Ali, C. Capineri, and M. Haklay. 2016. A review of volunteered geographic information quality assessment methods. International Journal of Geographical Information Science 8816:1–29.

Seresinhe, C. I., H. S. Moat, and T. Preis. 2017. Quantifying scenic areas using crowdsourced data. Environment and Planning B: Urban Analytics and City Science:26581351668730.

Seresinhe, C. I., T. Preis, and H. S. Moat. 2015. Quantifying the impact of scenic environments on health. Scientific Reports 5:16899.

Seymour, V. 2016. The Human–Nature Relationship and Its Impact on Health: A Critical Review. Frontiers in Public Health 4:1–12.

Seymour, V., and M. Haklay. 2017. Exploring Engagement Characteristics and Behaviours of Environmental Volunteers. Citizen Science: Theory and Practice 2:1–13.

Shaw, B. J., H. Draux, M. García Martín, J. Martin, and C. Bieling. 2017. Contributions of citizen science to landscape democracy: potentials and challenges of current approaches. Landscape Research:1–14.

Shirk, J. L., H. L. Ballard, C. C. Wilderman, T. Phillips, A. Wiggins, and R. Jordan. 2012. Public Participation in Scientific Research : a Framework for Deliberate Design. Ecology and Society 17:20.

Shoval, N., and M. Isaacson. 2007. Tracking tourists in the digital age. Annals of Tourism Research 34:141–159.

Sieber, J., and M. Pons. 2015. Assessment of urban ecosystem services using ecosystem services reviews and GIS-based tools. Procedia Engineering 115:53–60.

Silka, L. 2010. Community research in other contexts: learning from sustainability science. Journal of empirical research on human research ethics : JERHRE 5:3–11.

Silvertown, J. 2009. A new dawn for citizen science. Trends in ecology & evolution 24:467–71.

Silvertown, J. 2015. Have Ecosystem Services Been Oversold? Trends in Ecology & Evolution 30:641–648.

Simonite, T. 2013. The Decline of Wikipedia. http://www.technologyreview.com/featuredstory/520446/the-decline-of-wikipedia/.

De Smith, M. j, M. F. Goodchild, and P. A. Longley. 2007. Geospatial analysis: a Comprehensive Guide to Principles, Techniques and Software Tools. Page (M. J. De Smith, M. F. Goodchild, and P. A. Longley, Eds.). Troubador Publishing Ltd.

de Solla Price, D. J. 1963. Little Science , Big Science ... and Beyond. Columbia University Press, New York.

Spangenberg, J. H., C. Görg, D. T. Truong, V. Tekken, J. V. Bustamante, and J. Settele. 2014. Provision of ecosystem services is determined by human agency, not ecosystem functions. Four case studies. International Journal of Biodiversity Science, Ecosystem Services & Management 10:40–53.

Sparks, T. H., K. Huber, and P. Tryjanowski. 2008. Something for the weekend? Examining the bias in avian phenological recording. International journal of biometeorology 52:505–10.

Star, S. L., and J. R. Griesemer. 1989. Institutional Ecology , “Translations” and

Boundary Objects : Amateurs and Professionals in Berkeley â€™ s Museum of Vertebrate Zoology,1907-39. Social Studies of Science 19:387–420.

Steiger, E., R. Westerholt, and A. Zipf. 2016. Research on social media feeds – A GIScience perspective. Pages 237–254*in* C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, and R. Purves, editors.European Handbook of Crowdsourced Geographic Information. Ubiquity P. Ubiquity Press License: CC-BY 4.0., London, UK.

Stein, K., and S. Blaschke. 2010. Interlocking communication measuring collaborative intensity in social networks. From Sociology to Computing in Social Networks: Theory, Foundations and Applications:231–252.

Stein, K., D. Kremer, and C. Schlieder. 2015. Spatial Collaboration Networks of OpenStreetMap. Pages 167–186*in* J. J. Arsanjani, P. Mooney, A. Zipf, and M. Helbich, editors.OpenStreetMap in GIScience. Lecture No. Springer, Heidelberg, New York, Dordrecht, London.

Steinmann, R., E. Häusler, S. Klettner, M. Schmidt, and Y. Lin. 2013. Gender Dimensions in UGC and VGI: A Desk-Based Study. Proceeding of GI_Forum 2013 – Creating the GISociety:355–364.

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

Suh, B., G. Convertino, E. H. Chi, and P. Pirolli. 2009. The singularity is not near. Proceedings of the 5th International Symposium on Wikis and Open Collaboration - WikiSym '09 1:1.

Sui, D., S. Elwood, and M. Goodchild, editors. 2013a. Crowdsourcing geographic Knowledge: Volunteered geographic information (VGI) in theory and practice. Springer Netherlands, Dordrecht.

Sui, D., S. Elwood, and M. Goodchild, editors. 2013b. Crowdsourcing Geographic Knowledge. Springer Netherlands, Dordrecht.

Surmacki, A. 2005. What do data from birdwatchers notepads tell us? The case of the Bearded Tit (*Panurus biarmicus*) occurrence in western Poland. Ring 27:79–85.

Sutherland, W. J., editor. 1996. Ecological census techniques: a handbook. Cambridge University Press.

Swanson, B. 2007. The coming exaflood. The Wall Street Journal.

Tapscott, D., and A. D. Williams. 2008. Wikinomics: How Mass Collaboration Changes Everything. Page Journal of Communication. Portfolio.

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

Tenerelli, P., C. Püffel, S. Luque, C. Puffel, and S. Luque. 2017. Spatial assessment of aesthetic services in a complex mountain region: combining visual landscape properties with crowdsourced geographic information. Landscape Ecology 32:1097–1115.

The Antwerp Declaration. 2016. . http://antwerpdeclaration.info/.

The Economist. 2003, March. The revenge of geography. London. http://www.economist.com/node/1620794.

The Economist. 2010. The data deluge - Businesses, governments and society are only starting to tap its vast potential.

The Research Box. 2009. Experiencing landscapes: capturing the "cultural services" and "experiential qualities" of landscape. Page Report NECR024. Natural England.

Tieskens, K., N. Schulp, P. Verburg, and K. Tobias. 2014. Typology of cultural landscapes. Page HERCULES Sustainable futures for Europe's HERitage in CULtural landscapES: Tools for understanding, managing, and protecting landscape functions and values. HERCULES.

Tulloch, A. I. T., K. Mustin, H. P. Possingham, J. K. Szabo, and K. a. Wilson. 2012. To boldly go where no volunteer has gone before: Predicting volunteer activity to prioritize surveys at the landscape scale. Diversity and Distributions 19:465–480.

Tulloch, A. I. T., and J. K. Szabo. 2012. A behavioural ecology approach to understand volunteer surveying for citizen science datasets. Emu 112:313–325.

Turner, A. J. 2006. Introduction to Neogeography. Page OReilly Short Cuts. O'Reilly Media.

UK National Ecosystem Assessment. 2011a. The UK National Ecosystem Assessment Technical Report. Cambridge.

UK National Ecosystem Assessment. 2011b. Glossary. Pages 1452–1466UK National Ecosystem Assessment Technical Report.

UK National Ecosystem Assessment. 2014. UK National Ecosystem Assessment Follow-on: Synthesis of the Key Findings.

UN. 1993. Agenda 21: Earth Summit - The United Nations Programme of Action from Rio. United Nations.

United Cities and Local Governments - Committee on culture. 2008. Agenda 21 for culture. United Cities and Local Governments - Committee on culture, Barcelona.

United Kingdom. 2018. Data Protection Act 2018.

United Nations. 2015. The Millennium Development Goals Report. United Nations:72.

Upton, V., M. Ryan, C. O'Donoghue, and A. N. Dhubhain. 2015. Combining conventional and volunteered geographic information to identify and model forest recreational resources. Applied Geography 60:69–76.

Viégas, F. B., M. Wattenberg, J. Kriss, and F. Van Ham. 2007. Talk before you type:

Coordination in Wikipedia. Proceedings of the Annual Hawaii International Conference on System Sciences:1–10.

Vitolo, C., Y. Elkhatib, D. Reusser, C. J. A. Macleod, and W. Buytaert. 2015. Web technologies for environmental Big Data. Environmental Modelling & Software 63:185–198.

Vorstius, A. C., and C. J. Spray. 2015. A comparison of ecosystem services mapping tools for their potential to support planning and decision-making on a local scale. Ecosystem Services 15:75–83.

Vander Wal, T. 2007. Folksonomy Coinage and Definition. http://vanderwal.net/folksonomy.html.

Walden-Schreiner, C., Y.-F. Leung, and L. Tateosian. 2018. Digital footprints: Incorporating crowdsourced geographic information for protected area management. Applied Geography 90:44–54.

WCED. 1987. Our common future. World Commission on Environment and Development. Oxford University Press.

Westman, W. E. 1977. How Much Are Nature's Services Worth? Science 197:960 LP-964.

Wiggins, A. 2012. Crowdsourcing Scientific Work: A Comparative Study of Technologies , Processes , and Outcomes in Citizen Science. Syracuse University.

Wikipedia contributors. 2011. Verifiability, not truth. https://en.wikipedia.org/wiki/Wikipedia:Verifiability,_not_truth.

Wikipedia contributors. 2017a. Verifiability. https://en.wikipedia.org/w/index.php?title=Wikipedia:Verifiability&old id=701166094.

Wikipedia contributors. 2017b. WikiProject OpenStreetMap. https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Biophysics.

Wikipedia contributors. 2017c. OpenStreetMap. https://en.wikipedia.org/wiki/OpenStreetMap.

Wikipedia contributors. 2017d. How to import articles. https://en.wikipedia.org/wiki/Wikipedia:How_to_import_articles.

Wilkinson, D. M., and B. a Huberman. 2007. Cooperation and quality in wikipedia. Proceedings of the 2007 international symposium on Wikis WikiSym 07:157–164.

Willemen, L., A. Veldkamp, P. H. Verburg, L. Hein, and R. Leemans. 2012. A multi-scale modelling approach for analysing landscape service dynamics. Journal of Environmental Management 100:86–95.

William J. Hughes Technical Center. 2016. Global positioning system (GPS) standard positioning service (SPS) performance analysis report. Tech. Cntr. NSTB/WAAS T and E Team:147.

Wilson, K. A., and E. A. Law. 2016. How to Avoid Underselling Biodiversity with Ecosystem Services: A Response to Silvertown. Trends in Ecology and Evolution 31:332–333.

Wilson, M. W., and M. Graham. 2013. Guest editorial: Situating neogeography. Environment and Planning A 45:3–9.

Wilson, R. E., S. D. Gosling, and L. T. Graham. 2012. A Review of Facebook Research in the Social Sciences. Perspectives on Psychological Science 7:203–220.

Wood, S. a, A. D. Guerry, J. M. Silver, and M. Lacayo. 2013. Using social media to quantify nature-based tourism and recreation. Scientific reports 3:2976.

Woodward, G., A. J. Dumbrell, D. J. Baird, and M. Hajibabaei, editors. 2014. Big Data in Ecology. ADVANCES IN ECOLOGICAL RESEARCH 51:2–121.

Yang, A., H. Fan, N. Jing, Y. Sun, and A. Zipf. 2015. Temporal Analysis on Contribution Inequality in OpenStreetMap : A Comparative Study for Four Countries. ISPRS International Joiurnal of Geo-Information 5.

Yasseri, T., A. Spoerri, M. Graham, and J. Kertész. 2014. The most controversial topics in Wikipedia: A multilingual and geographical analysis. Global Wikipedia: International and cross-cultural issues in collaboration:178.

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

van Zanten, B. T., D. B. van Berkel, R. K. Meetemeyer, J. W. Smith, K. F. Tieskens, and P. H. Vergurg. 2016. Continental scale quatification of landscape values using social media data. Proceedings of the National Academy of Sciences:1–7.

Zielstra, D., and A. Zipf. 2010. A Comparative Study of Proprietary Geodata and Volunteered Geographic Information for Germany. Page 13th AGILE International Conference on Geographic Information Science 2010. Guimarães, Portugal.