Received: 29 July 2019 | Accepted: 20 January 2020
DOI: 10.1002/pan3.10083

RESEARCH ARTICLE

PEOPLE AND NATURE — BRITISH ECOLOGICAL SOCIETY

Using crowdsourced spatial data from Flickr vs. PPGIS for understanding nature's contribution to people in Southern Norway

Lorena Muñoz[1] | **Vera Helene Hausner**[1] | **Claire Runge**[1] | **Greg Brown**[2] | **Remi Daigle**[3]

[1]Department of Arctic and Marine Biology, UiT The Arctic University of Norway, Tromsø, Norway

[2]Department of Natural Resource Management and Environmental Sciences, California Polytechnic State University, San Luis Obispo, CA, USA

[3]Department of Oceanography, Dalhousie University, Halifax, NS, Canada

Correspondence
Lorena Muñoz
Email: lorena.munoz@uit.no

Funding information
Norges Forskningsråd, Grant/Award Number: 230330/E50/2014; CultES

Handling Editor: Laura Graham

Abstract

1. Crowdsourced data can provide spatially explicit data on the contribution of nature to people. Spatial information is essential for effectively managing the diverse relationships that people have with nature, but the potential and limits of using crowdsourcing data to generate maps for conservation purposes need further research.
2. Passive crowdsourcing tools include social media platforms where photos and user-generated tags are shared among users, whereas active crowdsourcing, such as public participatory geographic information system (PPGIS), provides an online platform for mapping place attributes such as values, experiences and preferences.
3. In this study, we assess the spatial information gained through using Flickr (a photo sharing platform) and PPGIS (an online mapping platform) platforms for conservation planning to understand differences and similarities on the spatial distribution of values captured by the two platforms, and to identify what environmental and infrastructure variables correlate best with the distribution of values. We test these tools in Southern Norway including protected areas and the surrounding zones.
4. We analysed non-spatial (using chi-square and Spearman rank correlation) and spatial (using clustering, Maxent and distribution overlap) data to identify differences between the two datasets and the values represented therein.
5. We found large differences in spatial distribution using these two datasets, with Flickr data concentrated outside the protected areas and near roads, whereas PPGIS provided more fine-scale data on diverse values in locations inaccessible by roads within the protected areas. Flickr can be used for generating regional scale data of scenic landscapes or routes, but PPGIS performs better for management of nature qualities appreciated by different user groups within protected areas. We discuss the pros and cons of using each data source and when each dataset is more suitable to be used in protected area management.

KEYWORDS
cluster analysis, management, maxent, nature qualities, protected area, social media, values, visitors

1 | INTRODUCTION

As anthropogenic pressures on nature increase across the globe, raising awareness of nature's contribution to people (NCP) has become one of the approaches for integrating conservation into policy (Pascual et al., 2017). Despite the growing body of research on the non-material contribution of nature to a good quality of life (Hirons, Comberti, & Dunford, 2016), tools for mainstreaming non-material contributions into ecosystem services assessments and decision-making are still under development (Costanza et al., 2017; Small, Munday, & Durance, 2017). The natural processes and features appreciated by people that positively contribute to their life are often referred to as nature qualities (Arler, 2000; Van den Bosch, Östergren, Grahn, Skärbäck, & Währborg, 2015) and are a central component of NCP. Bringing in diverse perspectives and values into conservation planning is costly, time-consuming and logistically challenging, but is important to find solutions that balance the needs of people with conservation objectives.

A wide range of methods and approaches have been used to elucidate the diverse perspectives on the cultural benefits provided by nature (Small et al., 2017; Teff-Seker & Orenstein, 2019; Tew, Simmons, & Sutherland, 2019). Among these are crowdsourcing methods which have the potential to deliver spatial information of NCP from a diverse range of citizens at a large scale of relevance to conservation (Bubalo, van Zanten, & Verburg, 2019). There are two main crowdsourcing approaches that have gained popularity in recent years: passive and active crowdsourcing. Passive crowdsourcing derives data from users leaving traces online on location and activity by sharing material on social media or by simply using their cell phones (Birenboim & Shoval, 2016; See et al., 2016). Social media derived from people sharing text or photos on an online platform, such as Flickr, has become particularly important for mapping recreation and aesthetic values appreciated by people in nature (Richards & Friess, 2015; van Zanten et al., 2016). Combining several content sharing platforms has been suggested for monitoring protected area popularity and temporal visitation patterns, using, for example, Instagram, Twitter and Flickr (Tenkanen et al., 2017). Active crowdsourcing, on the other hand, depends on users actively contributing with data through online platforms specifically designed to collect data about users or nature qualities (Ridding et al., 2018; Wolf, Brown, & Wohlfart, 2018). Data collection through online platforms could either openly recruit anyone to participate (volunteered geographic information [VGI]), or it could be based on targeted sampling of individuals to ensure representation of the population of interest (e.g. public participation geographic information system [PPGIS]; Brown, Kelly, & Whitall, 2014).

Although social media and online PPGIS platforms have both been shown to be useful tools for assessing the spatial distribution of values, each has their pros and cons. Social media data are less costly to collect and therefore allow the elicitation of values from a much larger pool of potential users on a broader scale (Toivonen et al., 2019). Social media data have been used to quantify nature-based tourism and recreation (Wood, Guerry, Silver, & Lacayo, 2013), tourism flows (Hawelka et al., 2014) or for mapping destinations and events that are highly visited by the public (Kisilevich, Keim, Andrienko, & Andrienko, 2013). The tags can also inform about how people value nature, how those values are distributed and the contribution of nature to the qualities appreciated by people (van Zanten et al., 2016). The photos can represent diverse activities and values including aesthetics, recreation, wildlife viewing and bio-cultural heritage (Toivonen et al., 2019). Moreover, photos taken by several people at a specific location can be associated with specific environmental characteristics of that area (Dunkel, 2015). Content analysis of photographs shared on social media has also been used to model the spatial distribution of values and non-material benefits with respect to landscape characteristics and infrastructure, and to indicate how changes in the landscape and infrastructure development can affect the overall visitor experience and distribution (Tenerelli, Demšar, & Luque, 2016; Walden-Schreiner, Leung, & Tateosian, 2018). However, social media have been shown to be unreliable at capturing some indirect-use and non-use values, whereas PPGIS is capable of capturing a wide range of values (Levin, Lechner, & Brown, 2017). The primary benefit of PPGIS surveys is the possibility to customize the tool to collect information on spatial values, preferences and experiences that are of direct relevance to protected area management (e.g. Brown & Weber, 2011). For example, PPGIS has been used to identify areas of value hotspots and the overlap of different user groups, to understand land use preferences, to address conflicts between different user groups, and to monitor tourism development preferences (Brown & Weber, 2013; Engen et al., 2018; Muñoz, Hausner, Brown, Runge, & Fauchald, 2019; Wolf et al., 2018). Participatory mapping surveys are customized for each case, which makes them suitable for surveying a wide range of people, which can include stakeholders, locals, visitors, experts, the general public and decision-makers (Brown & Kyttä, 2014). Thus, PPGIS can include voluntary participation (similar to social media), as well as targeted recruitment of a representative sample.

While the use of social media data has been compared to visitor data on a regional scale previously (Graham & Eigenbrod, 2019; Tenkanen et al., 2017), spatial data and the values identified using passive and active crowdsourcing tools have not been extensively evaluated using the same location. One exception is Levin et al. (2017) who compared the visitor density and values mapped by crowdsourcing tools in multiple protected areas. No one has to date compared the potential of active and passive crowdsourcing tools to provide spatial information of nature qualities on a finer scale of relevance to protected area management (i.e. within protected areas). The spatial distribution at this scale will depend on the profile of users captured by the different tools, the values people ascribe to nature and the spatial accuracy of the geolocations mapped using different platforms. If these tools are to be used to guide protected area management, it is important to understand the conditions that influence the results generated by each tool at this scale.

Here we examine the spatial distribution and the type of values generated by the two crowdsourced tools (Flickr and PPGIS), and their usefulness for informing protected area managers about the nature qualities that are important for different groups of people. We tested the crowdsourcing tools with respect to how they perform in capturing spatial information of the nature qualities that people care about in an iconic mountainous landscape in Norway, encompassing a cluster of protected areas that are visited by different domestic and international visitors. Our study differs from previous comparisons of Flickr data to visitor data by (a) the explicit focus on spatial information of relevance to protected area management, (b) the comparison of the values derived from using these two crowdsourcing tools and (c) their relationship to the locational profile of the Flickr/PPGIS users and environmental and infrastructure characteristics. We asked (a) Does the spatial distribution of values generated by Flickr versus PPGIS data differ? (b) How does the distribution of values using these two tools correlate with environmental and infrastructure variables? (c) How much do values overlap using Flickr versus PPGIS? and (d) Do international and domestic visitors map different attributes using the two tools? Finally, we discuss the pros and cons of using these tools for assessing NCP to inform protected area management.

2 | MATERIALS AND METHODS

2.1 | Study area

This study was conducted in Southern Norway and included Jotunheimen National Park (Jotunheimen NP), Breheimen National Park (Breheimen NP), Utladalen Protected Landscape (Utladalen PL) and the non-protected area surrounding these areas (Figure 1). Jotunheimen NP and Breheimen NP were originally designated for their wilderness and untouched nature covering 1,151 and 1,691 km^2, respectively, and have become major nature attractions in Norway. They are dominated by alpine vegetation and hold the highest peaks in Scandinavia and several glaciers and lakes. In 1980, at the same time as Jotunheimen was designated national park, the neighbouring area Utladalen was declared protected landscape with the aim to protect cultural landscapes (Ministry of Climate & Environment, 2014). The major difference between national parks and protected landscapes is the uses allowed. National parks are mainly designated to protect ecosystems and biological diversity allowing low levels of human use, whereas protected landscapes aim at conserving natural and cultural landscapes with high ecological and cultural values, and traditional use is an inherent objective for protection (Ministry of Climate & Environment, 2019). The study area has for a long time been used for traditional outdoor recreation, attracting visitors for the cabin-to-cabin hikes and climbing opportunities. The study area also includes several villages (e.g. Øvre Årdal, Beitostølen and Lom), which host a variety of cultural and recreational activities all year round, both in and around protected areas, such as music, film and food festivals, and guided tours by foot, bike, horse, dog-sledding or rafting (Jotunheimen, 2019).

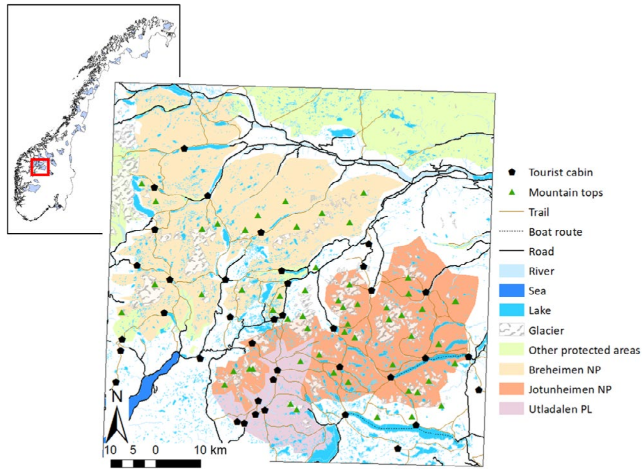

FIGURE 1 Map of the study area. Top left: Map of Norway with national parks shaded in blue and a red box indicating the study area location. Bottom left: Our study area with protected areas shaded in colour. The main touristic attractions are shown in the map (mountain tops, touristic cabins and glaciers)

2.2 | Data acquisition

2.2.1 | Flickr

Flickr is a free photo management and sharing platform where users can upload their pictures, geotag them and share them privately or publicly (Flickr, 2019). We retrieved information associated with 6,255 publicly available geotagged Flickr images on 4 April 2016 for our study area using the FLICKRGEOTAG R package (Daigle & Dunnington, 2018). The metadata that accompanied the images included de-identified (key-coded) photos and user ID codes, the country of origin of the Flickr user, text-based tags associated with each photo (which can be either user-specified or selected by Flickr's automated tagging algorithm), the coordinates (latitude and longitude in WGS84) of the image and the URL link to the photo. For the purpose of this study, we used the country of origin, the coordinates and the photo URL. For those users that did not report their country of origin (268 users), we estimated the contributors' home country from the median coordinates of all uploaded pictures. The photographs were taken between 2007 and 2016 with 34% of the images dated 2014–2015. Although social media data can identify changes in visitation from year to year (Tenkanen et al., 2017), we aggregated the data from 9 years for this study as Flickr data are temporally sparse in this region so that we could ensure sufficient sample to make robust conclusions. Also, we were not focusing on the changes over time in this study, but values that change more slowly (see Brown & Weber, 2012).

A detailed list of values was developed by five experts who had previous experience with Flickr and the case study area. We used the CICES V4.3 framework (Haines-Young & Potschin, 2013) to identify categories to code. A detailed definition of each value was discussed and agreed between the experts to avoid overlap between categories (see List 1 for the full list of values coded). Codes were trialled iteratively until agreeing upon a final list of values that could be extracted from the pictures. We knew from previous studies that recreation and scenic nature are the primary reasons for visiting protected areas (Levin et al., 2017; Muñoz et al., 2019), and biological diversity, wilderness and learning are the traditional international objectives of protecting land. Protected mountainous landscapes in Southern Norway also include traditional uses associated with historical land tenures (Hausner, Brown, & Lægreid, 2015). We therefore included harvesting, livelihood, social and heritage values relating to nature as possible qualities that visitors may appreciate.

Each value reflects the primary subject of the photograph. After the coding system had been developed, the content of each picture was manually examined and coded (by the author L.M.). We assigned one code to each picture based on the dominant feature of the picture, which could show, for example, an activity, wildlife or a landscape. After the content analysis, only the pictures taken in a natural setting were retained (4,038 photos) and those showing portraits, built environments and extractive activities were discarded from further analyses. From pictures taken in natural settings, four values had a similar definition and sufficient number of photos to compare with the PPGIS dataset (values 1–4 in List 1).

List 1 *The values used for coding Flickr photos adjusted from the CICES V4.3 framework. Only values 1–4 were used for comparisons with the values mapped in PPGIS*

1. *Biological diversity*: Dominant feature of the picture is plants, animals or other important ecosystem features. For example, pictures of wild animals or plants.
2. *Recreation*: Dominant feature of the picture is people doing physical recreational activities. For example, walking, hiking, climbing, boating.
3. *Scenic landscapes*: Dominant feature of the picture is an important place that is scenic, a distinctive landscape, wilderness or natural settings (could include people, but not the main focus). For example, scenic drives, scenic cruises, mountains, fjord, wilderness. Could be symbolic/spiritual values, which need to be determined ad hoc.
4. *Social*: Pictures taken primarily of social activities in natural setting, including organized activities. For example, alpine arrangements, bonfires, picnic.
5. *Harvest*: Dominant feature of the picture is people engaging in recreational harvest. For example, leisure hunting, fishing, picking berries, etc.
6. *Heritage*: Dominant feature of the picture is related to historical use of nature. For example, evidence of historical fishing and hunting, summer farms, etc.
7. *Learning*: Dominant feature of the picture is scientific or educational activities in nature or related to natural features. For example, school trips, field research, etc.
8. *Livelihood*: Dominant feature of the picture is related to local livelihoods/economy. For example, sheep farming, reindeer herding.

2.2.2 | Public participation geographic information system

The PPGIS is a GIS tool to map spatial attributes and important locations in an area. We conducted two online PPGIS surveys: a household survey combined with voluntary participation of locals, and a visitor survey with in-situ recruitment in the study area in October–December 2014 and July–September 2015, respectively. For the first survey, we invited a randomly selected set of 10% of the households in the municipalities in the study area to participate in the web-PPGIS study, contacting them by regular post. A reminder letter was sent 2 weeks after the first contact. Additionally, we used local organizations, newspaper and social media to recruit volunteers. During the peak visitor season to our study area in 2015, we recruited respondents to the second survey at recreational parking spots, either by direct contact or by leaflets placed on cars. Two reminders were sent by email to visitors recruited in the field.

In the PPGIS survey, we asked respondents to drag and drop georeferenced markers that represent one of the 12 values (see List 2 for the full list of values) onto a Google® map view, by zooming in and out as needed. People could place as many markers as they wanted, but were encouraged to place at least 20. They were free to place markers for as many, or as few, values as they wished. We refer to 'mapped value' as the georeferenced marker placed by participants on the map. We piloted the surveys on park managers whose feedback was used to improve the PPGIS platform. The Data Protection Official for Research for all the Norwegian universities and research institutes (Norwegian Centre for Research Data) approved the ethical treatment of the data in the project (CultES no. 230330/E50/2014) under the Personal Data Act 2000. The online survey included an informed consent for participation that respondents had to accept before completing the survey, where we informed participants about the purpose of the study and explained that data would be treated confidentially. Also, participants were informed that the study was voluntary, and that they could withdraw from it at any time or contact us through the provided email in case of any concerns regarding the study. For additional details about the survey, see Muñoz et al. (2019).

From the 12 values included in the mapping activity, four were comparable to the categories obtained by coding Flickr images: biological diversity, scenic landscapes, social value and recreation (values 1–4 in List 2). We used all values mapped in Flickr and PPGIS to identify the potential differences between international and domestic visitors for each platform (i.e. difference in clustering and ranking between user groups). We used the subset of four values that were comparable for PPGIS and the Flickr coding (see above) to compare the difference in spatial information obtained from these two platforms. When discussing results, we refer to either 'all values' (8 values in Flickr and 12 values in PPGIS) or 'four common values' (i.e. the ones that are comparable between the two datasets).

List 2 *The values used in the PPGIS survey adapted from Brown and Reed (2000) to the Norwegian context* (*Hausner et al., 2015*). *Only values 1–4 were used in direct comparisons with Flickr values*

1. *Biological diversity*: Areas that are important because they provide a variety of plants, wildlife and habitat.
2. *Recreation*: Areas that are important for outdoor recreation activities (e.g. camping, walking, skiing, alpine snowmobiling, cycling, horse riding).
3. *Scenic landscapes*: Areas that are important because they include beautiful nature and/or landscapes.
4. *Social value*: Areas that are important because they provide opportunities for social activities (e.g. associated with fireplaces, picnic tables, ski- or alpine arrangements, shelters, shared cabins, cabin complexes).
5. *Clean water/air*: Areas that are important because they provide clean water/air.
6. *Cultural value (including cultural identity)*: Areas that are important because of their historical value, or for passing down the stories, myths, knowledge and traditions, and/or to increase understanding of the way of life of our ancestors.
7. *Gathering*: Areas that are important for berries, mushroom or collecting herbs/plants.
8. *Hunting/fishing*: Areas that are important because of hunting and/or fishing.
9. *Spiritual value*: Areas that are important because they are valuable in their own right or have a deeper meaning; emotionally, spiritually or religious.
10. *Therapeutic*: Areas that are valuable because they make me feel better, either because they provide opportunities for physically activities important for my health and/or they give me peace, harmony and therapy.
11. *Wilderness and undisturbed nature*: Areas that are relatively untouched, providing for peace and quiet without too many disturbances.
12. *Special place*: Please describe why these places are special to you.

2.3 | Statistical analyses

2.3.1 | Density-based clustering for hotspot mapping

We conducted a density-based cluster analysis of all the values mapped to compare the areas with highest density of values (hotspots) in each dataset and to quantify the number of hotspots. To accomplish this, we used the 'density-based spatial clustering of applications with noise' (DBSCAN) algorithm (Ester, Kriegel, Sander, & Xiaowei, 1996) with a minimum of 10 neighbouring points within a 1,000 m search radius. In DBSCAN, points represent the geographical location of each Flickr photo or the mapped value location in PPGIS. This algorithm detects points that form clusters with irregular shapes and discards sparse points (Ester et al., 1996). The search radius was determined by visual inspection of the threshold of the k-nearest neighbour distances plot. DBSCAN forms clusters with core and border points. Core points are those that are surrounded by 10 points within the search radius. Ten points was selected as the minimum number of points to capture a diversity of values inside each cluster. Border points are those points that belong to a cluster because they are located inside the search radius of a core point, but do not have the requirements to be classified as a core point (i.e. they do not meet the requirement of a minimum 10 points in a 1,000 m search radius). The points that are not classified as either core or border points are discarded from the clusters. The resulting clusters are point clouds containing core and border points.

2.3.2 | Maximum entropy modelling for environmental and infrastructure variables

The purpose of the modelling was to test whether Flickr and PPGIS data are correlated with the same environmental and infrastructure characteristics. We developed the following 18 models to analyse the distribution of values: two overall models for all values in each dataset separately (i.e. Flickr and PPGIS), and 16 models for each unique combination of the four common values (the first four values in List 1 and List 2, we compared each domestic and international user group (n = 2), developed for each dataset). We selected the covariates based on previous research demonstrating how nature tourism is related to human infrastructure and environmental characteristics (Bagstad, Semmens, Ancona, & Sherrouse, 2016; Richards & Tunçer, 2018; Walden-Schreiner et al., 2018). Values were modelled against nine environmental and infrastructure variables (hereafter referred to as covariates); eight continuous variables: distance from trails, roads, touristic cabins, buildings (other infrastructures, e.g. houses, bridges), rivers, lakes, and mountain tops and glaciers and vegetation cover percentage; and one categorical variable: altitude (divided in 500-m elevation intervals; see Supporting Information Table S1). We extracted covariates from the N500 database developed by the Norwegian Mapping Authority (Kartverket), which contains among other things landscape characteristics and infrastructure (Kartverket, 2015). Mountain tops were manually georeferenced based on the protected area brochures published by the Norwegian Environmental Agency. Vegetation cover percentage was produced from CORINE2006 data (European Environmental Agency, 2015) and transformed to vegetation cover percentage. We reclassified the CORINE map by assigning 100% cover to vegetated areas, 50% cover to areas sparsely vegetated and 0% cover to areas artificial surfaces, rocks, non-vegetated areas and water bodies. The values for each pixel were interpolated using the nearest neighbour approach using a 3 × 3 kernel. We rasterized covariates in a 1,210,000 pixel raster with a 116.1 m pixel size. The raster layers provided distances to natural and human-made features and these were square root transformed to avoid skewedness towards the right end (long distances). We tested for correlation between covariates and found no indication to discard any of the covariates (Supporting Information Table S2).

We developed the 18 maximum entropy models using MaxEnt software version 3.4.0 (Phillips, Dudík, & Schapire [Internet]). Briefly, maximum entropy modelling compares the distribution of presences (e.g. sighting of a species) in environmental space (the set of covariates) against the background distribution of those covariates (Elith et al., 2011). The model compares the presence of points (i.e. values) against a set of randomly distributed background points to estimate the influence of environmental characteristics on the value distribution. Therefore, we removed duplicates from the model as MaxEnt works with presence data and 25% of the presence points were randomly selected as a test set during the internal validation of the model. We selected a random subset of 10,000 background points from the 1,210,000 grid cells in our study region. MaxEnt selected the regularization values and feature types, that is, hinge, product, linear and quadratic, that was best fit to the model. The output is a model that can predict the suitability of other areas for the values mapped by users. To identify those covariates that best explain the distribution of each value, we examined the permutation importance, which is a measure calculated by randomly selecting values for each of the covariates for each permutation during the training of the model, independent of the model path followed. The permutation importance measures how much the model relies on the given variable, normalized to percentages. In other words, the permutation importance is a measure of the contribution of a variable to the predictive ability of the model. We used these models to predict the suitability of the study area to contain the four common values. To assess how alike the predictions were for values mapped in different platforms (i.e. Flickr and PPGIS), we used the suitability maps for each value to calculate the niche overlap between the two datasets. MaxEnt is suitable for use with presence-only data such as that generated by Flickr or PPGIS, where the photo or PPGIS locations indicate the 'presence' of a value, but unmapped areas cannot be assumed to indicate the 'absence' of a value. Maximum entropy modelling has previously been used to model species distribution (Phillips & Dudík, 2008), but it is increasingly used for modelling ecosystem services and visitor distribution (Bagstad et al., 2016; Walden-Schreiner et al., 2018).

All analyses were conducted using the R Software version 3.4.1 (R Core Team, 2019) using 'DISMO' package for the Maxent model (Hijmans, Phillips, Leathwick, & Elith, 2017) and 'DBSCAN' package for the DBSCAN algorithm (Hahsler, Piekenbrock, Arya, & Mount, 2018).

2.3.3 | Comparing domestic and international visitors

We used exploratory analyses to describe and summarize differences in the mapped values by different user groups for each citizen-generated dataset. First, we identified differences between the values mapped by domestic and international visitors within each dataset using chi-square tests and then Spearman rank correlation tests. For each mapped common value, we compared standardized chi-square residuals of the proportion of values mapped compared to the total amount of values for domestic and international visitors, identifying those values that were outside the range −2 to 2 as being mapped significantly less or more often than the other group. We used the Spearman rank correlation to show the degree to which the two user groups (i.e. domestic and international visitors) are similar in their perception of value importance based on the ranks of mapped value frequencies (based on 8 data points in Flickr and 12 data points in PPGIS).

3 | RESULTS

In the Flickr dataset, 479 users geotagged a picture related to nature qualities inside the study area, from which 177 were domestic (Norwegians), 284 were international visitors and 18 had an unknown origin. Of the 479 users, 268 users did not report their origin. Using the median distance of all the photos that each of these individuals uploaded, we concluded that 100 were domestic visitors, 150 international visitors and 18 remained with no clear origin. From the 4,038 uploaded images, photos related to nature qualities primarily showed scenic landscapes (3,008 photos) and recreation (601). The median number of nature related photos uploaded by each user was 2, and 16 users uploaded more than 50 pictures inside the study area (3.3% of users; Supporting Information Figure S1). In the PPGIS dataset, 468 respondents were recruited, split between 332 domestic (Norwegians) and 136 international visitors. From 3,873 mapped values, the most commonly mapped value was recreation (1,176 markers) followed by scenic landscapes (1,070). The median number of mapped values by each user was 5, and five users (1%) were identified as 'supermappers' (those who mapped more than 50 values; Figure S1).

We tested differences in the spatial distribution of all values for the two datasets by creating density-based clusters to identify hotspots of values. The density cluster analysis resulted in 51 hotspots for the Flickr database and 36 hotspots for the PPGIS database (Figure 2) with 19.7% and 35.9% of the points remaining outside clusters. Figure 2 shows that places attractive to visitors are located along roads in the Flickr dataset, but are predominantly located inside protected areas in the PPGIS dataset (values inside PAs: 32.3% in Flickr and 77.4% in PPGIS).

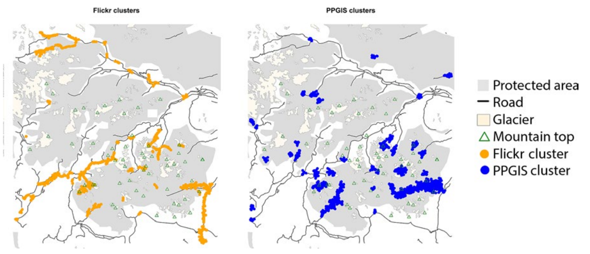

FIGURE 2 Clusters from the density-based clustering for Flickr (orange, left) and PPGIS (blue, right)

We compared 18 MaxEnt models to determine differences in the two datasets concerning the environmental and infrastructure covariates that explain the distribution of values. We used the permutation importance metric to understand the contribution of each covariate to the MaxEnt model, which contrary to the per cent contribution, does not depend on the order in which the covariates are entered into the model (Kalle, Ramesh, Qureshi, & Sankar, 2013). The MaxEnt models for all values in each dataset indicate that the location of values in Flickr was mainly explained by distance to motorized access, while the location of values in the PPGIS dataset was determined primarily by distance to mountain tops, glaciers and trails (Table 1). We further examined Maxent models for each comparable value, which confirmed that the values are explained by different environmental and infrastructure

TABLE 1 Permutation importance expressed in percentage on how much each model relies on each variable. Shaded numbers indicate the landscape or infrastructure covariates with the highest permutation importance percentage. We calculated the permutation importance (percentage of how much each variable contributes to the model) for all values in each dataset, and for the four comparable values for domestic and international visitors

	Motorized access distance	Building distance	Cabin distance	Lake distance	River distance	Topography	Mountain top/glacier distance	Trail distance	Vegetation
Flickr									
Overall	49.4	0.3	2.9	3.2	0.8	1.4	20.3	19.6	2.3
Biological diversity									
Domestic	66.1	2.3	1.7	2.9	1.4	0.4	4.9	13.4	6.9
International	55.5	2.1	8.5	0.0	0.1	3.0	6.6	12.3	11.9
Recreation value									
Domestic	26.4	0.3	2.8	4.0	3.7	0.1	43.6	17.7	1.4
International	47.9	2.7	1.7	1.4	2.1	2.3	15.5	20.6	5.9
Scenic value									
Domestic	36.2	0.9	2.5	7.3	1.4	2.6	27.5	19.5	2.0
International	56.9	0.3	3.0	1.3	0.6	0.2	14.2	21.0	2.4
Social value									
Domestic	0.0	0.0	36.2	6.7	0.0	0.7	51.4	4.7	0.3
International	17.0	1.5	4.8	8.0	0.1	7.3	22.5	38.3	0.5
PPGIS									
Overall	6.5	2.2	4.0	10.6	5.2	1.6	34.3	33.7	1.7
Biological diversity									
Domestic	4.1	6.6	17.6	1.1	7.2	3.7	42.1	9.9	7.8
International	0.0	1.8	10.3	0.7	3.2	1.6	35.3	44.1	3.0
Recreation value									
Domestic	7.0	1.9	6.5	11.5	5.6	0.3	38.0	28.9	0.3
International	7.9	5.9	7.5	4.4	4.3	4.0	26.4	38.5	1.1
Scenic value									
Domestic	8.6	2.1	2.5	7.7	6.1	2.3	34.3	33.7	2.7
International	11.7	1.4	8.9	3.5	5.5	1.5	22.4	41.9	3.1
Social value									
Domestic	14.3	2.5	4.8	4.9	1.7	11.5	20.5	25.8	13.9
International	0.0	0.7	22.6	0.0	0.1	0.4	4.6	71.5	0.0

TABLE 2 Results for the overlap of predictions for Flickr and PPGIS datasets resulting from MaxEnt models. Values range from 0 (no overlap) to 1 (identical distribution). The last row contains the overlap between the MaxEnt habitat suitability prediction for all values in Flickr and PPGIS for all users together

	Domestic	International
Biological	0.68	0.80
Recreation	0.89	0.83
Scenic	0.94	0.83
Social	0.83	0.64
Flickr versus PPGIS (model for all values and users together)	0.94	

covariates in each dataset. It also showed that domestic and international visitors correlate differently to covariates in the PPGIS dataset. According to the permutation importance metric (Table 1), the location of values found in the Flickr dataset was heavily influenced by distance to motorized access, with three exceptions: domestic-visitors-related recreation and social values to mountain tops and glaciers, and international-visitors-related social values to trails. Values in the PPGIS dataset differed from these results as they were less influenced by infrastructure and more by proximity to mountain tops, glaciers and trails. Domestic visitors mapped values closer to mountain tops and glaciers, with the exception of social values which were mostly related to trails. Most values mapped by international visitors were related to distance from trails.

Most of the Flickr values (60%) were within the first 500 m from roads, compared with 23% in the PPGIS dataset. For trails, only 34% of values from the Flickr data were within 500 m from trails compared to 50% of the PPGIS data.

We measured the percentage overlap of the predicted spatial distribution probability of values based on MaxEnt analysis of Flickr versus PPGIS data (Table 2; see maps in Figures S2a–d). Whereas different environmental and infrastructure characteristics have a stronger influence on value distribution in the two datasets, it appears that the spatial overlap between Flickr and PPGIS is relatively good, at least for recreation and scenic values (Table 2).

We used chi-square tests to assess differences in values between domestic and international visitors within the two datasets (Table 3). In the Flickr dataset, domestic visitors uploaded more images representing social values (127 photos vs. 31) and recreational values, whereas international visitors took significantly more photos of scenic landscapes (1,695 photos by internationals vs. 1,238 by domestic). PPGIS data revealed that domestic visitors mapped significantly more values representing cultural, gathering, hunting and fishing, and therapeutic values than international visitors who mapped more clean water, recreation and wilderness values.

The Spearman rank correlation confirmed the differences between domestic and international visitors within each dataset. The Spearman rank correlation test showed that domestic and international visitors in the Flickr database were highly positively correlated (rho = 0.96, $p = 0.0002$) in the ranked themes of contributed photos. The types of mapped values in PPGIS for domestic and international visitors were not as highly correlated (rho = 0.58, $p = 0.05$) based on frequency rankings.

TABLE 3 Standardized residuals for chi-square tests. Numbers below −2 or above 2 indicate that the value of domestic and international visitors has been mapped significantly less or more than would be expected within those two datasets (shaded). In brackets, the percentage a value was photographed/mapped by domestic and international visitors for each dataset

	Domestic	International	Pooled frequency (rank)
Flickr			
Biological	−0.82 (3.7)	0.82 (4.2)	4
Recreation	8.59 (19.9)	−8.59 (10.1)	2
Scenic	−10.52 (66.8)	10.52 (81.5)	1
Social	8.55 (6.9)	−8.55 (1.5)	3
Extraction	1.29 (0.4)	−1.29 (0.1)	8
Harvest	0.48 (0.5)	−0.48 (0.4)	7
Heritage	−0.28 (0.6)	0.28 (0.7)	6
Livelihood	−0.37 (1.3)	0.37 (1.5)	5
PPGIS			
Biological	−0.21 (4.6)	0.21 (4.8)	6
Recreation	−3.27 (28.9)	3.27 (34.3)	1
Scenic	−0.73 (27.3)	0.73 (28.5)	2
Social	1.78 (3.7)	−1.78 (2.5)	8
Clean water	−5.42 (6.6)	5.42 (11.9)	3
Cultural	4.94 (4.7)	−4.94 (1.3)	7
Gathering	2.96 (2.0)	−2.96 (0.7)	12
Hunt/fish	8.08 (8.1)	−8.08 (1.1)	5
Special place	0.92 (3.3)	−0.92 (2.7)	9
Spiritual	−1.32 (1.5)	1.32 (2.1)	11
Therapeutic	3.33 (3.4)	−3.33 (1.4)	10
Wilderness	−3.08 (6.0)	3.08 (8.8)	4

4 | DISCUSSION

We found large differences in the spatial data generated by passive versus active crowdsourcing methods. Flickr and PPGIS datasets differ substantially in both the types and locations of values mapped. Values represented in Flickr photos were located closer to roads than those mapped in the PPGIS dataset, which were predominantly located inside PAs and often associated with trails, mountain tops and glaciers. Despite these differences, the predicted spatial distribution of values generated by models applied to these two datasets showed substantial overlap, especially for scenic and recreational values, indicating that both datasets capture similar environmental and landscape characteristics. However, the overlap in value distribution suitability is lower when comparing domestic visitors in Flickr against domestic visitors in PPGIS (the same applies for international visitors). These differences, and the differences in the infrastructure and environmental variables that relate most to the distribution of values, may indicate that each crowdsourcing method gathers different information that is suitable at different scales (fine scale for PPGIS and regional scale for Flickr).

Our study results, consistent with Sonter, Watson, Wood, and Ricketts (2016), demonstrate that the value distribution can differ, or even be contradictory, depending on the data source, type of values mapped and the local contexts. As values mapped using the Flickr dataset are drawn from photographs, they can only represent visited places. In contrast, PPGIS allows the placement of a wider diversity of values, including areas that have not been visited, but that are important for the respondent (e.g. existence values). This may be one of the reasons why the clusters of values mapped using Flickr data are located in different places than those mapped using PPGIS. Moreover, images uploaded in Flickr might not be georeferenced according to the coordinates of the nature quality (e.g. the mountain photographed), but rather be placed where the picture was taken (Zielstra & Hochmair, 2013; e.g. the road from which the mountain was photographed). While tools that use elevation models or Google imagery to identify the location of scenic values from photographs are available (e.g. the Scenic Quality Package in InVEST; The Natural Capital Project, 2019), such tools are yet to be developed for more intangible values such as 'social' or 'special place'.

We found visitor infrastructure to be the most important factor explaining the spatial distribution of values in Flickr. Flickr tends to emphasize the importance of roads, and about 60% of the pictures were located within 500 m of a road. The fact that there is a high proportion of values found near roads does not mean that roads increase nature values. As shown by Kulczyk, Woźniak, and Derek (2018), the distribution of values can be locally affected by infrastructure despite nature being the true attraction in the region. Such data will not fully capture the fine-scale distribution of nature qualities that are appreciated in landscapes more distant from roads. Despite the strong bias towards roads, passive crowdsourced data can be valuable for identifying tourism hotspots and scenic routes on a regional and subregional scale and for informing management actions (e.g. Alivand & Hochmair, 2017). Contrary to our Flickr dataset, van Zanten et al. (2016) found hills and mountains to be the strongest predictors of scenic and recreation values using social media data. They controlled for accessibility using distance to big cities and travel time. Similarly, Kim, Kim, Lee, Lee, and Andrada (2019) found nature attractions such as beaches and waterfalls explained the distribution of Flickr data more than cultural sites and tourist facilities (i.e. accommodation venues and restaurants). These results indicate that the importance of infrastructure can differ depending on the local context. In our case, mountain tops and glaciers were the main predictor of recreational value for domestic visitors in the Flickr dataset, and for multiple values mapped by domestic visitors in the PPGIS dataset. Thus, both datasets can provide valuable information about NCP, confirmed by the high overlap between Flickr and PPGIS in the spatial MaxEnt models.

Differences among domestic and international visitors with respect to the use and appreciation of nature qualities within protected areas have previously been documented (Shultis, 1989; Tyrväinen, Mäntymaa, & Ovaskainen, 2014), but few studies have compared the spatial distribution of values among these two visitor groups. Spatially explicit analyses are important for detecting potential overlap of conflicting values of relevance to protected area management. Increasing tourism may have a low impact on local recreation if visitors and locals use different areas and value different nature qualities (Muñoz et al., 2019; Sonter et al., 2016). We found domestic visitors to upload more photos of recreation and social values and less scenic landscapes compared with international visitors. Similar results have been reported by Walden-Schreiner et al. (2018) and Fagerholm et al. (2019). Data from PPGIS captured a higher diversity of values compared to Flickr, with domestic visitors appreciating cultural, hunting and fishing, and gathering values more than international visitors, who mapped more values related to recreation, wilderness and clean water. The attachment of different groups of people to a place can be key to understand nature qualities that need to be managed, and to discern management actions to avoid conflicts among users (Fagerholm et al., 2019; Gundersen, Mehmetoglu, Vistad, & Andersen, 2015). The difference in mapped NCP can also be indicative of more deeply rooted cultural differences, as determined, for example, by the country one resides in Brown et al. (2015). Our study shows that PPGIS captures better the differences between domestic and international visitors than Flickr does, and will likely be more useful when developing strategies for tourism development and management.

5 | ADDITIONAL ADVANTAGES AND LIMITATIONS OF FLICKR AND PPGIS

As previous studies have concluded, crowdsourced data are a valuable source for assessing NCP. However, each method has their advantages and limitations that need to be carefully considered depending on the research questions to be addressed.

The first difference between the two platforms relies on the type of values that can be mapped. In PPGIS, values are generally pre-defined and the definition is available for the respondents, who decide which listed value they ascribe to a given place. However, for social media data, a code is assigned by experts based on the photographs or keywords (see e.g. Oteros-Rozas, Martin-Lopez, Fagerholm, Bieling, & Plieninger, 2018). While defining values in a PPGIS platform is flexible and can include a wide range of values because the process is deductive, the coding of social media pictures is an inductive process where themes are limited to those that can be identified visually. While this is likely to be reasonably accurate for values such as recreation (as represented by a photo of a person in skis for instance), some judgement on the part of the expert is involved for other values such as 'social' and 'spiritual' that are difficult to identify in this way. A limitation of this study lies in the fact that the PPGIS platform was used to map values that are not comparable to the values obtained in the Flickr photographs. The optimal would have been to include only the four comparable values to test for differences in the spatial distribution between the two datasets. On the other hand, by utilizing the full potential for mapping a diversity of values in PPGIS surveys, we could assess the potential of each platform to identify differences between visitors. We found passive crowdsourced data such as Flickr to be unreliable for capturing the full range of values and the importance of protected areas, including typical conservation values (biodiversity, wilderness and clean nature) and those important to local culture (cultural heritage, harvesting and social values; see also Levin et al., 2017).

Second, the accuracy and precision of these methods can be difficult to assess. The spatial accuracy of photo sharing platforms can be assessed through the positional error between the geotagged photo and the actual location of the picture. By visually matching photographs with ArcMap aerial imagery to estimate the camera position of the image, Zielstra and Hochmair (2013) found median errors in geospatial accuracy of Flickr images ranging from 46 to 1,606 m in different locations. For PPGIS, the accuracy of attributes that represent subjective judgements cannot be directly assessed against authoritative data (Brown & Kyttä, 2014). For spatial variables where accuracy could be evaluated, Brown (2012) and Cox, Morse, Anderson, and Marzen (2014) concluded that PPGIS respondents were able to accurately identify areas of native vegetation and suitable habitat for threatened species. In addition to the accuracy, the resolution of the data also affects the method. For example, Flickr has been shown to capture visitor distribution at coarse resolutions (several kilometres; Graham & Eigenbrod, 2019; Mancini, Coghill, & Lusseau, 2018; van Zanten et al., 2016), while PPGIS performs well at fine resolutions (Munro, Kobryn, Palmer, Bayley, & Moore, 2017).

Third, researchers need to make choices between the number of participants, representativeness and time frame available when using these different crowdsourced data. Crowdsourced data might be biased towards different users depending on the type of social media platform, knowledge about an area or place of residence (Brown & Kyttä, 2014; Bubalo et al., 2019; Ruths & Pfeffer, 2014). Demographic data are often not reported by the social media platforms. Also, there are studies showing that social media users are not representative of the general population, with educated people over-represented and gender bias shifting over time on different platforms (Li, Goodchild, & Xu, 2013; Mellon & Prosser, 2017; Mislove, Lehmann, Ahn, Onnela & Rosenquist, 2011). The social media platform used can provide different results. For example, Hausmann et al. (2018) found that Flickr users post more pictures related to biodiversity than Instagram users, who post more photos of people. However, Instagram performs better at estimating visitor rates than Flickr and Twitter (Tenkanen et al., 2017).

In our case, there was no visitor data available to assess whether the PPGIS data were biased towards mid-aged males and educated participants as shown in similar studies (Brown et al., 2015; Bubalo et al., 2019). Sampling design plays a crucial role in capturing a representative sample of the population or a targeted population segment (Brown, 2017; Brown & Kyttä, 2014; Brown et al., 2019). However, although data on visitation and visitor distribution provided by social media have previously been validated against local knowledge and field surveys (Kim et al., 2019), there is no available true representation of the spatial distribution of values with which Flickr and PPGIS data can be assessed.

6 | CONCLUSIONS

Crowdsourced data from passive and active sources can be a useful tool to inform managers about the spatial distribution of NCP in protected areas. Our results show that crowdsourced data provide fine-scale information on a diversity of values that people associate to protected areas, and the differences between user groups that are relevant for management. The methods differ in the distribution of values people ascribe to nature, for example in PPGIS a high proportion of values is located inside protected areas, whereas in Flickr they are more clustered and closer to roads. Although both methods are good at capturing scenic and recreation values, Flickr is more limited on the values that can be interpreted from pictures, whereas in PPGIS the values that are difficult to show in a picture can be captured (e.g. spiritual or inspirational values). We recommend a careful consideration of the type of data needed (in terms of values, explanatory variables and type of respondents) and logistical constraints (required quantity of data, scale and accuracy). To overcome some of the limitations of crowdsourcing data, combining these tools with field surveys could combine the benefits of both approaches, delivering large-scale datasets from a broad user sample along with more detailed and specific information on NCP and the nature qualities that are valued by different groups of people.

ACKNOWLEDGEMENTS

This study was funded by CultES—Assessing spatially explicit cultural ecosystem services for adaptive management in the Alpine North, Environmental Research Program, Norwegian Research Council, 230330/E50/2014. We would like to thank the two anonymous reviewers and the associate editor for their constructive comments on earlier versions of the draft.

CONFLICT OF INTEREST

Nothing to declare.

AUTHORS' CONTRIBUTIONS

L.M. and V.H.H. conceived the ideas of the paper. All authors contributed to data gathering design of Flickr and/or through a dedicated PPGIS platform; R.D., V.H.H. and L.M. collected the data for Flickr using an API, and for PPGIS through a household survey and through volunteer recruitment on media and in situ; L.M. coded Flickr data and prepared environmental and infrastructure data, and L.M., V.H.H. and C.R. analysed the data; L.M. and V.H.H. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA AVAILABILITY STATEMENT

The data are publicly available at DataverseNO https://doi.org/10.18710/VQLTM8 (Muñoz, Hausner, Runge, Brown, & Daigle, 2020).

ORCID

Lorena Muñoz https://orcid.org/0000-0002-7514-6384
Vera Helene Hausner https://orcid.org/0000-0001-9825-0419

REFERENCES

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

Arler, F. (2000). Aspects of landscape or nature quality. *Landscape Ecology*, *15*(3), 291–302. https://doi.org/10.1023/A:1008192301500

Bagstad, K. J., Semmens, D. J., Ancona, Z. H., & Sherrouse, B. C. (2016). Evaluating alternative methods for biophysical and cultural ecosystem services hotspot mapping in natural resource planning. *Landscape Ecology*, *32*(1), 77–97. https://doi.org/10.1007/s10980-016-0430-6

Birenboim, A., & Shoval, N. (2016). Mobility research in the age of the smartphone. *Annals of the American Association of Geographers*, *106*(2), 283–291. https://doi.org/10.1080/00045608.2015.1100058

Brown, G. (2012). An empirical evaluation of the spatial accuracy of public participation GIS (PPGIS) data. *Applied Geography*, *34*, 289–294. https://doi.org/10.1016/j.apgeog.2011.12.004

Brown, G. (2017). A review of sampling effects and response bias in internet participatory mapping (PPGIS/PGIS/VGI). *Transactions in GIS*, *21*(1), 39–56. https://doi.org/10.1111/tgis.12207

Brown, G., Hausner, V. H., Grodzińska-Jurczak, M., Pietrzyk-Kaszyńska, A., Olszańska, A., Peek, B., … Lægreid, E. (2015). Cross-cultural values and management preferences in protected areas of Norway and Poland. *Journal for Nature Conservation*, *28*, 89–104. https://doi.org/10.1016/j.jnc.2015.09.006

Brown, G., Kelly, M., & Whitall, D. (2014). Which “public”? Sampling effects in public participation GIS (PPGIS) and volunteered geographic information (VGI) systems for public lands management. *Journal of Environmental Planning and Management*, *57*(2), 190–214. https://doi.org/10.1080/09640568.2012.741045

Brown, G., & Kyttä, M. (2014). Key issues and research priorities for public participation GIS (PPGIS): A synthesis based on empirical research. *Applied Geography*, *46*, 122–136. https://doi.org/10.1016/j.apgeog.2013.11.004

Brown, G., Rhodes, J., Lunney, D., Goldingay, R., Fielding, K., Garofano, N., … McAlpine, C. (2019). The influence of sampling design on spatial data quality in a geographic citizen science project. *Transactions in GIS*, *23*(6), 1184–1203. https://doi.org/10.1111/tgis.12568

Brown, G., & Reed, P. (2000). Validation of a forest values typology for use in national forest planning. *Forest Science*, *46*(2), 240–247.

Brown, G., & Weber, D. (2011). Public participation GIS: A new method for national park planning. *Landscape and Urban Planning*, *102*(1), 1–15. https://doi.org/10.1016/j.landurbplan.2011.03.003

Brown, G., & Weber, D. (2012). Measuring change in place values using public participation GIS (PPGIS). *Applied Geography*, *34*, 316–324. https://doi.org/10.1016/j.apgeog.2011.12.007

Brown, G., & Weber, D. (2013). Using public participation GIS (PPGIS) on the Geoweb to monitor tourism development preferences. *Journal of Sustainable Tourism*, *21*(2), 192–211. https://doi.org/10.1080/09669582.2012.693501

Bubalo, M., van Zanten, B. T., & Verburg, P. H. (2019). Crowdsourcing geo-information on landscape perceptions and preferences: A review. *Landscape and Urban Planning*, *184*, 101–111. https://doi.org/10.1016/j.landurbplan.2019.01.001

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

Cox, C., Morse, W., Anderson, C., & Marzen, L. (2014). Applying Public Participation Geographic Information Systems to wildlife management. *Human Dimensions of Wildlife*, *19*(2), 200–214. https://doi.org/10.1080/10871209.2014.871663

Daigle, R., & Dunnington, D. (2018). FlickrRgeotag. *Zenodo*. https://doi.org/10.5281/zenodo.1314717

Dunkel, A. (2015). Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning*, *142*, 173–186. https://doi.org/10.1016/j.landurbplan.2015.02.022

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, *17*(1), 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Engen, S., Runge, C., Brown, G., Fauchald, P., Nilsen, L., & Hausner, V. (2018). Assessing local acceptance of protected area management using public participation GIS (PPGIS). *Journal for Nature Conservation*, *43*, 27–34. https://doi.org/10.1016/j.jnc.2017.12.002

Ester, M., Kriegel, H.-P., Sander, J., & Xiaowei, X. U. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD*, *96*, 226–231.

European Environmental Agency. (2015). Corine Land Cover (CLC) 2006 V17. Copernicus Program. https://land.copernicus.eu/pan-european/corine-land-cover/clc-2006/view

Fagerholm, N., Torralba, M., Moreno, G., Girardello, M., Herzog, F., Aviron, S., … Plieninger, T. (2019). Cross-site analysis of perceived ecosystem service benefits in multifunctional landscapes. *Global Environmental Change*, *56*, 134–147. https://doi.org/10.1016/j.gloenvcha.2019.04.002

Flickr. 2019. About Flickr. Retrieved from https://www.flickr.com/about

Graham, L. J., & Eigenbrod, F. (2019). Scale dependency in drivers of outdoor recreation in England. *People and Nature*, *1*(3), 406–416. https://doi.org/10.1002/pan3.10042

Gundersen, V., Mehmetoglu, M., Vistad, O. I., & Andersen, O. (2015). Linking visitor motivation with attitude towards management restrictions on use in a national park. *Journal of Outdoor Recreation and Tourism*, *9*, 77–86. https://doi.org/10.1016/j.jort.2015.04.004

Hahsler, M., Piekenbrock, M., Arya, S., & Mount, D. (2018). Dbscan: Density based clustering of applications with noise (DBSCAN) and related algorithms. Retrieved from https://cran.r-project.org/package=dbscan

Haines-Young, R., & Potschin, M. (2013). Common International Classification of Ecosystem Services (CICES): Consultation on Version 4, August–December 2012. EEA Framework Contract No EEA/IEA/09/003. Retrieved from https://cices.eu/content/uploads/sites/8/2012/07/CICES-V43_Revised-Final_Report_29012013.pdf

Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V., & Di Minin, E. (2018). Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas. *Conservation Letters*, *11*(1), e12343. https://doi.org/10.1111/conl.12343

Hausner, V. H., Brown, G., & Lægreid, E. (2015). Effects of land tenure and protected areas on ecosystem services and land use preferences in Norway. *Land Use Policy*, *49*, 446–461. https://doi.org/10.1016/j.landusepol.2015.08.018

Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns.

Cartography and Geographic Information Science, *41*(3), 260–271. https://doi.org/10.1080/15230406.2014.890072

Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2017). Dismo: Species distribution modelling. Retrieved from http://rspatial.org/sdm/

Hirons, M., Comberti, C., & Dunford, R. (2016). Valuing cultural ecosystem services. *Annual Review of Environment and Resources*, *41*(1), 545–574. https://doi.org/10.1146/annurev-environ-110615-085831

Jotunheimen. 2019. Jotunheimen. Retrieved from jotunheimen.com

Kalle, R., Ramesh, T., Qureshi, Q., & Sankar, K. (2013). Predicting the distribution pattern of small carnivores in response to environmental factors in the Western Ghats. *PLoS ONE*, *8*(11), e79295–e79295. https://doi.org/10.1371/journal.pone.0079295

Kartverket. (2015). N500 database. Land division, Cartography section. Contact person Signy Berge, Projection EUREF89 UTM zone33.

Kim, Y., Kim, C.-K., Lee, D. K., Lee, H.-W., & Andrada, R. I. I. T. (2019). Quantifying nature-based tourism in protected areas in developing countries by using social big data. *Tourism Management*, *72*, 249–256. https://doi.org/10.1016/j.tourman.2018.12.005

Kisilevich, S., Keim, D., Andrienko, N., & Andrienko, G. (2013). Towards acquisition of semantics of places and events by multi-perspective analysis of geotagged photo collections BT – Geospatial visualisation. In A. Moore & I. Drecki (Eds.), *Geospatial visualisation* (pp. 211–233). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-12289-7_10

Kulczyk, S., Woźniak, E., & Derek, M. (2018). Landscape, facilities and visitors: An integrated model of recreational ecosystem services. *Ecosystem Services*, *31*, 491–501. https://doi.org/10.1016/j.ecoser.2018.02.016

Levin, N., Lechner, A. M., & Brown, G. (2017). An evaluation of crowdsourced information for assessing the visitation and perceived importance of protected areas. *Applied Geography*, *79*, 115–126. https://doi.org/10.1016/j.apgeog.2016.12.009

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

Mancini, F., Coghill, G. M., & Lusseau, D. (2018). Using social media to quantify spatial and temporal dynamics of nature-based recreational activities. *PLoS ONE*, *13*(7), 1–19. https://doi.org/10.1371/journal.pone.0200565

Mellon, J., & Prosser, C. (2017). Twitter and Facebook are not representative of the general population: Political attitudes and demographics of british social media users. *Research & Politics*, *4*(3), 2053168017720008. https://doi.org/10.1177/2053168017720008

Ministry of Climate and Environment. (2014). Forskrift Om Utladalen Landskapsvernområde, Luster Og Årdal Kommunar, Sogn Og Fjordane, Norway. Retrieved from https://lovdata.no/dokument/LTII/forskrift/2014-11-14-1399

Ministry of Climate and Environment. (2019). Nasjonalparker Og Andre Verneområder. Retrieved from https://www.regjeringen.no/no/tema/klima-og-miljo/naturmangfold/innsiktsartikler-naturmangfold/nasjonalparker-og-andre-verneomrader/id2076286/

Mislove, A., Lehmann, S., Ahn, Y.-Y., Onnela, J.-P., & Rosenquist, J. N. (2011). Understanding the demographics of Twitter users. In *Proceedings of the fifth international AAAI conference on weblogs and social media* (Vol. 2932, pp. 554–557). Barcelona, Spain: AAAI Press. Retrieved from https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2816

Muñoz, L., Hausner, V. H., Brown, G., Runge, C., & Fauchald, P. (2019). Identifying spatial overlap in the values of locals, domestic- and international tourists to protected areas. *Tourism Management*, *71*, 259–271. https://doi.org/10.1016/j.tourman.2018.07.015

Muñoz, L., Hausner, V. H., Runge, C., Brown, G., & Daigle, R. (2020). Replication Data for: Using crowdsourced spatial data from Flickr vs. PPGIS for understanding nature's contribution to people in Southern Norway. *DataverseNO*, https://doi.org/10.18710/VQLTM8

Munro, J., Kobryn, H., Palmer, D., Bayley, S., & Moore, S. A. (2017). Charting the coast: Spatial planning for tourism using public participation GIS. *Current Issues in Tourism*, *22*(4), 486–504. https://doi.org/10.1080/13683500.2017.1346589

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

Pascual, U., Balvanera, P., Díaz, S., Pataki, G., Roth, E., Stenseke, M., … Yagi, N. (2017). Valuing nature's contributions to people: The IPBES approach. *Current Opinion in Environmental Sustainability*, *26–27*, 7–16. https://doi.org/10.1016/j.cosust.2016.12.006

Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, *31*(2), 161–175. https://doi.org/10.1111/j.0906-7590.2008.5203.x

Phillips, S. J., Dudík, M., & Schapire, R. E. (n.d.). Maxent software for modeling species niches and distributions. Retrieved from http://biodiversityinformatics.amnh.org/open_source/maxent/

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Richards, D. R., & Friess, D. A. (2015). A rapid indicator of cultural ecosystem service usage at a fine spatial scale: Content analysis of social media photographs. *Ecological Indicators*, *53*, 187–195. https://doi.org/10.1016/j.ecolind.2015.01.034

Richards, D. R., & Tunçer, B. (2018). Using image recognition to automate assessment of cultural ecosystem services from social media photographs. *Ecosystem Services*, *31*, 318–325. https://doi.org/10.1016/j.ecoser.2017.09.004

Ridding, L. E., Redhead, J. W., Oliver, T. H., Schmucki, R., McGinlay, J., Graves, A. R., … Bullock, J. M. (2018). The importance of landscape characteristics for the delivery of cultural ecosystem services. *Journal of Environmental Management*, *206*, 1145–1154. https://doi.org/10.1016/j.jenvman.2017.11.066

Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, *346*(6213), 1063–1064. https://doi.org/10.1126/science.346.6213.1063

See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., … Rutzinger, M. (2016). Crowdsourcing, citizen science or volunteered geographic information? The current state of crowdsourced geographic information. *ISPRS International Journal of Geo-Information*, *5*(5), 55–https://doi.org/10.3390/ijgi5050055

Shultis, J. D. (1989). Images and use of New Zealand's protected areas by domestic and international visitors. *GeoJournal*, *19*(3), 329–335. https://doi.org/10.1007/BF00454580

Small, N., Munday, M., & Durance, I. (2017). The challenge of valuing ecosystem services that have no material benefits. *Global Environmental Change*, *44*, 57–67. https://doi.org/10.1016/j.gloenvcha.2017.03.005

Sonter, L. J., Watson, K. B., Wood, S. A., & Ricketts, T. H. (2016). Spatial and temporal dynamics and value of nature-based recreation, estimated via social media. *PLoS ONE*, *11*(9), e0162372. https://doi.org/10.1371/journal.pone.0162372

Teff-Seker, Y., & Orenstein, D. E. (2019). The 'desert experience': Evaluating the cultural ecosystem services of drylands through walking and focusing. *People and Nature*, *1*(2), 234–248. https://doi.org/10.1002/pan3.28

Tenerelli, P., Demšar, U., & Luque, S. (2016). Crowdsourcing indicators for cultural ecosystem services: A geographically weighted approach for mountain landscapes. *Ecological Indicators*, *64*, 237–248. https://doi.org/10.1016/j.ecolind.2015.12.042

Tenkanen, H., Di Minin, E., Heikinheimo, V., Hausmann, A., Herbst, M., Kajala, L., & Toivonen, T. (2017). Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-18007-4

Tew, E. R., Simmons, B. I., & Sutherland, W. J. (2019). Quantifying cultural ecosystem services: Disentangling the effects of management from landscape features. *People and Nature*, *1*(1), 70–86. https://doi.org/10.1002/pan3.14

The Natural Capital Project. (2019). InVEST scenic quality model. Retrieved from https://invest.readthedocs.io/en/latest/final_ecosystem_services/scenic_quality.html

The Norwegian Directorate for Nature Management. (2019). Breheimen. From perennial ice to lush valleys with transhumance farms. *The Norwegian Directorate for Nature Management*. Retrieved from https://www.miljodirektoratet.no/globalassets/dokumenter/publikasjoner/brosjyrer/breheimen_np_e_230511.pdf

Toivonen, T., Heikinheimo, V., Fink, C., Hausmann, A., Hiippala, T., Järv, O., … Di Minin, E. (2019). Social media data for conservation science: A methodological overview. *Biological Conservation*, *233*, 298–315. https://doi.org/10.1016/j.biocon.2019.01.023

Tyrväinen, L., Mäntymaa, E., & Ovaskainen, V. (2014). Demand for enhanced forest amenities in private lands: The case of the Ruka-Kuusamo tourism area, Finland. *Forest Policy and Economics*, *47*, 4–13. https://doi.org/10.1016/j.forpol.2013.05.007

Van den Bosch, M., Östergren, P.-O., Grahn, P., Skärbäck, E., & Währborg, P. (2015). Moving to serene nature may prevent poor mental health – Results from a Swedish longitudinal cohort study. *International Journal of Environmental Research and Public Health*, *12*(7), 7974–7989. https://doi.org/10.3390/ijerph120707974

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

Walden-Schreiner, C., Leung, Y.-F., & Tateosian, L. (2018). Digital footprints: Incorporating crowdsourced geographic information for protected area management. *Applied Geography*, *90*, 44–54. https://doi.org/10.1016/j.apgeog.2017.11.004

Wolf, I. D., Brown, G., & Wohlfart, T. (2018). Applying public participation GIS (PPGIS) to inform and manage visitor conflict along multi-use trails. *Journal of Sustainable Tourism*, *26*(3), 470–495. https://doi.org/10.1080/09669582.2017.1360315

Wood, S. A., Guerry, A. D., Silver, J. M., & Lacayo, M. (2013). Using social media to quantify nature-based tourism and recreation. *Scientific Reports*, *3*, 2976.https://doi.org/10.1038/srep02976

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

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Muñoz L, Hausner VH, Runge C, Brown G, Daigle R. Using crowdsourced spatial data from Flickr vs. PPGIS for understanding nature's contribution to people in Southern Norway. *People Nat*. 2020;2:437–449. https://doi.org/10.1002/pan3.10083