Journal of Applied Ecology

MISS FRANCESCA MANCINI (Orcid ID : 0000-0003-4085-4978)

Handling Editor: Yolanda Wiersma

Quantifying wildlife watchers' preferences to investigate the overlap between

recreational and conservation value of natural areas

Authors: Francesca Mancini[a], George M. Coghill[b], David Lusseau[a]

[a]Institute of Biological and Environmental Sciences, University of
Aberdeen, Zoology Building, Tillydrone Avenue, Aberdeen AB24 2TZ,
UK.

[b]Department of Computing Science, University of Aberdeen, Meston
Building, Meston Walk, Aberdeen AB24 3UE, UK.

Corresponding author: Francesca Mancini

Room 418, Zoology Building, Tillydrone Avenue, Aberdeen AB24 2TZ,

UK. Telephone: 01224 274106. Email: r03fm14@abdn.ac.uk

1. Nature-based recreation substantially benefits human wellbeing, for example, by improving

   physical and mental health. However, recreation can also have severe ecological impacts.

   The recreational value of landscapes and natural areas is often used to generate support for

   public spending in conservation. However, we still don't know whether nature-based

recreationists place greater recreational value on natural areas that have high conservation value compared to other green spaces.

2. Here, we determine which attributes of nature-based tourism provide recreational services. We used pictures of wildlife posted on Flickr to quantify wildlife watching activities in Scotland. We then determined the environmental variables key to attracting wildlife watchers to a destination, such as protected areas (PAs), the perceived naturalness, and the presence of different types of infrastructure.

3. Infrastructure best predicts the intensity of wildlife watching activities in Scotland, while areas of high natural value are rarely used. PAs are weak attractors of wildlife watchers, with PAs designated to protect threatened habitats or species having low recreational value. In accessible and highly visited areas, higher biodiversity increases the intensity of wildlife watching activities.

4. *Synthesis and applications.* Areas of high natural and conservation value and areas of high recreational value do not tend to overlap. Recreational ecosystem services are mainly provided by the wider countryside and highly transformed landscapes, as opposed to wild ecosystems and protected areas designated to protect environmental features of high conservation value. These results question the synergy between the goals of recreation and those of conservation and the use of recreation as a justification for economic investment in conservation. During wildlife watching activities most people experience an urbanised, highly transformed nature; it will be important to determine how this human-dominated nature can influence support for conservation of wild and remote areas.

Introduction

Recreation is one of the key cultural ecosystem services offered by nature. It provides substantial benefits for human wellbeing, including improving physical and mental health (Sandifer, Sutton-Grier & Ward 2015). The recreational value of landscapes and natural areas is often used to generate support for public spending in conservation (Wilkie & Carpenter 1999; Balmford *et al.* 2009). For example, Protected Areas (PAs) are under increasing pressure to generate income (Walpole, Goodwin & Ward 2001). The role of nature-based tourism in conservation is still debated. While tourism can generate economic benefits for local communities and conservation (Krüger 2005), it faces "leakages": revenues are often lost from the local area and very little is reinvested into conservation (Sandbrook 2010). Importantly it is not clear whether nature-based recreationists are more likely to value natural areas of high conservation value.

All nature-based recreational activities, including hiking and wildlife watching, can have ecological impacts as severe as declines in species richness and shifts in community composition from native to non-native species (Reed & Merenlender 2008). We need to understand the role of environmental features of high conservation value (e.g., threatened species or high naturalness) in driving attraction of recreationists to a site to determine whether increased recreation might lead to conflicts with conservation goals. Most studies looking at wildlife watchers' and nature recreationists' preferences have focused on visitors to PAs (De Vos *et al.* 2016; Sessions *et al.* 2016; Sonter *et al.* 2016; Baum, Cumming & De Vos 2017), thus missing an important part of nature recreationists that use green spaces in urban areas or the wider countryside. Our view on preferences is therefore likely biased towards specialists, who will have different preferences from the general public. Also, they might have underestimated the recreational value of some environmental features or infrastructure that are not represented in PAs.

Here we identify the natural and human characteristics of a destination that predict high intensity of recreational activities. We focused on wildlife watching in Scotland, where nature-based tourism contributes £1.4 billion per year to the economy, creating 39000 FTE (Full Time Equivalent) jobs, with £153 million attributable to wildlife watching alone (Bryden *et al.* 2010). Surveys are the most common tools to quantify recreational activities (Casado-Arzuaga *et al.* 2014; Peña, Casado-Arzuaga & Onaindia 2015) and questionnaires are mostly used to understand visitors' preferences (van Zanten *et al.* 2016b). The widespread use of social media offers now a great opportunity to use crowdsourced datasets, allowing us to sample more people. A few studies have compared geotagged photographs uploaded on social media to visitor statistics obtained through more traditional methods (e.g. surveys or censuses) (Wood *et al.* 2013; Levin, Kark & Crandall 2015; Sessions *et al.* 2016; Hausmann *et al.* 2017b; Levin, Mark & Brown 2017), and all of these studies have demonstrated that data from social media is a reliable indicator of intensity of recreational activities and preferences. We used the number of users of the photo-sharing website Flickr that posted photographs of wildlife taken in Scotland as our measure of intensity of wildlife watching activities (Mancini, Coghill & Lusseau 2018). We assume, as in previous studies, that the more people photograph and share information about a particular location, the higher its recreational value and we interpret spatial concentration of social media content as an indicator of the popularity of a wildlife watching destination (Gliozzo, Pettorelli & Haklay 2016; van Zanten *et al.* 2016a).

The potential of ecosystems to provide recreational services depends on different factors, such as their beauty, naturalness and biodiversity, and the presence of a PA, but accessibility and infrastructure are also crucial (Maes *et al.* 2011; Peña, Casado-Arzuaga & Onaindia 2015). For example, large areas of Scotland have semi-natural landscapes that show minimal signs of current human influence. These can be a variety of habitat types, from mountains to undeveloped coastline, but they are all characterised by high naturalness and biodiversity (therefore high conservation value), low accessibility and no infrastructure. Areas that have higher naturalness and biodiversity, protected under certain designations (Site of Special Scientific Interest or Marine Protected Areas)

will also have a high conservation value because they were designated to protect threatened species or important habitats, but they will provide better accessibility and more infrastructure compared to wild areas. Other PA designations, such as Country Parks (CNTRY) or Local Nature Reserves (LNR) state as their main goal to provide opportunities for the public to enjoy nature close to where they live and were designated with the dual objective of preserving important natural and cultural heritage and providing people with recreational opportunities. Therefore, these areas will have lower conservation value compared to PAs such as Marine Protected Areas and to wild areas. Following this conceptual framework, we tested the effect of different types of PA designations, naturalness, biodiversity, accessibility and presence of recreational infrastructure on the intensity of wildlife watching activities in Scotland. We aimed to assess whether wildlife watchers place greater recreational value on natural areas that have high conservation value compared to other green spaces.

Materials and Methods

Data collection

Wildlife watching – To quantify wildlife watching activities in Scotland, we queried the Flickr Application Programming Interface (API) for photographs of wildlife taken in Scotland between 2005 and 2015. We used 4 keywords to select only relevant photographs: "bird", "seal", "whale" and "dolphin", which are the main groups of charismatic wildlife watched in Scotland (Curtin 2013). We used packages RCurl (Lang and the CRAN team, 2015), XML (Lang and the CRAN team, 2015b) and httr (Wickham, 2016) in R (R Core Team, 2015) to communicate with the API, and request and download the data. We downloaded the metadata associated with the photos: photograph and user ID, the date when the photo was taken, the geographic coordinates of where it was taken, user tags and description. Using the user tags associated with the photos, we eliminated pictures that were

not relevant (such as photos of statues or paintings and photos taken in zoos). The tags were

examined and a list of keywords for non-relevant photos was compiled, then, following a method

similar to that used in (van Zanten *et al.* 2016a), we used that list of keywords to filter out irrelevant

photos. In order to avoid bias coming from having a small number of very active users, we used the

combination of user ID and date to delete multiple photos from the same user on the same day.

Therefore, the number of data points in the dataset (41203) represents the number of Flickr visitor

days (FVD; Fig. 1). For details on this data collection procedure see (Mancini, Coghill & Lusseau 2018

and Data Sources)

Infrastructure – Information about infrastructure was downloaded from Google Places API Web

Service and Ordnance Survey (see Table S1 in Supporting Information and Data sources). Google

Places API was queried to obtain the locations of tourist accommodations, tour operators offering

wildlife recreational activities, airports, bus stops, train stations and car parks (Fig. 1). We used the

packages httr (Wickham 2016) to query the API and jsonlite (Ooms, 2014) to format the data, and

we used the Google Places API Radar Search Services to perform the search for the different types of

infrastructure on the Google Places API. Because of the limits in the search results retuned by this

service (maximum 200 places), we created a grid of 2x2 Km cells over Scotland and used the

coordinates of the cell centres for the search, with a radius of 2.5 Km. This search produced

duplicates, which were then removed using the place ID. We used the "type" argument in the

Google Places API Radar Search Services to download information on the different types of

infrastructure: "lodging", "airport", "bus_station", "train_station" and "parking". We further filtered

the airport list by only selecting those with airline services according to OurAirports, a website that

provides a collection of aviation data for airports around the world.

(http://ourairports.com/countries/GB/SCT/airports.html?show=scheduled). For tour operators, we

used both "type=tour_operator" and "keyword=wildlife" to restrict the search to wildlife-related

tours only. This search returned some non-relevant results, therefore we went through each location

returned by the API and manually selected only those operations that charged money for some kind

of wildlife watching infrastructure, such as boat tours, guided walks, visitor centres with cameras on bird nests or bird hides. Each of these searches only returned the place ID, we then used these IDs in Place Details Requests to obtain the coordinates and name of each infrastructure. In total we obtained the locations of 14057 tourist accommodations, 22 airports, 41166 bus stops, 372 train stations, 693 car parks and 112 tour operators.

We also downloaded a shapefile of roads from Ordnance Survey Open Data (Table S1, Data sources and Fig. 1) to test the effect of the presence of roads on intensity of wildlife watching.

Environmental data – All the environmental variables tested were obtained from Scottish Natural Heritage (SNH) Natural Spaces application (Table S1 and Data Sources), except the biodiversity records which were downloaded from the Global Biodiversity Information Facility (GBIF, Table S1 and Data sources).

From SNH Natural Spaces we downloaded shapefiles with boundaries for PAs in Scotland (Fig. 1). These areas include: local designations, such as Local Nature Reserves (LNR), Marine Consultation Areas (MCAs), SNH Nature Reserves (NR) and Country Parks (CNTRY); national designations, such as Sites of Special Scientific Interest (SSSI), National Nature Reserves (NNR), National Parks (NP) and Marine Protected Areas (MPA); international designations such as Council of Europe Diploma Sites (COUNEUR), Natura Sites (Special Areas of Conservation – SAC – and Special Protection Areas – SPA), Ramsar sites, Biosphere (BIOSPH) and Biogenetic (BIOGEN) reserves and World Heritage Sites (WHS).

From SNH Natural Spaces we also downloaded a raster of perceived naturalness in Scotland at a resolution of 1 m (Fig. 1). This naturalness score was estimated previously from a land use map, where each land class was given a naturalness score from 1 (low naturalness) to 5 (high naturalness); then a focal statistical window of 250 m was passed over the dataset averaging the naturalness value to account for surrounding areas (for more details see Carver et al. 2008).

Records of species occurrence were downloaded from GBIF (Table S1 and Data sources). We queried the GBIF database for occurrences of *Phocidae*, *Aves* and *Cetacea* in Scotland, which returned 333,105 occurrences (Fig. 1).

Data manipulation

We showed previously that FVD counts is an unbiased measure of wildlife watching intensity at a resolution as low as 10x10 Km(Mancini, Coghill & Lusseau 2018). Therefore, we aggregated all variables previously described at this resolution. Data were processed in ArcGIS 10.4.1 (Table S2).

The response variable, intensity of wildlife watching (Count_WW), was obtained by counting the number of FVD in each 10x10 Km cell (Fig. 1). Because the naturalness score was only available for terrestrial locations, we excluded all the photos that were taken at sea (135 photographs). The polygon shapefiles representing the boundaries of the different types of PAs were processed by calculating the area of each cell that was covered by that type of PA (Area_*). For marine PAs we calculated the distance between the centres of each cell to the boundaries of the nearest reserve instead (Dist_*). We also created an aggregated variable, where we combined all the types of terrestrial PAs into one layer and then calculated for each cell the area covered by any type of PA (Area_PA). For each cell we also calculated an average naturalness value by calculating the mean of the naturalness score across all the raster cells within our 10x10 Km grid cells (Mean_Nat). We then calculated the number of and the distance from the centres of each cell to the nearest airport (Dist_Air), car park (Dist_CarPark), tour operator (Dist_TourOp), train station (Dist_Train) and road (Dist_Road), and the number of tourist accommodations (Count_Hotels) and bus stops (Count_Bus) in each cell. As we did for PAs, we calculated an aggregated infrastructure variable which consisted of the count of any type of infrastructure inside each cell (Count_Inf). Species richness was calculated by counting the number of unique species recorded in the GBIF data in each cell (Species).

For occurrences of marine species, we only selected those that were within 1 Km from the coast, we then assigned each occurrence to their nearest coastal grid cell and added these to the number of terrestrial species occurrences in the same cell.

Analysis

The aim of this analysis was to understand how environmental and infrastructure variables explained the variance in the intensity of wildlife watching at a 10x10 Km resolution. First, we wanted to know which aspect of a destination (e.g. the infrastructure, the presence of PAs or its perceived naturalness) is most important in explaining intensity of wildlife watching. Then we tested the effect of each specific infrastructure, type of PA or environmental variable (naturalness or species richness). Therefore, we performed our model selection in three steps. We first regressed our response variable, the intensity of wildlife watching against aggregated variables (area covered by a PA, number of infrastructure) and mean naturalness score (Table S1). We then fitted an infrastructure model, where we estimated the effect of each type of infrastructure on the response variable, and an environmental model where we tested the effect of each different type of PA and of the mean naturalness score on the intensity of wildlife watching. We adopted this approach to avoid overfitting due to the high number of covariates that we would have had to use in one full model. For the same reason and due to the limited sample size (1132 observations), we were not able to test for the effect of interactions between some of the explanatory variables, which would have resulted in low degrees of freedom and over fitted models. Moreover, given the degree of correlation between some of our explanatory variables, for example different types of infrastructure, we didn't fit interaction terms to avoid over inflation. We also performed an analysis in which all the predictors selected in the infrastructure and environmental models were tested together to directly compare their effects on the intensity of wildlife watching (Table S3).

The spatial distribution of species records from GBIF overlapped the distribution of FVD. Some areas of Scotland are difficult to access, the terrain is challenging and generally are only visited by experienced hikers. In these areas both the intensity of wildlife watching and the number of GBIF records were zero or very low (Fig. S1 in Appendix S1). As a consequence of this, we decided to subset the dataset according to the number of GBIF records so we only considered those areas where enough data were present. The distribution of this variable was bimodal, with one mode around 10 and one around 10000 records (Fig. S2 in Appendix S1). We fitted a mixture model to the number of GBIF records to obtain a classification of our observations using the function densityMclust in the R package mclust (Fraley & Raftery 2002; Fraley *et al.* 2012) to classify our observations into two groups, high and low GIBIF records. We then only used the observations belonging to the "high" group to test the effect of the number of species on the intensity of wildlife watching.

Given the skewness of the response variable we log_{10} transformed it. We checked for collinearity between our explanatory variables by estimating Variance Inflation Factors (VIF). All the variables had VIF < 2 except the area covered by a Site of Special Scientific Interest and area covered by a terrestrial Special Area of Conservation for which VIF was > 3; therefore, these two variables were not used together in the same models.

We first fitted a linear model using log_{10}–transformed intensity of wildlife watching as response variable and all the other variables as explanatory variables. Inspection of the residuals from this model revealed the presence of spatial autocorrelation. In order to account for these spatial patterns we used linear regressions fitted with Generalised Least Squares, which allows for the inclusion of autocorrelation structures. We used the function gls in the R package nlme (Pinheiro *et al.* 2016). All the covariates were centred around the mean and scaled by their standard deviation so that coefficients were comparable.

Aggregated model – The first model we fitted is described in Eq. 1:

$$\log_{10}(Count\_WW) \sim Area\_PA + Count\_Inf + Mean\_Nat \qquad \text{Eq.1}$$

We fitted 5 models with different correlation structures (rational quadratic, spherical, linear, Gaussian and exponential) with restricted maximum likelihood (REML) and then used AIC to select the retained correlation structure. We then used diagnostic plots and variograms of the model residuals to check that model assumptions were not violated.

Environmental model – This model is described by Eq. 2:

$$\log_{10}(Count\_WW) \sim Area\_WHS + Area\_SAC\_L + Area\_SSSI + Dist\_MSAC + Dist\_MPA +$$
$$Dist\_MCA + Area\_SPA + Area\_RAMSA + Area\_NR + Area\_NNR + Area\_LNR +$$
$$Area\_COUNE + Area\_CNTRY + Area\_BIOSP + Area\_BIOGE + Area\_NP + Mean\_Nat \quad \text{Eq. 2}$$

After choosing the best autocorrelation structure using the same method we used for the aggregated model, we performed model validation by visually inspecting diagnostic plots and variograms of model residuals, then we performed model selection. Given the high number of variables we decided to use a data mining approach. We used the function dredge in the R package MuMIn (Barton 2016) to fit models with all the possible combinations of the explanatory variables (without interactions), except for models that contained the two collinear variables together. All the models were fitted with maximum likelihood (ML) and after AIC and weights were calculated we refitted all the models with a ΔAIC < 5 with REML to obtain unbiased coefficients. We then calculated averaged coefficients across the best (ΔAIC < 5) models.

Infrastructure model – In this model we tested the effect of the different types of infrastructure:

$$\log_{10}(Count\_WW) \sim Dist\_Air + Count\_Bus + \log_{10}(Count\_Hotel) + Dist\_CarPark +$$
$$Dist\_TourOp + Dist\_Train + Dist\_Road \qquad \text{Eq. 3}$$

Model validation and selection were performed as for the environmental model.

Species richness model – After sub setting the dataset according to the number of GBIF records (mixing probabilities: low 0.13, high 0.87; means: low 18, high 25118), we used the variables that were selected as important by the model selection procedure on the environmental model and used them as covariates together with the number of species in this model:

$$\log_{10}(Count\_WW) \sim Dist\_MSAC + Area\_LNR + Area\_CNTRY + Area\_NP + Mean\_Nat +$$
$$Species \qquad\qquad\qquad \text{Eq. 2}$$

For every model we then produced maps of model residuals (averaged for the environmental and infrastructure models) to investigate how well our models fit the data and identify any spatial patterns left unexplained.

Results

We retained an exponential correlation structure for all models.

Aggregated model – All the variables had a significant effect on the intensity of wildlife watching (Fig. 2). The amount of infrastructure present in each 10x10 Km cell had a positive effect on the response variable, the intensity of wildlife watching (Fig. 2 right; coefficient: 0.2, SE: 0.02, t-value: 10.4, DF: 1131, p-value < 0.001). The area of the cell that is occupied by a protected area also had a positive, but weaker, effect on the response variable (Fig. 2 left; coefficient: 0.06, SE: 0.02, t-value: 2.45, DF: 1131, p-value < 0.01), while the mean naturalness score of the area had a negative effect (Fig. 2 middle; coefficient: -0.3, SE: 0.03, t-value: -10.12, DF: 1131, p-value < 0.001).

Environmental model – Wildlife watching intensity increased as the distance from a marine Special Area of Conservation (coefficient: -0.16, adjusted SE: 0.05, z: 3.4) and the area's mean naturalness score (coefficient: -0.4, adjusted SE: 0.03, z: 13.8) decreased and as the area covered by a Country

Park (coefficient: 0.04, adjusted SE: 0.01, z: 2.7), the area covered by a Local Nature Reserve (coefficient: 0.05, adjusted SE: 0.01, z: 3.4) and the area covered by a National Park (coefficient: 0.1, adjusted SE: 0.03, z: 3.4) increased (Fig 3).

Infrastructure model – number of tourist accommodations (coefficient: 0.4, adjusted SE: 0.01, z: 25.3) and number of bust stations (coefficient: 0.07, adjusted SE: 0.01, z: 4.3) increased the intensity of wildlife watching in each 10x10 Km cell (Fig.4), while distance from an airport had a negative effect (Fig. 4; coefficient: -0.08, adjusted SE: 0.03, z: 2.7).

Species richness model – In areas where there was sufficient access for people to visit, the strongest positive effect on the intensity of wildlife watching was that of the number of species (Fig. 5; coefficient: 0.5, SE: 0.04, t-value: 13.36, DF: 760, p-value < 0.001). The other variables maintained similar effects as in the previous environmental model: distance from marine Special Areas of Conservation (coefficient: -0.08, SE: 0.03, t-value: -2.35, DF: 760, p-value < 0.05), area covered by a Local nature Reserve (coefficient: 0.03, SE: 0.01, t-value: 2.01, DF: 760, p-value < 0.05), area covered by a Country Park (coefficient: 0.04, SE: 0.02, t-value: 2.64, DF: 760, p-value < 0.05), area covered by a National Park (coefficient: 0.08, SE: 0.03, t-value: 2.63, DF: 760, p-value < 0.05), and mean naturalness score (coefficient: - 0.23, SE: 0.02, t-value: -8.58, DF: 760, p-value < 0.001).

The maps of model residuals (Fig. 6) showed some spatial patterns left unexplained by the models. The aggregated and environmental model (Fig. 6 top) seemed to be under-predicting the intensity of wildlife watching in the more populated part of Scotland, the central belt of Scotland, especially around Glasgow and Edinburgh, in the area around Inverness and the Cairngorms National Park, on the west coast and in North and South Shetland. The same models were over-predicting the intensity of wildlife watching in the Highland and in the southern regions of Ayrshire and Dumfries and Galloway. The residuals from the species richness model seemed to present a very similar pattern (Fig. 6 bottom-left). The infrastructure model provided the best fit to the data, with less over- and under-prediction compared to the other ones (Fig. 6 bottom-right).

Discussion

Interactions with wildlife were common outside PAs which have been the primary focus of studies on recreational ecosystem services so far (Balmford *et al.* 2009; De Vos *et al.* 2016; Sessions *et al.* 2016; Sonter *et al.* 2016; Baum, Cumming & De Vos 2017). Secondly, wildlife watchers in Scotland experience wildlife mostly in areas that have low conservation value, where nature is easily accessible and facilities are provided, while areas that have very high naturalness are rarely used (Figs. 2,4,6). The preferred PAs, Country Parks, Local Nature Reserves and National Parks (Fig 3), are all areas that provide opportunities for the public to enjoy nature close to where they live (Fig. S3 in Appendix S1). They were designated with the dual objective of preserving important natural and cultural heritage and providing people with recreational opportunities. As such, they are better connected to infrastructure to facilitate access. National Nature Reserves, Sites of Special Scientific Interest or Special Protection Areas, which are managed for the conservation of threatened species or important habitats, have limited infrastructure in their vicinity because of their remit. They are also not intensely used by wildlife watchers, except for marine Special Areas of Conservation, where we found a higher intensity of wildlife watching (Fig. 3). We think this is because coastal Special Areas of Conservation are very close to main cities in the East (e.g. Inverness, Dundee and Edinburgh) and to very popular tourist areas in the West (Fig. S4 in Appendix S1).

The nature that the people experience during wildlife watching activities is not the one on which conservation programmes are focussed; it is a nature with a strong human dimension. In the context of wildlife watching, wild landscapes and environmental features that are priorities for conservation do not attract recreationists as other types of low-conservation value green space do. Hence, it is important to rethink the position of green spaces in providing recreational ecosystem services. The concept that high conservation value does not enhance recreational ecosystem services was proposed by a recent study (Hornigold, Lake & Dolman 2016) where the authors compared visitation to similar natural areas inside or outside Sites of Special Scientific Interest as a proxy for the effect of

high conservation value on the likelihood of site visitation. Our study expands on this work by directly comparing the effect of infrastructure vs that of naturalness on the intensity of wildlife watching and by testing the effect of every type of PA designation at a national scale. The value of biodiversity and natural areas for recreational purposes is often highlighted to support economic investment for conservation (Walpole, Goodwin & Ward 2001; Balmford *et al.* 2009). Given the results of our study, it is likely that the value of green, and blue, spaces that are not special for threatened species and habitats is underestimated. In built areas people are attracted to sites with increased species richness (Fig. 5) and there is evidence that the mental health benefits of green spaces increase with biodiversity (Fuller *et al.* 2007; Luck *et al.* 2011). In an urban context managing for increased species richness and recreation could be a sustainable approach to meet multiple sustainable development goals (SDG 3,6,8,10,11,12,15; United Nations, 2015).

There are some limitations in the data and methodology we used in this study. While many people take photographs during recreational activities, only some post their photographs online; social media users tend to be under 35, well-educated and earn a higher income than people who do not post on social media (Lo *et al.* 2011). Besides demographic biases, not all that is experienced is posted online, which means that we are not capturing all the experiences of wildlife watchers but only those that the social media users considered worthy of sharing. There are also issues with sampling bias related to the type of social media used; in this case Flickr users tend to be more experienced nature recreationists who are nature enthusiasts compared to Instagram users who tend to be younger and more interested in charismatic megafauna and other recreational activities (Hausmann *et al.* 2017b). However, the results from a previous validation study(Mancini, Coghill & Lusseau 2018), together with other recent publications (Wood *et al.* 2013; Levin, Kark & Crandall 2015; Sessions *et al.* 2016; Hausmann *et al.* 2017b; Levin, Mark & Brown 2017), give us enough confidence that the data used in this study is a reliable proxy for wildlife watching activities in Scotland. The choice of wildlife that we focused on could also have affected the spatial patterns of wildlife watching we found. The way in which charisma is defined can influence the effect that

charismatic species have on spatial patterns of wildlife watching (Booth *et al.* 2011; Hausmann *et al.* 2017a; Arbieu *et al.* 2017). However, while seals are one of Scotland's Big Five (https://scottishwildlifetrust.org.uk/news/can-you-spot-all-of-scotlands-big-5/) and dolphins are usually considered a charismatic species, the photographs of birds are likely to include species such as golden eagles, also one of Scotland's big Five, but also species such as sparrows and blackbirds not usually considered charismatic. There are also other factors that might be important in influencing the spatial patterns of wildlife watching in Scotland, such as marketing and satisfaction. However, the limited availability of data to quantify these effects limits their inclusion in studies at this scale and resolution. Lastly, there are many more types of human interactions with nature that this study does not include, both recreational, such as outdoor sports or dog walking, and non-recreational, such as spiritual or social-cohesion experiences. Although it is possible that photographs of wildlife were taken while conducting one of these other activities, the data used in this study can only be representative of wildlife watching and, therefore, all the conclusions made are not generalizable to other types of interactions with nature. An interesting extension of this analysis would be to compare total number of Flickr photographs, to wildlife photographs. This would allow us to determine which sites are used primarily for wildlife watching, which sites are used for wildlife watching only secondarily or during other activities and which activities are mostly associated with wildlife recreation.

Conclusions

The widespread use of the Internet and social media in the nation we targeted allowed us to quantify wildlife watching activities outside PAs where visitor numbers are not usually monitored. Thanks to this wider sampling we could make inferences about drivers of the attractiveness, and hence the value, of the natural environment outside PAs. We found little overlap between areas of high conservation value and areas of high recreational value (Fig. 2- 4). This is a positive result for conservation, because it means that those environmental features that are priorities for

conservation are to some extent protected from intensive recreational activities. However, this result also questions the synergy between the goals of conservation and those of recreation. Attracting visitors to natural areas requires improving access and building infrastructure, which could compromise the integrity of the natural features that we want to protect. For example, there are 4447 species in the IUCN Red List of Threatened Species that are threatened by development of "tourism & recreation areas" (threat 1.3, IUCN 2018). In managing natural areas, it is important to find a balance between giving people the opportunity to enjoy nature and achieve conservation goals. From our results it seems that designations such as Country Parks and Local Nature Reserves achieve this balance by attracting higher intensities of nature recreational activities and leaving areas of higher conservation value such as Marine Protected Areas or Sites of Special Scientific Interest relatively undisturbed. Moreover, the notion that the general public does not value naturalness as a recreational service, as shown by our results in a wildlife watching context and by Hornigold, Lake and Dolman (2016) in a wider recreational context, is very important for environmental management. The recreational value of natural areas is often used to support public spending for their conservation (Wilkie & Carpenter 1999; Balmford *et al.* 2009), but some natural areas, those that are most important for conservation, attract lower numbers of recreationists. If this human-dominated nature is providing most recreational ecosystem services, the general public might be less inclined to support conservation actions to save more remote natural areas that they do not experience directly (Wells & Lekies 2006). PAs are under increasing pressure to provide financial justification for their existence (Walpole, Goodwin & Ward 2001) and tourism is their main tool to generate income. In the attempt to develop PAs into tourism destinations, more infrastructure might be built to attract more people, which could have negative consequences for their conservation objectives.

Authors' contributions

FM, DL and GMC conceived the ideas and designed methodology; FM collected and analysed the data and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data accessibility

All the scripts used to collect and analyse the data are available via Zenodo DOIs: 10.5281/zenodo.1126604. (Mancini 2017) and 10.5281/ZENODO.1126605 (Mancini 2018).

References

Arbieu, U., Grünewald, C., Martín-López, B., Schleuning, M. & Böhning-Gaese, K. (2017) Large mammal diversity matters for wildlife tourism in Southern African Protected Areas: Insights for management. *Ecosystem Services*.

Balmford, A., Beresford, J., Green, J., Naidoo, R., Walpole, M. & Manica, A. (2009) A global

perspective on trends in nature-based tourism. *PLoS biology*, **7**, e1000144.

Barton, K. (2016) MuMIn: Multi-Model Inference.

Baum, J., Cumming, G.S. & De Vos, A. (2017) Understanding Spatial Variation in the Drivers of

Nature-based Tourism and Their Influence on the Sustainability of Private Land Conservation.

Ecological Economics, **140**, 225–234.

Booth, J.E., Gaston, K.J., Evans, K.L. & Armsworth, P.R. (2011) The value of species rarity in

biodiversity recreation: A birdwatching example. *Biological Conservation*, **144**, 2728–2732.

Bryden, D.M., Westbrook, S.R., Burns, B., Taylor, W.A. & Anderson, S. (2010) *Assessing the Economic*

Impacts of Nature Based Tourism in Scotland. Scottish Natural Heritage Commissioned Report

No. 398.

Carver, S., Comber, L., Fritz, S., McMorran, R., Taylor, S. & Washtell, J. (2008) *Wildness Study in the*

Cairngorms National Park. Report Commissioned by the Cairngorms National Park Authority

and Scottish Natural Heritage.

Casado-Arzuaga, I., Onaindia, M., Madariaga, I. & Verburg, P.H. (2014) Mapping recreation and

aesthetic value of ecosystems in the Bilbao Metropolitan Greenbelt (northern Spain) to support

landscape planning. *Landscape Ecology*, **29**, 1393–1405.

Curtin, S. (2013) Lessons from Scotland: British wildlife tourism demand, product development and

destination management. *Journal of Destination Marketing and Management*, **2**, 196–211.

Fraley, C. & Raftery, A.E. (2002) Model-based Clustering, Discriminant Analysis and Density

Estimation. *Journal of the American Statistical Association*, **97**, 611–631.

Fraley, C., Raftery, A. e., Murphy, B. & Scrucca, L. (2012) *Mclust Version 4 for R: Normal Mixture*

Modeling for Model-Based Clustering, Classification, and Density Estimation Technical Report

No. 597.

Fuller, R.A., Irvine, K.N., Devine-Wright, P., Warren, P.H. & Gaston, K.J. (2007) Psychological benefits
of greenspace increase with biodiversity. *Biology letters*, **3**, 390–4.

Gliozzo, G., Pettorelli, N. & Haklay, M. (Muki). (2016) Using crowdsourced imagery to detect cultural
ecosystem services: a case study in South Wales, UK. *Ecology and Society*, **21**, art6.

Hausmann, A., Toivonen, T., Heikinheimo, V., Tenkanen, H., Slotow, R. & Di Minin, E. (2017a) Social
media reveal that charismatic species are not the main attractor of ecotourists to sub-Saharan
protected areas. *Scientific Reports*, **7**, 1–9.

Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V. & Di Minin, E.
(2017b) Social Media Data can be used to Understand Tourists´ Preferences for Nature-based
Experiences in Protected Areas. *Conservation Letters*, **00**, 1–10.

Hornigold, K., Lake, I. & Dolman, P. (2016) Recreational Use of the Countryside: No Evidence that
High Nature Value Enhances a Key Ecosystem Service. *PLOS ONE*, **11**, e0165043.

IUCN. (2018) The IUCN red list of threatened species, http://www.iucnredlist.org/search

Krüger, O. (2005) The role of ecotourism in conservation: panacea or Pandora's box? *Biodiversity and
Conservation*, **14**, 579–600.

Lang, D.T. & the CRAN team. (2015a) RCurl: General Network (HTTP/FTP/...) Client Interface for R. R
package version 1.95-4.7.

Lang, D.T. & the CRAN team. (2015b) XML: Tools for Parsing and Generating XML Within R and S-
Plus. R package version 3.98-1.3.

Levin, N., Kark, S. & Crandall, D. (2015) Where have all the people gone? Enhancing global
conservation using night lights and social media. *Ecological Applications*, **25**, 2153–2167.

Levin, N., Mark, A. & Brown, G. (2017) An evaluation of crowdsourced information for assessing the importance of protected areas. *Applied Geography*, **79**, 115–126.

Lo, I.S., McKercher, B., Lo, A., Cheung, C. & Law, R. (2011) Tourism and online photography. *Tourism Management*, **32**, 725–731.

Luck, G.W., Davidson, P., Boxall, D. & Smallbone, L. (2011) Relations between Urban Bird and Plant Communities and Human Well-Being and Connection to Nature. *Conservation Biology*, **25**, 816–826.

Maes, J., Hauck, J., Paracchini, M.L., Ratamäki, O., Termansen, M., Perez-Soba, M., Kopperoinen, L., Rankinen, K., Schägner, J.P., Henrys, P., Cisowska, I., Zandersen, M., Jax, K., Notte, A. La, Leikola, N., Pouta, E., Smart, S., Hasler, B., Lankia, T., Andersen, H.E., Lavalle, C., Vermaas, T., Alemu, M.H., Scholefi, P., Batista, F., Pywell, R., Hutchins, M., Blemmer, M., Fonnesbech-Wulff, A., Vanbergen, A.J., Münier, B., Baranzelli, C., Roy, D., Thieu, V., Zulian, G., Kuussaari, M., Thodsen, H., Alanen, E.-L., Egoh, B., Sørensen, P.B., Braat, L. & Bidoglio, G. (2011) *A Spatial Assessment of Ecosystem Services in Europe: Methods, Case Studies and Policy Analysis - Phase 1. Synthesis Report*.

Mancini, F. (2017) FrancescaMancini/Flickr-API: Release for publication. DOI: 10.5281/ZENODO.1126605

Mancini, F. (2018) FrancescaMancini/WildlifeWatching-SocialMedia: Quantifying wildlife watchers' preferences to investigate the overlap between recreational and conservation value of natural areas. DOI: 10.5281/ZENODO.1126605

Mancini, F., Coghill, G.M. & Lusseau, D. (2018) Using social media to quantify spatial and temporal dynamics of wildlife tourism activities. *PLOS ONE*, **13**, e0200565.

Ooms, J. (2014) The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R

Objects.

Peña, L., Casado-Arzuaga, I. & Onaindia, M. (2015) Mapping recreation supply and demand using an

ecological and a social evaluation approach. *Ecosystem Services*, **13**, 108–118.

Pinheiro, J., Bates, D., Debroy, S., Sarkar, D. & R Core Team. (2016) nlme: Linear and Nonlinear Mixed

Effects Models.

R Core Team. (2015) R: A Language and Environment for Statistical Computing. R Foundation for

Statistical Computing, Vienna, Austria.

Reed, S.E. & Merenlender, A.M. (2008) Quiet, Nonconsumptive Recreation Reduces Protected Area

Effectiveness. *Conservation Letters*, **1**, 146–154.

Sandbrook, C.G. (2010) Local economic impact of different forms of nature-based tourism.

Conservation Letters, **3**, 21–28.

Sandifer, P.A., Sutton-Grier, A.E. & Ward, B.P. (2015) Exploring connections among nature,

biodiversity, ecosystem services, and human health and well-being: Opportunities to enhance

health and biodiversity conservation. *Ecosystem Services*, **12**, 1–15.

Sessions, C., Wood, S.A., Rabotyagov, S. & Fisher, D.M. (2016) Measuring recreational visitation at

U.S. National Parks with crowd-sourced photographs. *Journal of Environmental Management*,

183, 703–711.

Sonter, L.J., Watson, K.B., Wood, S.A., Ricketts, T.H., Danielson, P. & Xian, G. (2016) Spatial and

Temporal Dynamics and Value of Nature-Based Recreation, Estimated via Social Media. *PLOS*

ONE, **11**, e0162372.

United Nations. (2015) *Transforming Our World: The 2030 Agenda for Sustainable Development*.

De Vos, A., Cumming, G.S., Moore, C.A., Maciejewski, K. & Duckworth, G. (2016) The relevance of

spatial variation in ecotourism attributes for the economic sustainability of protected areas. *Ecosphere*, **7**, e01207.

Walpole, M.J., Goodwin, H.J. & Ward, K.G.R. (2001) Pricing Policy for Tourism in Protected Areas: Lessons from Komodo National Park, Indonesia. *Conservation Biology*, **15**, 218–227.

Wells, N.M. & Lekies, K.S. (2006) Nature and the Life Course: Pathways from Childhood Nature Experiences to Adult Environmentalism. *Children, Youth and Environments*, **16**, 1–24.

Wickham, H. (2016) httr: Tools for Working with URLs and HTTP. R package version 1.1.0.

Wilkie, D.S. & Carpenter, J. (1999) Can nature tourism help finance protected areas in the Congo Basin? *Oryx*, **33**, 333–339.

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.

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

van Zanten, B.T., Verburg, P.H., Scholte, S.S.K. & Tieskens, K.F. (2016b) Using choice modeling to map aesthetic values at a landscape scale: Lessons from a Dutch case study. *Ecological Economics*, **130**, 221–231.

Data sources

Flickr Services. The App Garden. API Documentation. Available: https://www.flickr.com/services/api/

Google Maps Platform - Geo-location APIs. Google Maps Platform. Google Cloud. Available: https://cloud.google.com/maps-platform/

Digimap Ordnance Survey (2017). Strategi. Available:

https://www.ordnancesurvey.co.uk/opendatadownload/products.html#STRTGI

Scottish Natural Heritage (2017). Natural Spaces. Available: https://gateway.snh.gov.uk/natural-

spaces/index.jsp

GBIF.org (08 March 2017) GBIF Occurrence Download https://doi.org/10.15468/dl.lckabh

flickr

Number of photos

- 0 - 9
- 10 - 26
- 27 - 50
- 51 - 88
- 89 - 144
- 145 - 251
- 252 - 413
- 414 - 637
- 638 - 1143
- 1144 - 1959

Miles
0 20 40   80   120

Environment

CNTRY
LNR
WHS
NNR
SSSI
RAMSAR
SAC_L
SPA
Cairngorms NP
COUNEUR
BIOGEN
NR
Loch Lomond NP
BIOSPH
MCA
MSAC
MPA

Naturalness

High : 256

Low : 1

Species Records

Miles
0 20 40   80   120

Miles
0 20 40   80   120

Infrastructure

- Airports
- TrainStations
- CarParks
- BusStations

Motorway
A Road
B Road
Minor Road

- Hotels
- Tour Operators

Miles
0 20 40   80   120

Miles
0 20 40   80   120

Miles
0 20 40   80   120

Aggregated model

Environmental model

Infrastructure model

Species richness model

Residuals

1

0

-1