Using social media to quantify spatial and temporal dynamics of wildlife tourism activities

Running title: Wildlife watching and social media

Word count: 5080

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. Wildlife tourism is a profitable industry that can affect the conservation status of targeted populations. Tourist behaviour plays a key role in the success of sustainable management strategies. Traditionally, visitor numbers are obtained through surveys, which are expensive and limited in coverage and resolution. Recently, data from social media have been used to quantify visitation. However, we do not know at which scale the use of this proxy is appropriate, especially outside protected areas. Here, we validated for the first time the use of a dataset obtained from the photo-sharing website Flickr as a proxy for wildlife tourism in Scotland.
2. We used photos uploaded on Flickr to estimate visitation in the Cairngorms National Park (CNP) and compared this dataset to a time series of visitor numbers obtained from the CNP authority. Then, we compared the spatial distribution of photographs of birds, seals, whales and dolphins taken in Scotland and uploaded on Flickr to a dataset obtained from a 2014-15 Scotland-wide wildlife tourism survey.
3. Wavelet analysis showed that the two time series are significantly correlated and synchronised. The results of the spatial validation showed that both the presence and the number of pictures uploaded on Flickr are correlated to survey data at different scales. Finally, kernel density maps of the wildlife pictures revealed spatio-temporal trends in wildlife watching hotspots that confirmed the validity of this dataset.
4. Both temporal and spatial trends in the distribution of pictures uploaded on Flickr displayed similar patterns to those observed in datasets obtained using traditional methods. This was true for different spatial scales and for locations inside and outside protected areas. Therefore, this method allowed us to quantify visitation even in areas that are not monitored. In conclusion, despite limitations and challenges, data from social media offer great potential to study wildlife tourism at different spatial and temporal scales.

Keywords: conservation, Moran eigenvector spatial filtering, big data, Flickr, time series analysis, wildlife watching

1. Introduction

Recreation is a key cultural ecosystem service provided by nature. Tourism is often a primary income for local communities (Curtin 2003; Silva 2013), it can dominate national economies and play a key role in nations' macroeconomics (O'Connor et al., 2009). Nature-based tourism involves interactions with the natural environment and it represents a big component of global recreation. This type of recreation is an important issue because of its economic contribution to conservation (Gossling 1999), the health benefits it brings to humans (Russell *et al.* 2013) and its role in alleviating poverty (Ferraro & Hanauer 2014).Nature-based tourism, such as wildlife watching, was initially welcomed by conservation and environmental organisations as an eco-friendly alternative to other consumptive activities, such as hunting and fishing (Tisdell & Wilson 2002). However, there is growing evidence that these activities, if not managed properly, can have negative effects on the environment (McClung *et al*. 2004; Reed & Merenlender 2008; Pirotta & Lusseau 2015). Quantifying temporal and spatial patterns of wildlife watching activities can help management by identifying areas that are under too much pressure from the tourism industry and areas that could be sustainably developed to redistribute this pressure.

In cases where wildlife watching activities require specialised infrastructures (e.g. whale watching boat trips) or are performed inside well-monitored protected areas (Balmford *et al.* 2015), visitor numbers can be recorded relatively easily. When access to a nature-based recreation site is not monitored and wildlife watching activities are conducted independently in more natural and remote areas, data on visitation can be obtained through surveys, which are very expensive, time consuming and limited in coverage. Bigger datasets on tourism (e.g. airport arrivals) are useful for global scale inferences, but it is difficult to distinguish wildlife tourism from general tourism. As a result, detailed data of spatial and temporal patterns of wildlife watching activities are rare. The widespread use of the Internet and the popularity of smartphones and social media websites offers ecologists the opportunity to use the data generated by their billions of users (Worthington *et al.* 2012; Leighton *et* *al.* 2016). Following promising results from a first study validating the use of data from the photo-sharing website Flickr as a proxy for visitation (Wood *et al.* 2013), recent studies have started to use data from social media to make inferences on nature-based tourism dynamics (Keeler *et al.* 2015; Levin *et al.* 2015). However, only average numbers of pictures taken inside national parks and paid attractions were validated as a proxy (Wood *et al.* 2013), without testing whether temporal or spatial trends were reproducing patterns observed in datasets collected in more traditional ways. The validation was conducted on a global scale and only for well-monitored tourism destinations. Therefore, we still don't know if this proxy is valid to quantify this ecosystem service in more natural and remote areas and at a finer resolution.

Here we aimed to assess the spatial and temporal scale at which such methods could be used, using unique data available in Scotland. Bird and wildlife watching contributes roughly £127 million per year to the Scottish economy (Bryden *et al.* 2010), generating 2763 full-time equivalent (FTE) jobs (Blake *et al.* 2010). The main groups of charismatic wildlife that attract tourists to Scotland are birds, seals, whales and dolphins (Curtin 2013). We tested whether temporal patterns of pictures taken at a nature-based tourism site (the Cairngorms National Park – CNP) and posted on Flickr were reproducing patterns found in a dataset obtained from the CNP authority. We then tested the validity of the spatial distribution of geotagged pictures of wildlife posted on Flickr as a proxy for spatial patterns of wildlife watching activities in Scotland using a dataset collected through a Scotland-wide wildlife tourism survey (Land Use Consultants 2016). Finally, we use this validation to assess the geography of wildlife tourism in Scotland.

2. Materials and methods

2.1. Data collection

The CNP visitation dataset contained the number of visitor days to the park (a person spending at least a portion of a day at the CNP) per month from 2009 to 2014. These visitor days were estimated by the STEAM model (STEAM visitor days – SVD) (Global Tourism Solutions Ltd. 2006), which uses data on accommodation occupancy rates, visitor surveys, number of visitors to paid attractions and other data supplied by the CNP authority. Data from Flickr were collected through Flickr application programming interface (API https://www.flickr.com/services/api/) and R (R Core Team 2015), using the packages RCurl (Lang & the CRAN team 2015; version 1.95.4.7), XML (Lang & the CRAN team 2015b; version 3.98.1.3) and httr (Wickham 2016; version 1.1.0). Dates and geographic coordinates associated with the pictures were used to select only those taken in the CNP between 2009 and 2014 (R code available at https://github.com/FrancescaMancini/Flickr-API). We downloaded different metadata associated with the photos: picture and photographer ID, the date when the photo was taken, the geographic coordinates of where the picture was taken and the tags. In order to avoid duplicates, we used the combination of photographer-ID and date to delete multiple photos from the same user on the same day. Therefore, the number of data points in the dataset represents the number of Flickr visitor days (FVD).

We also used keywords (bird, seal, dolphin and whale) to query Flickr API and select all the photos of the main groups of wildlife sought by tourists taken in Scotland and we downloaded the same metadata as for the CNP photos. Again, we deleted multiple photos from the same user on the same day and, using the tags associated with the photos, we eliminated all the pictures that were not relevant, such as pictures of statues or paintings and pictures taken in zoos. We also obtained data from the Scottish marine recreation and tourism survey (Land Use Consultants 2016). The dataset (available at http://live-marinedatascotland.getnucivic.com/dataset/scottish-marine-recreation-and-tourism-survey-2015) contained spatial information on trips to coastal areas in Scotland made by the survey respondents between October 2014 and October 2015 to conduct bird or wildlife watching activities. From the dataset obtained from Flickr we selected only the pictures taken in that time period. A buffer of 2 Km inside the coastline was created in ArcGIS (ESRI 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute) to select only pictures taken in a coastal environment. In ArcGIS we created 3 grids (5 Km, 10 Km and 20 Km) and counted the number of FVD and of visitor days from the survey (SuVD) in each cell.

2.2. Temporal validation

In order to test whether the temporal patterns shown by the two time series were similar, we used Wavelets Analysis (WA) (Cazelles *et al.* 2008). WA decomposes the variance of the time series in its oscillating components, thus detecting significant periodicities. The advantage of this method compared to other spectral decompositions is that WA does not assume stationarity of the time series, but it allows the main frequency component to change through time by estimating the signal's spectral characteristics as a function of time (the wavelet power spectra). We used the Morlet mother wavelet to perform this decomposition. This continuous and complex function allows the extraction of both time-dependent amplitude and phase of the time series. WA also allows the analysis of patterns of covariation between two time series. We compared the time series of SVD and FVD using the wavelet coherence, which identifies the linear correlation between two signals. In order to assess statistical significance of the association between the two time series we used a random noise resampling scheme, where the null hypothesis tested is that the association between the two signals is not different from that expected by chance alone (Ménard *et al.* 2007). We also computed the phase difference to test whether the two time series were synchronised or out of phase. This analysis was conducted in R using the package WaveletComp version 1.0 (R code available at https://github.com/FrancescaMancini/Flickr-Statistical-Analysis) (Roesch & Schmidbauer 2014).

2.3. Spatial validation

We compared the spatial distribution of wildlife tourists obtained from Flickr and the one obtained from the survey at three different spatial scales: 5 Km, 10 Km and 20 Km. We fitted Generalised Linear Models (GLM) to the three datasets using the number of FVD in each cell as the response variable and the number of SuVD in each cell as explanatory variable. The data contained a high number of zeros, so we fitted a GLM with a binomial error distribution to the presence/absence of FVD in a cell and a GLM with a negative binomial error distribution to the number of FVD present.

This allowed us to test two hypotheses: 1) the probability of finding at least one picture by one user on Flickr is higher for areas with higher SuVD 2) the number of users posting pictures on Flickr is associated to the number of SuVD in the same area. Since densely populated areas tend to have a higher average number of Flickr users (Fig. A1), we used population abundance for each grid cell as model weights. The residuals of the GLMs were spatially correlated and directional variograms, estimated using R package gstat version 1.1.0 (R code available at https://github.com/FrancescaMancini/Flickr-Statistical-Analysis) (Pebesma 2004), showed that the spatial autocorrelation was anisotropic (Fig. A2). We therefore used spatial eigenvector mapping (SEVM) to derive explanatory variables for the GLMs (Griffith & Peres-Neto 2006). This method decomposes a matrix of relationships between data points into eigenvectors that capture spatial effects. The eigenvectors can then be included as explanatory variables in the GLM to remove the effect of spatial autocorrelation on the analysis. First we used a Delaunay triangulation of the grid cells centres to define neighbours, to which we assigned row-standardised spatial weights. A set of Moran eigenvectors (ME) were then calculated from these weights and those that best reduced the spatial autocorrelation of residuals were selected and included as spatial covariates in the GLMs. We used AIC to select only the ME that improved the explanatory power of the model. This analysis was conducted in R using the package spdep version 0.5.92 (R code available at https://github.com/FrancescaMancini/Flickr-Statistical-Analysis) (Bivand *et al.* 2013; Bivand & Piras 2015).

2.4. Wildlife watching hotspots

We investigated spatio-temporal patterns in wildlife watching hotspots by producing density maps of the geotagged pictures posted on Flickr. We used a two-dimensional kernel density estimator from the R package ggplot2 (Wickham 2009), where the bandwidth is calculated using Scott's rule of thumb (Scott 1992) (R code available at https://github.com/FrancescaMancini/Flickr-Statistical-Analysis).

3. Results

In total, we downloaded metadata on 29,336 pictures (4699 unique FVD) taken in the CNP between 2009 and 2014 and uploaded on Flickr.

The query to Flickr API returned 92,229 results (36,998 FVD) for pictures with the word "bird" in tags, title or description taken in Scotland between 2005 and 2015. From the search with the word "seal" we obtained 2571 FVD and from the search with the word "dolphin" and the word "whale" we obtained 1634 FVD.

3.1. Temporal validation

The power spectra of the Flickr and CNP survey time series were very similar, with significant 12-month cycles throughout the 5-year period (Fig. A3a-b), showing that visitation has a strong seasonal component in both measures. This similarity was supported by the strong coherence between the Flickr and the empirical time series around the same 12-month oscillations, which was also constant through time (Fig. 1a). The phase difference for this 12-month cycle was constant around 0 (Fig.1b), indicating that the two time series were synchronised. Cross-correlation was also significant at a period of 6 and 3 months, but this was not consistent throughout the time period (Fig.1a).

3.2. Spatial validation

The survey dataset indicated that the areas more intensely used for birds and wildlife watching are around the west coast of Scotland, the Moray Firth, the Firth of Forth and the Tay estuary (Fig. A4). Spatial distribution of the pictures from Flickr also identified the last three as areas of high visitation (Fig. A1). The Flickr dataset also contained pictures taken on the west coast of Scotland, but not in the same density as shown by the survey dataset. This area of the country is not highly populated, so there might be a certain bias in the number of Flickr users uploading pictures. When the Flickr data was normalised by population size, the west coast appeared as an area of high visitation (Fig. A5).

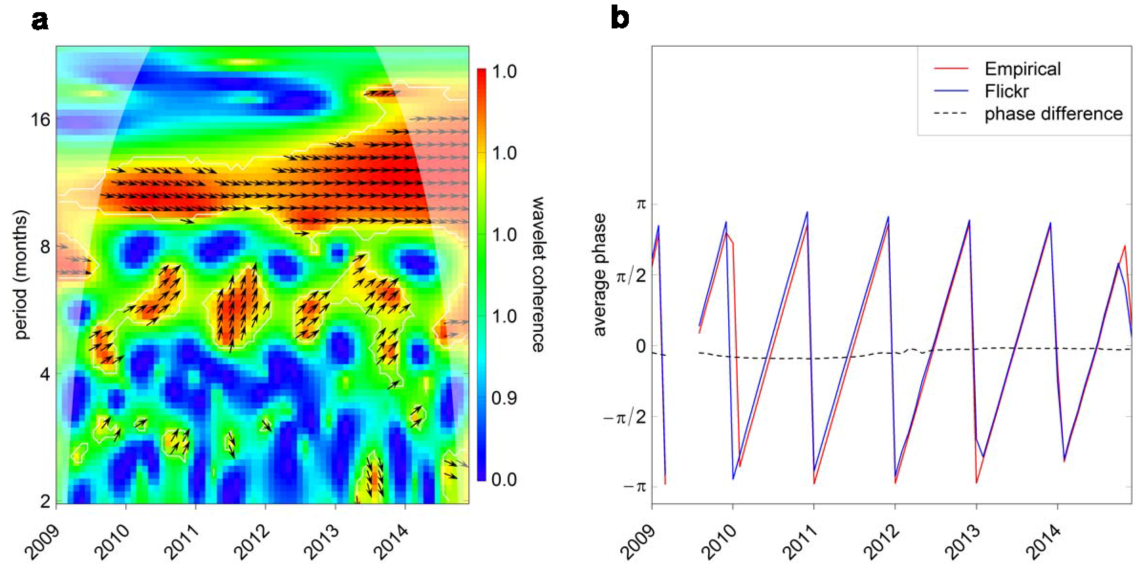

Fig.1. Results of wavelet analysis. a) Wavelet coherence between the two time series. Colour code from dark blue (low values) to red (high values). The arrows indicate synchrony of the two time series: arrows pointing to the right means the oscillations are synchronised. Arrows are only plotted within white contour lines that indicate significance. The shaded area near the edges in the graphs is the cone of influence, and indicates the range of the graph where the results are not reliable because of edge effects. b) Phases of the oscillations of the two time series (blue and red lines) computed in the 8-16 periodic band where there is significant correlation. The dotted line is the phase difference.

The number of SuVD was associated with the presence of pictures posted on Flickr by a user for a certain area (Fig. 2). This was true for each of the spatial scales tested (20km: coefficient = 0.37, *SE* = 0.04, *Z* = 9.5, *p-value* < 0.001; 10km: estimate = 0.36, SE = 0.02, Z = 14.7, *p-value* < 0.001; 5km: estimate = 0.28, *SE* = 0.001, *Z* = 21.1, *p-value* < 0.001). The higher the number of visitors captured by the survey, the higher the probability of finding photos on Flickr taken in that area.

The number of SuVD was related to the number of Flickr pictures at 10km and 20km resolution but not 5km (Fig. 3; 20km: coefficient = 0.1, *SE* = 0.01, *Z* = 8.1, *p-value* < 0.001; 10km: coefficient = 0.02, *SE* = 0.001, *Z* = 3.4, *p-value* < 0.001; 5km: coefficient = 0.01, *SE* = 0.006, *Z* = 1.7, *p-value* > 0.05). A higher number of visitors captured by the survey corresponds to a higher number of FVD in the same grid cell.

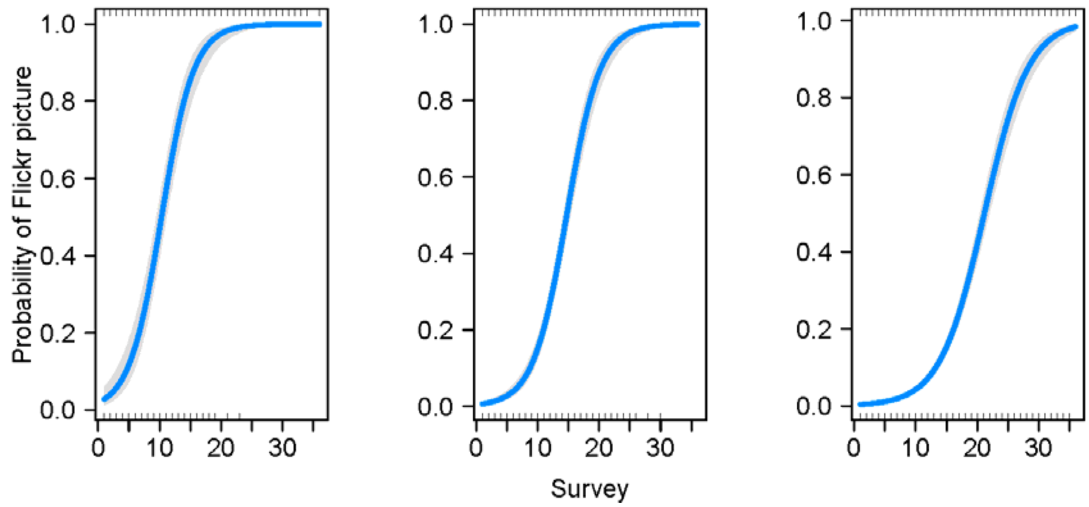

Fig. 2. Results of binomial GLMs. Left: results at the 20 Km resolution; centre: results at the 10Km resolution; right: results at the 5Km resolution. Predictions from the models (blue line) are plotted on the response scale with confidence intervals (shaded areas around the prediction curve). Tick marks on the x-axis represent data.

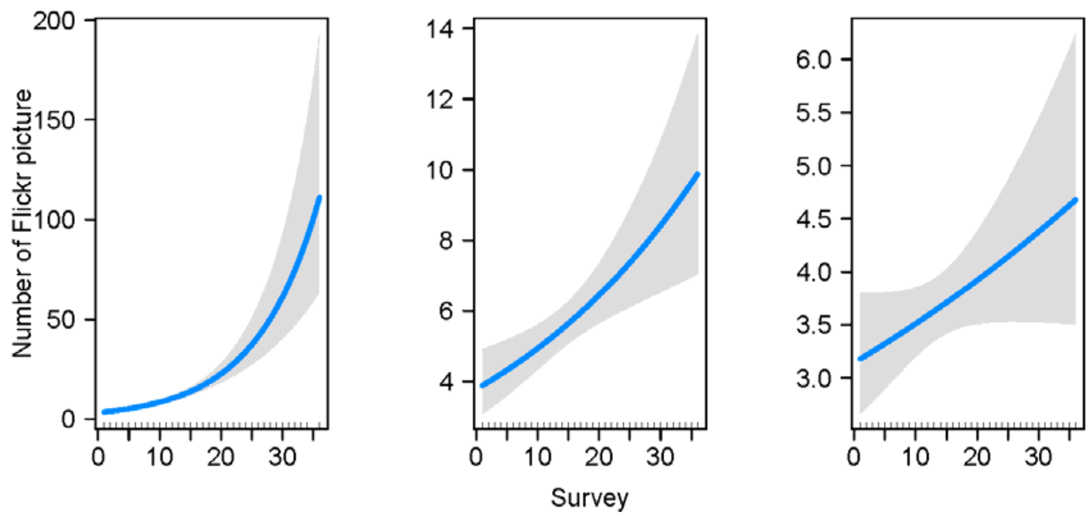

Fig. 3. Results of negative binomial GLMs. Left: results at the 20 Km resolution; centre: results at the 10Km resolution; right: results at the 5Km resolution. Predictions from the models (blue line) are plotted on the response scale with confidence intervals (shaded areas around the prediction curve). Tick marks on the x-axis represent data.

3.3. Wildlife watching hotspots

The density maps revealed spatio-temporal patterns of wildlife watching hotspots in Scotland. Birdwatching (Fig. 4) seems to be concentrated around Edinburgh and Glasgow, however when FVD in Edinburgh and Glasgow were excluded from the dataset, hotspots in the Moray Firth, Orkney, Shetland, the Isle of Mull and the North-West coast started to appear (Fig. A6). A seasonal plot of the same pictures shows that this high density around urban areas spreads out towards the West coast and the islands during spring and summer (Fig. A7).

Seal watching seems to be concentrated initially around the west coast, the Firth of Forth and Shetland (Fig. 5). It is worth noticing the appearance of another hotspot from 2008 corresponding to Newborough in Aberdeenshire, becoming very important after 2011.

Dolphin and whale watching maps showed a consistent hotspot at Chanonry Point in the Moray Firth (Fig. 6), with the appearance of a second hotspot from 2013 in Aberdeen.

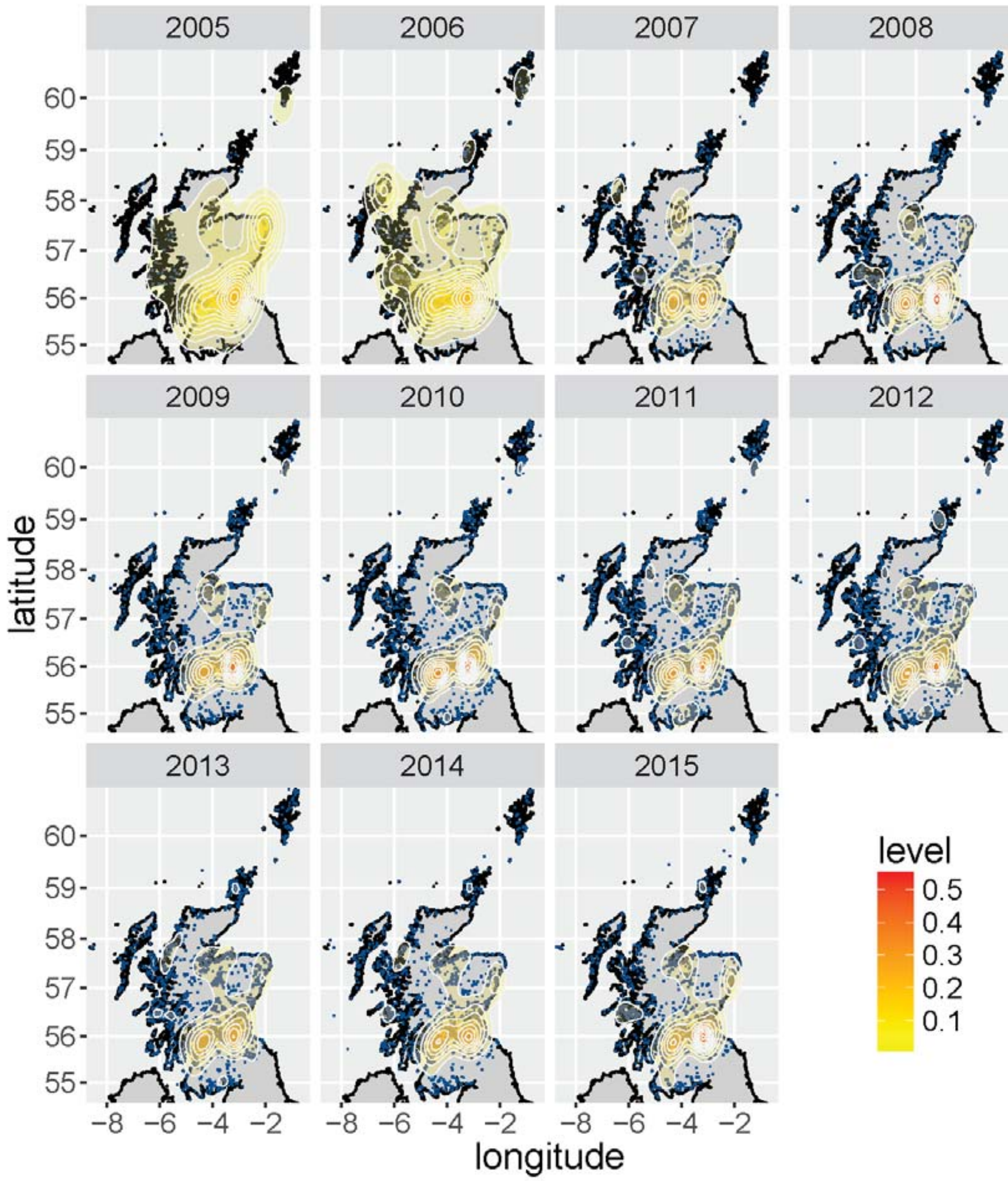

Fig.4. Bird watching density maps. Each panel represents the density of FVD in a different year, from 2005 to 2015. The blue dots on the maps are the data. Different colours represent different density levels, from low (yellow) to high (red). White lines represent contour lines for different density levels.

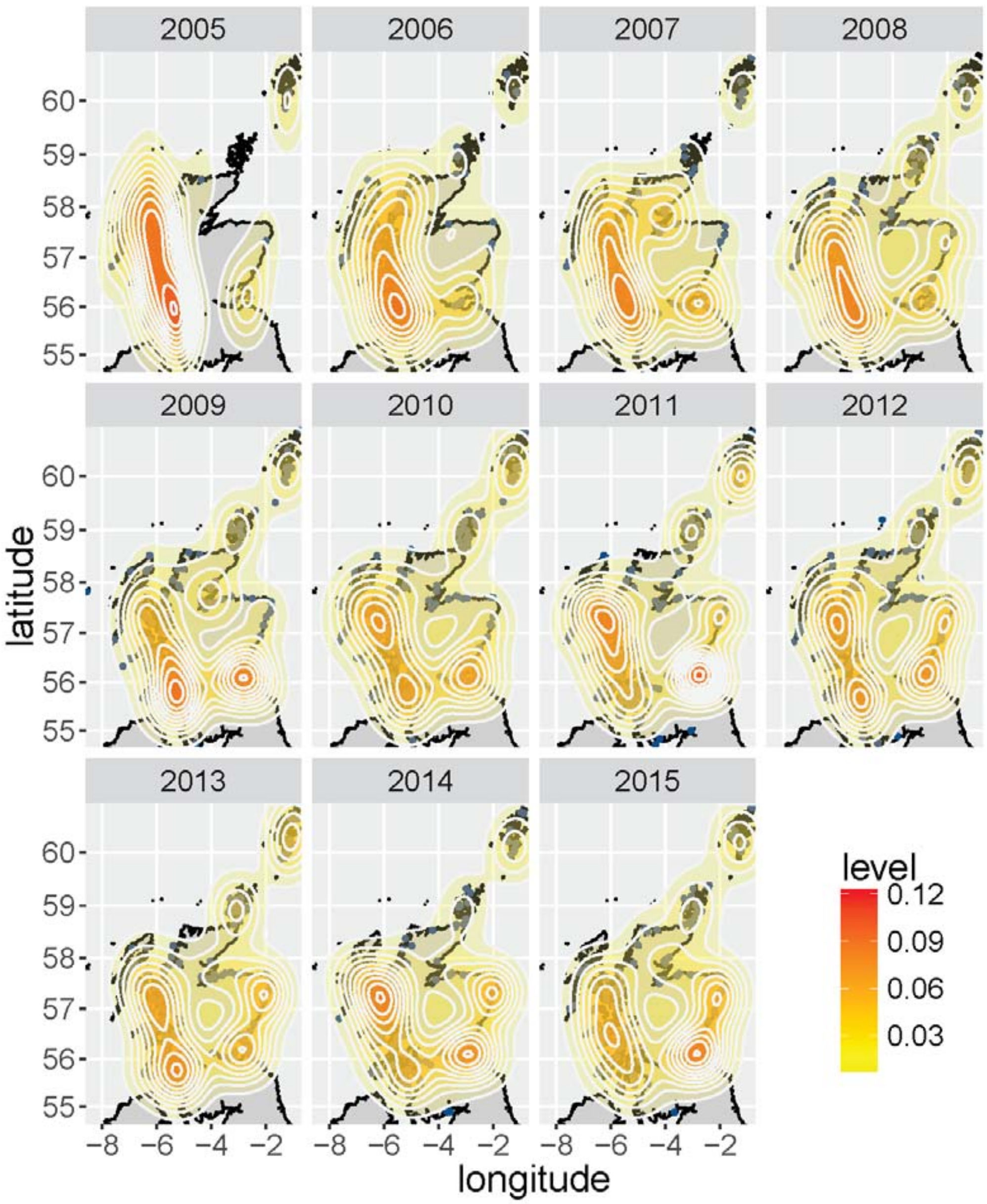

Fig.5. Seal watching density maps. Each panel represents the density of FVD in a different year, from 2005 to 2015. The blue dots on the maps are the data. Different colours represent different density levels, from low (yellow) to high (red). White lines represent contour lines for different density levels.

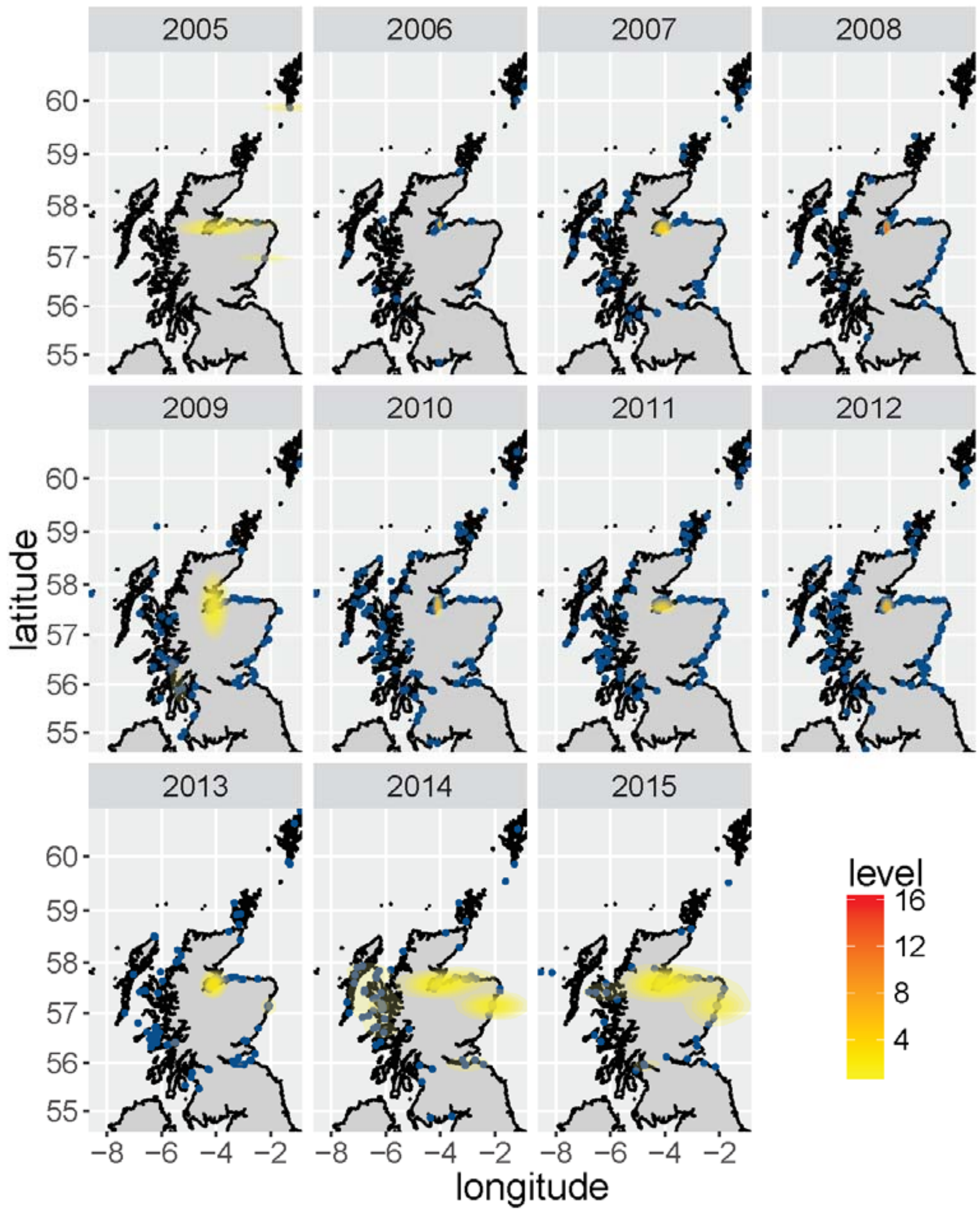

Fig.6. Dolphin and whale watching density maps. Each panel represents the density of FVD in a different year, from 2005 to 2015. The blue dots on the maps are the data. Different colours represent different density levels, from low (yellow) to high (red).

4. Discussion

Nature-based recreation is difficult to measure, especially when it does not require the use of infrastructures and it is performed in remote areas. Monitoring tourism in protected areas can also be difficult when access to the park is not subject to an entrance fee. The popularity of smartphones, GPS devices and social media is allowing people to generate large amounts of spatial and temporal data (Worthington *et al.* 2012; Leighton *et al.* 2016), recording where they go and what they see.

We show for the first time that we can use social media to quantify temporal and spatial patterns of wildlife tourism at spatial and temporal scales that are relevant to management actions (Stevens *et al.* 2007) (see Appendix B for further analysis and validation of temporal trends). For the first time, time series obtained from the website Flickr were compared to time series obtained from traditional survey and modelling methods. The two datasets revealed the same seasonal trend, with peaks of visitation during the summer and low numbers of visitors during the winter months. We also found a strong correlation between the seasonal oscillations of the two time series. In areas where a survey detected a high number of visitors it was more likely to find at least one picture posted on Flickr. This was true for different spatial resolutions, from 5 to 20km. Number of users taking pictures is also related to the number of visitors obtained from a survey at a resolution of 20 to 10 Km. We could not estimate the number of visitors from the number of Flickr users at a 5km resolution.

Spatio-temporal trends of wildlife watching hotspots confirm the validity of FVD as a proxy for visitation. The majority of pictures of birds that were taken around Edinburgh and Glasgow, were taken inside greenspaces and urban parks, such as Hogganfield Park in Glasgow and Bawsinch and Duddingston Scottish Wildlife Trust reserve in Edinburgh (Fig. A1). The density maps also detected a change in the bird watching hotspots with seasons (Fig. A6 and A7), consistent with a movement from the area around Edinburgh and Glasgow to more remote areas on the west coast, the Moray Firth and the islands. We were therefore able to capture the movement of people towards tourism destination (Blake *et al.* 2010; Land Use Consultants 2016). The seal hotspot map (Fig. 5) reveals high activity in the Firth of Forth, Tay estuary and the west coast where special areas of conservation (SACs) and haul out sites are present for both grey and harbour seals (Morris *et al.* 2014). This map also shows the appearance of a seal watching point in Newborough after 2008. This site now holds 26% of grey seals (*Halichoerus grypus*) and 1% of harbour seals (*Phoca vitulina*) in the East Coast of

Scotland Seal Management Area and has recently been proposed as a new designated haul-out site to protect the seals (Marine Scotland 2015). The whale and dolphin watching density map (Fig. 6) also reveals the emergence of a dolphin watching hotspot: Aberdeen harbour. The hotspot only appears in 2013, after the launch by the Royal Society for the Protection of Birds (RSPB) of dolphin watching events from Aberdeen harbour as part of the "Dates with nature" projects (http://www.rspb.org.uk/discoverandenjoynature/seenature/datewithnature/details.aspx?id=340366). This result indicates that such organised events can attract tourists and create a hotspot, offering an opportunity for managers to shift tourists' attention from destination that are unsustainably used to unexploited ones.

Data from social media still presents some limitations that need to be acknowledged, some of which also apply to traditional sampling methods (Wood et al., 2013; Li et al., 2013). Also, there is a bias resulting from densely populated areas having more Flickr users than sparsely populated ones (Levin et al., 2015; Fig. A1). Different species of wildlife may be more or less suited to be photographed: for instance, there are far more pictures of birds than dolphins uploaded on Flickr, partly due to the fact that encounters with dolphins are less common than encounters with birds, and partly to the fact that taking a picture of a dolphin is more difficult than taking a picture of a bird and it might require specialised equipment. Therefore, it might not be possible to compare volume of tourism dedicated to different species. Furthermore, the perceived value of a trip may influence whether an individual takes or shares photographs, producing a bias against images from visitors who visit areas closer to their home (Wood *et al.* 2013). Half of the respondents to the marine recreation and tourism survey lived within one mile of the coast (Land Use Consultants 2016) and they might have reported using an area where they would not normally take pictures because of its proximity to home. This could explain some of the differences between the two datasets.

In conclusion, despite limitations, the number of geotagged pictures uploaded on Flickr can be used as a proxy for wildlife watching activities at spatial and temporal scales that are relevant for ecosystem management in global regions where this social media is prevalently used. This opens new avenues to study tourist behaviour and decisions. The fact that we can use this data at a scale as fine as 10 Km means that we can now make more precise inference on tourists' preferences on larger areas. This information has also implications for wildlife tourism management and conservation of targeted species. First, we can now easily and cheaply quantify wildlife tourism in areas that are not monitored, allowing us to assess whether some areas, therefore wildlife populations, are receiving too much pressure from tourism. Secondly, organised events such as the RSPB "Date with Nature" can attract tourists and create a recreational hotspot. This would be a good strategy to redirect recreational activities by moving them away from overcrowded sites to unexploited ones, thus relieving pressure on those wildlife populations that are overexploited.

Acknowledgements

This work was funded by the University of Aberdeen and Scottish Natural Heritage (SNH) and their support is gratefully acknowledged. We thank MASTS (the Marine Alliance for Science and Technology for Scotland) for their role in funding this work and B. Leyshon and F. Manson (SNH) for fruitful discussion. All the data used in this study are publicly available and were collected following terms and conditions of the data providers. The data are anonymised and we did not collect any personal information.

Data accessibility

All the data and R scripts are available at https://github.com/FrancescaMancini/Flickr-API and https://github.com/FrancescaMancini/Flickr-Statistical-Analysis.

References

Balmford, A., Green, J.M.H., Anderson, M., Beresford, J., Huang, C., Naidoo, R., Walpole, M. & Manica, A. (2015). Walk on the Wild Side: Estimating the Global Magnitude of Visits to Protected Areas. *PLoS Biology*, **13**, 1–6.

Bivand, R., Hauke, J. & Kossowski, T. (2013). Computing the Jacobian in Gaussian spatial autoregressive models: An illustrated comparison of available methods. *Geographical Analysis*, **45**, 150–179.

Bivand, R. & Piras, G. (2015). Comparing Implementations of Estimation Methods for Spatial Econometrics. *Journal of Statistical Software*, **63**, 1–36.

Blake, A., Curtin, S., Richards, S., Vaughan, R., Edwards, J. & Fletcher, J. (2010). *The Economic Impact of Wildlife Tourism in Scotland. Report by the International Centre for Tourism and Hospitality Research Bournemouth University for Scottish Government Social Research*.

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.*

Cazelles, B., Chavez, M., Berteaux, D., Ménard, F., Vik, J.O., Jenouvrier, S. & Stenseth, N.C. (2008). Wavelet analysis of ecological time series. *Oecologia*, 287–304.

Curtin, S. (2013). Lessons from Scotland: British wildlife tourism demand, product development and destination management. *Journal of Destination Marketing and Management*, **2**, 196–211.

Curtin, S. (2003). Whale-Watching in Kaikoura: Sustainable Destination Development? *Journal of Ecotourism*, **2**, 173–195.

ESRI. (2013). ArcGIS Desktop: Release 10.2. Redlands CA: Environmental Systems Research Institute.

Ferraro, P.J. & Hanauer, M.M. (2014). Quantifying causal mechanisms to determine how protected areas affect poverty through changes in ecosystem services and infrastructure. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 4332–4337.

Global Tourism Solutions Ltd. (2006). Overview of STEAM, GTS. URL http://www.globaltourismsolutions.co.uk/steam-model

Gossling, S. (1999). Ecotourism : a means to safeguard biodiversity and ecosystem functions ? *Ecological Economics*, **29**, 303–320.

Griffith, D.A. & Peres-Neto, P.R. (2006). Spatial Modeling in Ecology : the Flexibility of Eigenfunction Spatial Analyses. *Ecology*, **87**, 2603–2613.

Keeler, B.L., Wood, S.A., Polasky, S., Kling, C., Filstrup, C.T. & Downing, J.A. (2015). Recreational demand for clean water: Evidence from geotagged photographs by visitors to lakes. *Frontiers in Ecology and the Environment*, **13**, 76–81.

Land Use Consultants. (2016). Scottish Marine Recreation and Tourism Survey 2015.

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

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. URL https://cran.r-project.org/package=XML.

Leighton, G.R.M., Hugo, P.S., Roulin, A. & Amar, A. (2016). Just Google it: Assessing the use of Google Images to describe geographical variation in visible traits of organisms. *Methods in Ecology and Evolution*.

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.

Marine Scotland. (2015). *Consultation on Possible Designation of a Seal Haul-out Site. Proposal in respect of the Ythan Estuary*.

McClung, M. R., Seddon, P. J., Massaro, M. & Setiawan, A. N. (2004). Nature-based tourism impacts on yellow-eyed penguins Megadyptes antipodes: does unregulated visitor access affect fledging weight and juvenile survival? *Biological Conservation*, **119**, 279-285.Ménard, F., Marsac, F., Bellier, E. & Cazelles, B. (2007). Climatic oscillations and tuna catch rates in the

Indian Ocean: A wavelet approach to time series analysis. *Fisheries Oceanography*, **16**, 95–104.

Morris, C., Duck, C., Lonergan, M., Baxter, J., Middlemas, S. & Walker, I. (2014). *Method used to identify key seal haul-out sites in Scotland for designation under the Marine (Scotland) Act Section 117*.

Pebesma, E.J. (2004). Multivariable geostatistcs in S: the gstat package. *Computers & Geosciences*, **30**, 683–691.

Pirotta, E. & Lusseau, D. (2015). Managing the wildlife tourism commons. *Ecological Applications*, **25**, 729–741.

R Core Team. (2015). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.r-project.org/.

Reed, S.E. & Merenlender, A.M. (2008). Quiet, Nonconsumptive Recreation Reduces Protected Area Effectiveness. *Conservation Letters*, **1**, 146–154.

Roesch, A. & Schmidbauer, H. (2014). WaveletComp: Computational Wavelet Analysis. R package version 1.0. URL https://cran.r-project.org/package=WaveletComp.

Russell, R., Guerry, A.D., Balvanera, P., Gould, R.K., Basurto, X., Chan, K.M. a., Klain, S., Levine, J. & Tam, J. (2013). *Humans and Nature: How Knowing and Experiencing Nature Affect Well-Being*.

Scott, D.W. (1992). *Multivariate Density Estimation: Theory, Practice and Visualization* (D.W. Scott, Ed.). John Wiley & Sons, Inc., Hoboken, NJ, USA.

Silva, L. (2013). How ecotourism works at the community-level: the case of whale-watching in the Azores. *Current Issues in Tourism*, **18**, 196–211.

Stevens, C.J., Fraser, I., Mitchley, J. & Thomas, M.B. (2007). Making ecological science policy-relevant: Issues of scale and disciplinary integration. *Landscape Ecology*, **22**, 799–809.

Tisdell, C. & Wilson, C. (2002). Ecotourism for the survival of sea turtles and other wildlife.

Biodiversity and Conservation, **11**, 1521–1538.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer Internatinal Publishing.

Wickham, H. (2016). httr: Tools for Working with URLs and HTTP. R package version 1.1.0. URL https://cran.r-project.org/package=httr.

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.

Worthington, J.P., Silvertown, J., Cook, L., Cameron, R., Dodd, M., Greenwood, R.M., McConway, K. & Skelton, P. (2012). Evolution MegaLab: a case study in citizen science methods. *Methods in Ecology and Evolution*, **3**, 303–309.