

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

2

3 *Running title:* Wildlife watching and social media

4 *Word count:* 5080

5 *Authors:* Francesca Mancini^a, George M. Coghill^b, David Lusseau^a

6 ^aInstitute of Biological and Environmental Sciences, University of Aberdeen,
7 Zoology Building, Tillydrone Avenue, Aberdeen AB24 2TZ, UK.

8 ^bDepartment of Computing Science, University of Aberdeen, Meston
9 Building, Meston Walk, Aberdeen AB24 3UE, UK.

10 *Corresponding author:* Francesca Mancini

11 Room 418, Zoology Building, Tillydrone Avenue, Aberdeen AB24 2TZ, UK.
12 Telephone: 01224 274106. Email: r03fm14@abdn.ac.uk

13

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

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

40 1. Introduction

41 Recreation is a key cultural ecosystem service provided by nature. Tourism is often a primary income
42 for local communities (Curtin 2003; Silva 2013), it can dominate national economies and play a key
43 role in nations' macroeconomics (O'Connor et al., 2009). Nature-based tourism involves interactions
44 with the natural environment and it represents a big component of global recreation. This type of
45 recreation is an important issue because of its economic contribution to conservation (Gossling
46 1999), the health benefits it brings to humans (Russell *et al.* 2013) and its role in alleviating poverty
47 (Ferraro & Hanauer 2014). Nature-based tourism, such as wildlife watching, was initially welcomed
48 by conservation and environmental organisations as an eco-friendly alternative to other
49 consumptive activities, such as hunting and fishing (Tisdell & Wilson 2002). However, there is
50 growing evidence that these activities, if not managed properly, can have negative effects on the
51 environment (McClung *et al.* 2004; Reed & Merenlender 2008; Pirotta & Lusseau 2015). Quantifying
52 temporal and spatial patterns of wildlife watching activities can help management by identifying
53 areas that are under too much pressure from the tourism industry and areas that could be
54 sustainably developed to redistribute this pressure.

55 In cases where wildlife watching activities require specialised infrastructures (e.g. whale watching
56 boat trips) or are performed inside well-monitored protected areas (Balmford *et al.* 2015), visitor
57 numbers can be recorded relatively easily. When access to a nature-based recreation site is not
58 monitored and wildlife watching activities are conducted independently in more natural and remote
59 areas, data on visitation can be obtained through surveys, which are very expensive, time consuming
60 and limited in coverage. Bigger datasets on tourism (e.g. airport arrivals) are useful for global scale
61 inferences, but it is difficult to distinguish wildlife tourism from general tourism. As a result, detailed
62 data of spatial and temporal patterns of wildlife watching activities are rare. The widespread use of
63 the Internet and the popularity of smartphones and social media websites offers ecologists the
64 opportunity to use the data generated by their billions of users (Worthington *et al.* 2012; Leighton *et*

65 *al.* 2016). Following promising results from a first study validating the use of data from the photo-
66 sharing website Flickr as a proxy for visitation (Wood *et al.* 2013), recent studies have started to use
67 data from social media to make inferences on nature-based tourism dynamics (Keeler *et al.* 2015;
68 Levin *et al.* 2015). However, only average numbers of pictures taken inside national parks and paid
69 attractions were validated as a proxy (Wood *et al.* 2013), without testing whether temporal or
70 spatial trends were reproducing patterns observed in datasets collected in more traditional ways.
71 The validation was conducted on a global scale and only for well-monitored tourism destinations.
72 Therefore, we still don't know if this proxy is valid to quantify this ecosystem service in more natural
73 and remote areas and at a finer resolution.

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

85 **2. Materials and methods**

86 *2.1. Data collection*

87 The CNP visitation dataset contained the number of visitor days to the park (a person spending at
88 least a portion of a day at the CNP) per month from 2009 to 2014. These visitor days were estimated
89 by the STEAM model (STEAM visitor days – SVD) (Global Tourism Solutions Ltd. 2006), which uses

90 data on accommodation occupancy rates, visitor surveys, number of visitors to paid attractions and
91 other data supplied by the CNP authority. Data from Flickr were collected through Flickr application
92 programming interface (API <https://www.flickr.com/services/api/>) and R (R Core Team 2015), using
93 the packages RCurl (Lang & the CRAN team 2015; version 1.95.4.7), XML (Lang & the CRAN team
94 2015b; version 3.98.1.3) and httr (Wickham 2016; version 1.1.0). Dates and geographic coordinates
95 associated with the pictures were used to select only those taken in the CNP between 2009 and
96 2014 (R code available at <https://github.com/FrancescaMancini/Flickr-API>). We downloaded
97 different metadata associated with the photos: picture and photographer ID, the date when the
98 photo was taken, the geographic coordinates of where the picture was taken and the tags. In order
99 to avoid duplicates, we used the combination of photographer-ID and date to delete multiple photos
100 from the same user on the same day. Therefore, the number of data points in the dataset represents
101 the number of Flickr visitor days (FVD).

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

116 *2.2. Temporal validation*

117 In order to test whether the temporal patterns shown by the two time series were similar, we used
118 Wavelets Analysis (WA) (Cazelles *et al.* 2008). WA decomposes the variance of the time series in its
119 oscillating components, thus detecting significant periodicities. The advantage of this method
120 compared to other spectral decompositions is that WA does not assume stationarity of the time
121 series, but it allows the main frequency component to change through time by estimating the
122 signal's spectral characteristics as a function of time (the wavelet power spectra). We used the
123 Morlet mother wavelet to perform this decomposition. This continuous and complex function allows
124 the extraction of both time-dependent amplitude and phase of the time series. WA also allows the
125 analysis of patterns of covariation between two time series. We compared the time series of SVD
126 and FVD using the wavelet coherence, which identifies the linear correlation between two signals. In
127 order to assess statistical significance of the association between the two time series we used a
128 random noise resampling scheme, where the null hypothesis tested is that the association between
129 the two signals is not different from that expected by chance alone (Ménard *et al.* 2007). We also
130 computed the phase difference to test whether the two time series were synchronised or out of
131 phase. This analysis was conducted in R using the package WaveletComp version 1.0 (R code
132 available at <https://github.com/FrancescaMancini/Flickr-Statistical-Analysis>) (Roesch & Schmidbauer
133 2014).

134 *2.3. Spatial validation*

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

141 This allowed us to test two hypotheses: 1) the probability of finding at least one picture by one user
142 on Flickr is higher for areas with higher SuVD 2) the number of users posting pictures on Flickr is
143 associated to the number of SuVD in the same area. Since densely populated areas tend to have a
144 higher average number of Flickr users (Fig. A1), we used population abundance for each grid cell as
145 model weights. The residuals of the GLMs were spatially correlated and directional variograms,
146 estimated using R package *gstat* version 1.1.0 (R code available at
147 <https://github.com/FrancescaMancini/Flickr-Statistical-Analysis>) (Pebesma 2004), showed that the
148 spatial autocorrelation was anisotropic (Fig. A2). We therefore used spatial eigenvector mapping
149 (SEVM) to derive explanatory variables for the GLMs (Griffith & Peres-Neto 2006). This method
150 decomposes a matrix of relationships between data points into eigenvectors that capture spatial
151 effects. The eigenvectors can then be included as explanatory variables in the GLM to remove the
152 effect of spatial autocorrelation on the analysis. First we used a Delaunay triangulation of the grid
153 cells centres to define neighbours, to which we assigned row-standardised spatial weights. A set of
154 Moran eigenvectors (ME) were then calculated from these weights and those that best reduced the
155 spatial autocorrelation of residuals were selected and included as spatial covariates in the GLMs. We
156 used AIC to select only the ME that improved the explanatory power of the model. This analysis was
157 conducted in R using the package *spdep* version 0.5.92 (R code available at
158 <https://github.com/FrancescaMancini/Flickr-Statistical-Analysis>) (Bivand *et al.* 2013; Bivand & Piras
159 2015).

160 *2.4. Wildlife watching hotspots*

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

166 **3. Results**

167 In total, we downloaded metadata on 29,336 pictures (4699 unique FVD) taken in the CNP between
168 2009 and 2014 and uploaded on Flickr.

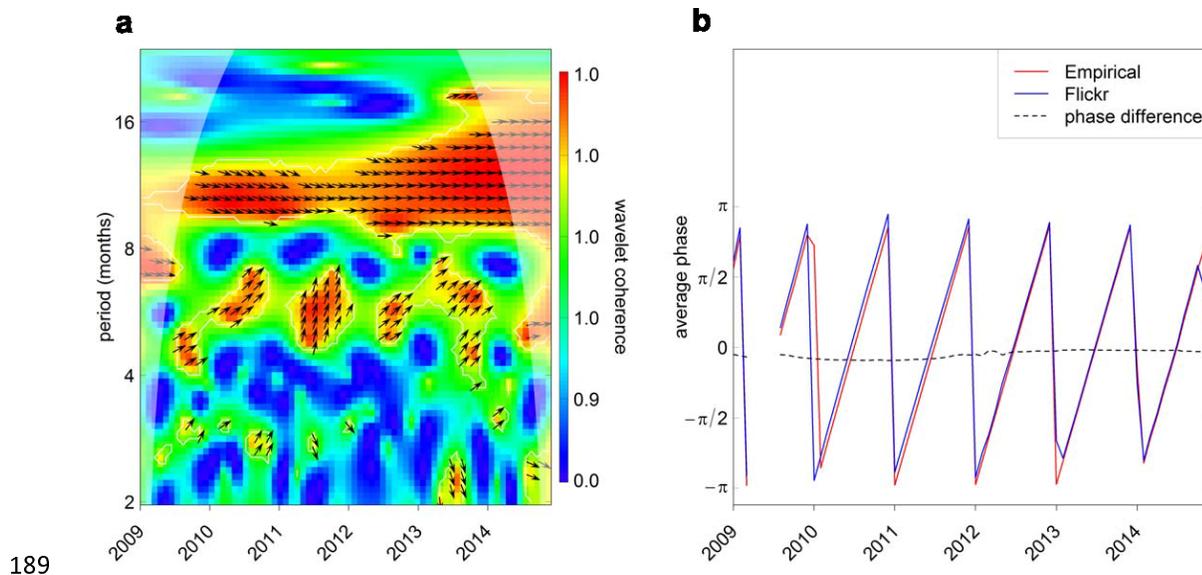
169 The query to Flickr API returned 92,229 results (36,998 FVD) for pictures with the word “bird” in
170 tags, title or description taken in Scotland between 2005 and 2015. From the search with the word
171 “seal” we obtained 2571 FVD and from the search with the word “dolphin” and the word “whale”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

173 *3.1. Temporal validation*

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

181 *3.2. Spatial validation*

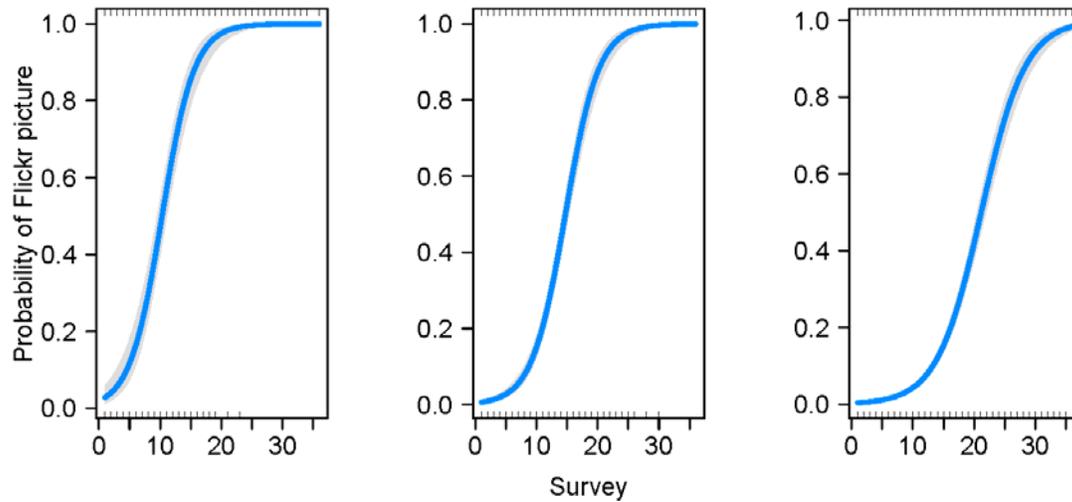
182 The survey dataset indicated that the areas more intensely used for birds and wildlife watching are
183 around the west coast of Scotland, the Moray Firth, the Firth of Forth and the Tay estuary (Fig. A4).
184 Spatial distribution of the pictures from Flickr also identified the last three as areas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



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

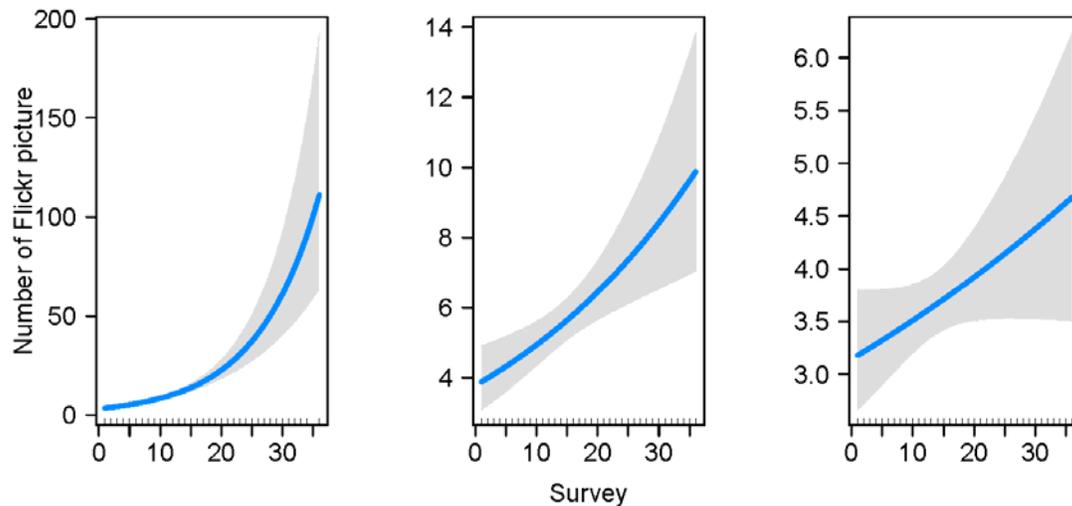
197 The number of SuVD was associated with the presence of pictures posted on Flickr by a user for a
198 certain area (Fig. 2). This was true for each of the spatial scales tested (20km: coefficient = 0.37, $SE =$
199 0.04, $Z = 9.5$, $p\text{-value} < 0.001$; 10km: estimate = 0.36, $SE = 0.02$, $Z = 14.7$, $p\text{-value} < 0.001$; 5km:
200 estimate = 0.28, $SE = 0.001$, $Z = 21.1$, $p\text{-value} < 0.001$). The higher the number of visitors captured by
201 the survey, the higher the probability of finding photos on Flickr taken in that area.

202 The number of SuVD was related to the number of Flickr pictures at 10km and 20km resolution but
203 not 5km (Fig. 3; 20km: coefficient = 0.1, $SE = 0.01$, $Z = 8.1$, $p\text{-value} < 0.001$; 10km: coefficient = 0.02,
204 $SE = 0.001$, $Z = 3.4$, $p\text{-value} < 0.001$; 5km: coefficient = 0.01, $SE = 0.006$, $Z = 1.7$, $p\text{-value} > 0.05$). A
205 higher number of visitors captured by the survey corresponds to a higher number of FVD in the same
206 grid cell.



207

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



211

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

216

217 *3.3. Wildlife watching hotspots*

218 The density maps revealed spatio-temporal patterns of wildlife watching hotspots in Scotland.

219 Birdwatching (Fig. 4) seems to be concentrated around Edinburgh and Glasgow, however when FVD

220 in Edinburgh and Glasgow were excluded from the dataset, hotspots in the Moray Firth, Orkney,

221 Shetland, the Isle of Mull and the North-West coast started to appear (Fig. A6). A seasonal plot of

222 the same pictures shows that this high density around urban areas spreads out towards the West

223 coast and the islands during spring and summer (Fig. A7).

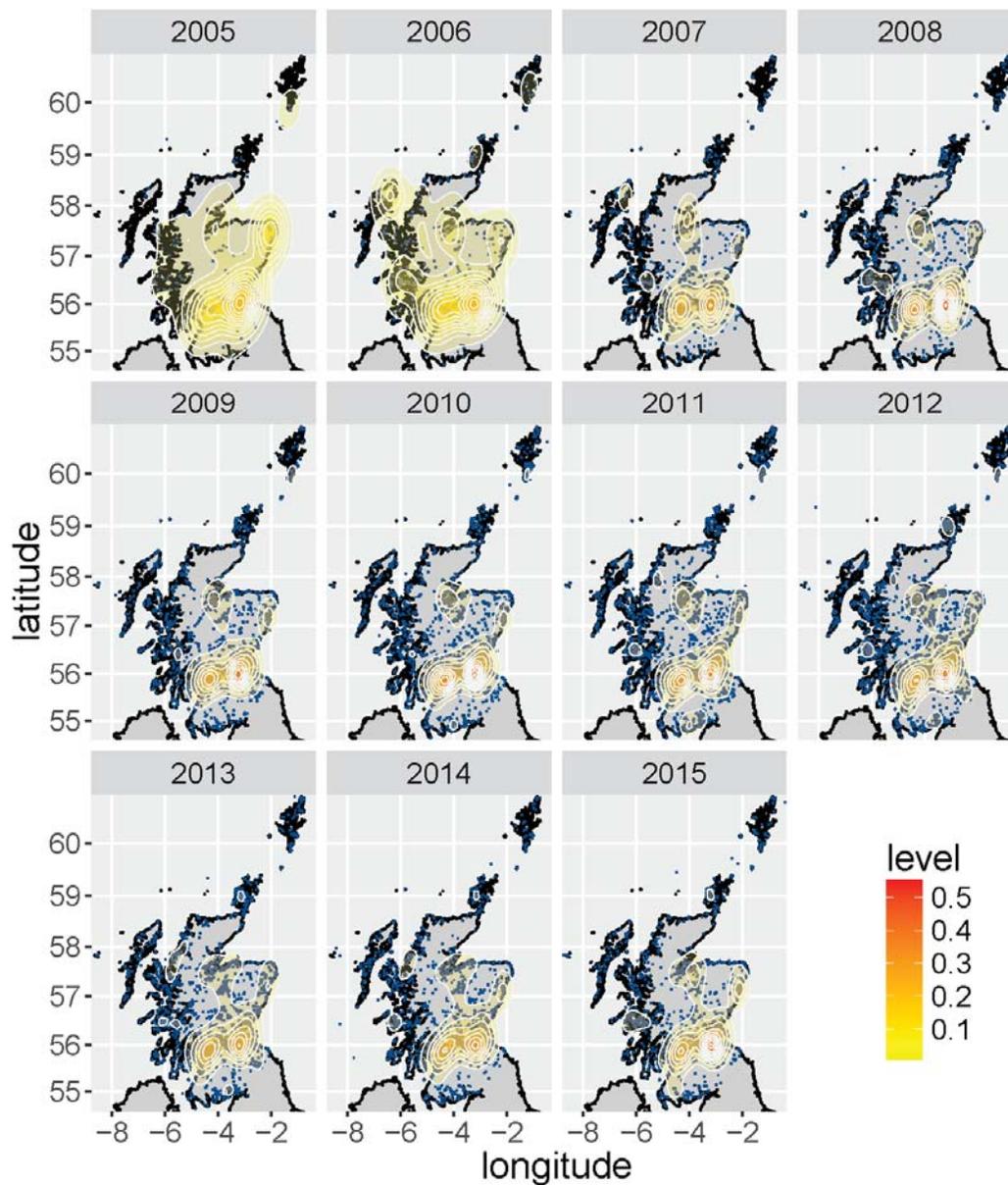
224 Seal watching seems to be concentrated initially around the west coast, the Firth of Forth and

225 Shetland (Fig. 5). It is worth noticing the appearance of another hotspot from 2008 corresponding to

226 Newborough in Aberdeenshire, becoming very important after 2011.

227 Dolphin and whale watching maps showed a consistent hotspot at Chanonry Point in the Moray Firth

228 (Fig. 6), with the appearance of a second hotspot from 2013 in Aberdeen.

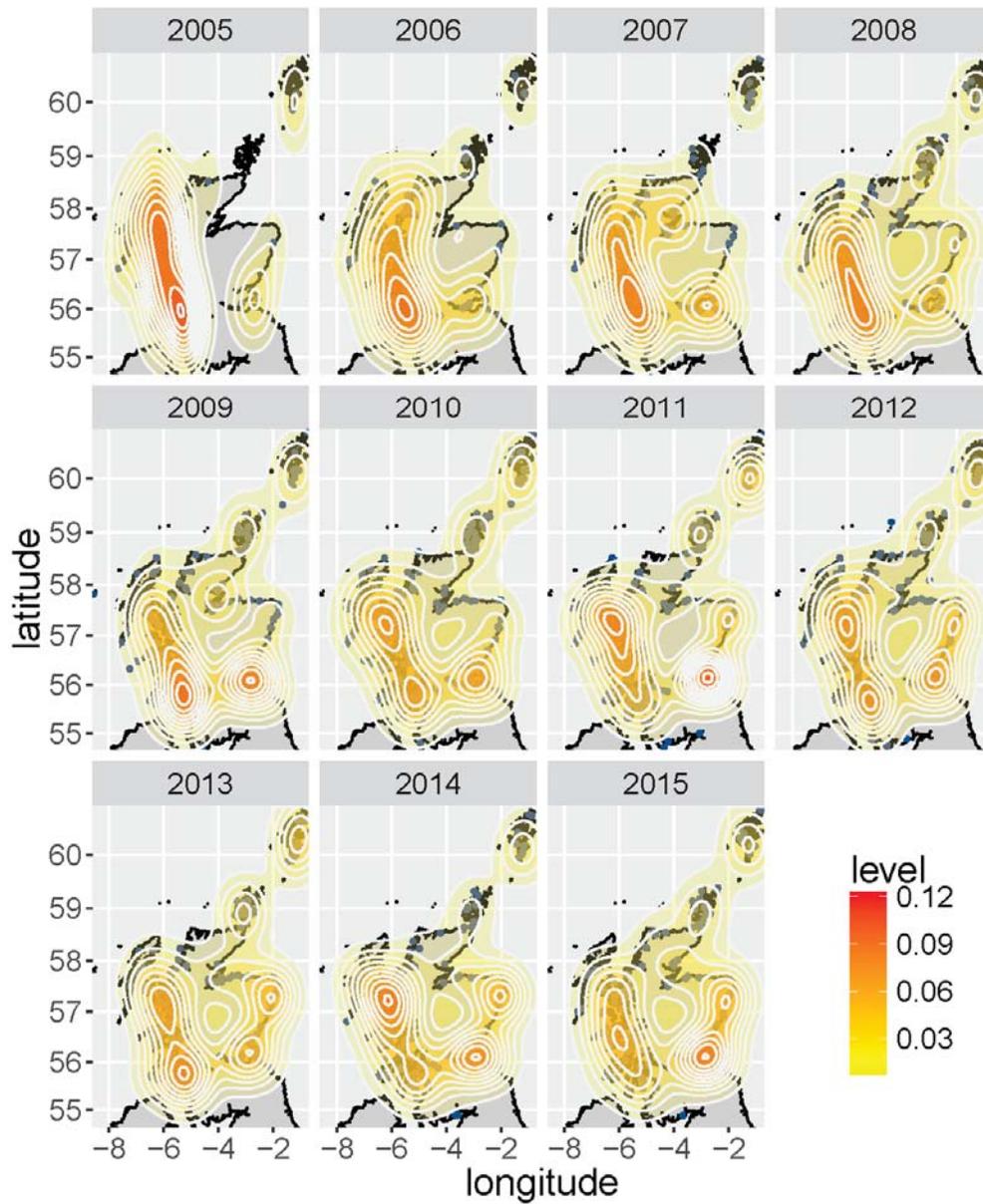


229

230 Fig. 4. Bird watching density maps. Each panel represents the density of FVD in a different year, from 2005 to
231 2015. The blue dots on the maps are the data. Different colours represent different density levels, from low
232 (yellow) to high (red). White lines represent contour lines for different density levels.

233

234



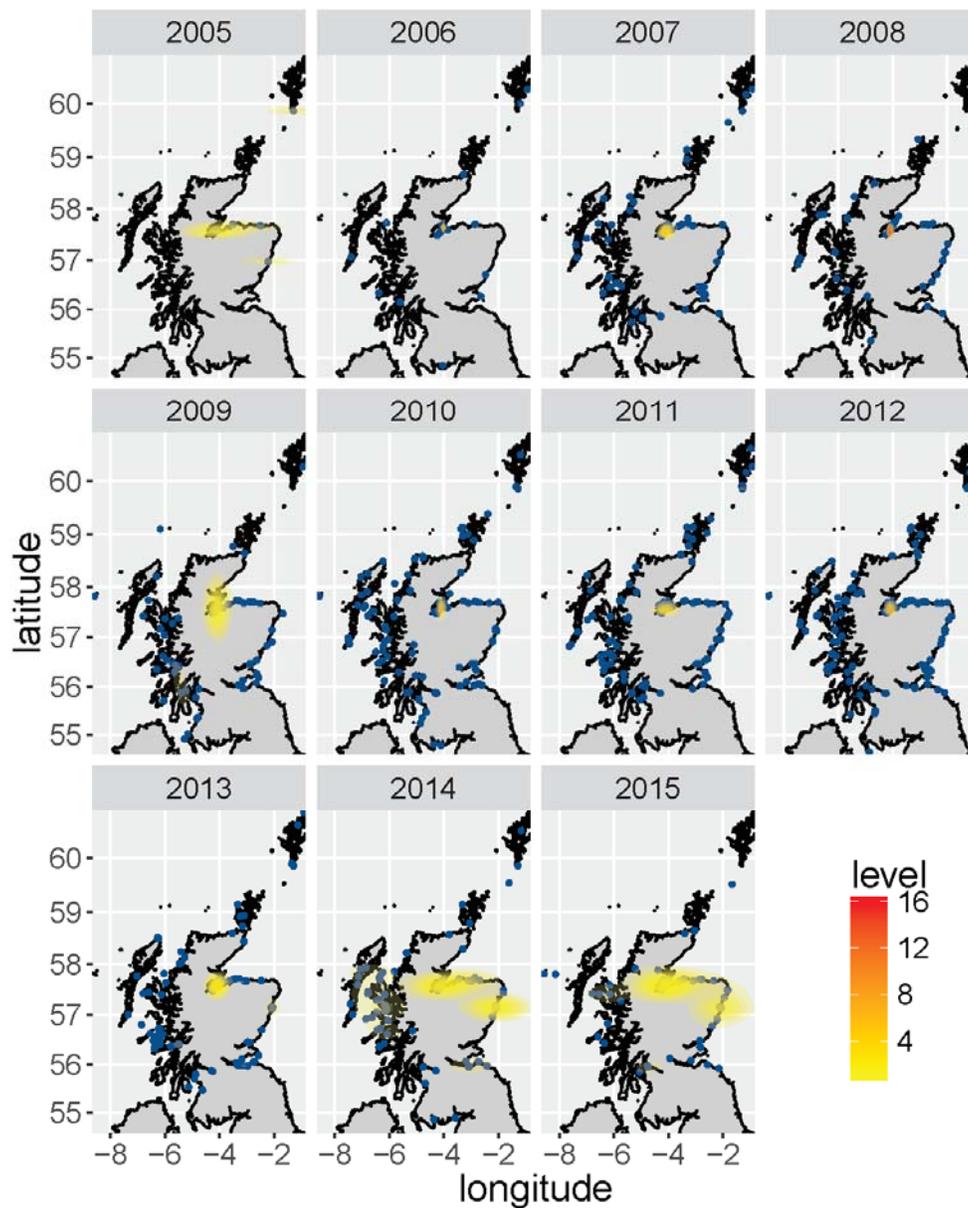
235

236 Fig.5. Seal watching density maps. Each panel represents the density of FVD in a different year, from 2005 to

237 2015. The blue dots on the maps are the data. Different colours represent different density levels, from low

238 (yellow) to high (red). White lines represent contour lines for different density levels.

239



240

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

244 4. Discussion

245 Nature-based recreation is difficult to measure, especially when it does not require the use of
246 infrastructures and it is performed in remote areas. Monitoring tourism in protected areas can also
247 be difficult when access to the park is not subject to an entrance fee. The popularity of smartphones,

248 GPS devices and social media is allowing people to generate large amounts of spatial and temporal
249 data (Worthington *et al.* 2012; Leighton *et al.* 2016), recording where they go and what they see.

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

261 Spatio-temporal trends of wildlife watching hotspots confirm the validity of FVD as a proxy for
262 visitation. The majority of pictures of birds that were taken around Edinburgh and Glasgow, were
263 taken inside greenspaces and urban parks, such as Hogganfield Park in Glasgow and Bawsinch and
264 Duddingston Scottish Wildlife Trust reserve in Edinburgh (Fig. A1). The density maps also detected a
265 change in the bird watching hotspots with seasons (Fig. A6 and A7), consistent with a movement
266 from the area around Edinburgh and Glasgow to more remote areas on the west coast, the Moray
267 Firth and the islands. We were therefore able to capture the movement of people towards tourism
268 destination (Blake *et al.* 2010; Land Use Consultants 2016). The seal hotspot map (Fig. 5) reveals high
269 activity in the Firth of Forth, Tay estuary and the west coast where special areas of conservation
270 (SACs) and haul out sites are present for both grey and harbour seals (Morris *et al.* 2014). This map
271 also shows the appearance of a seal watching point in Newborough after 2008. This site now holds
272 26% of grey seals (*Halichoerus grypus*) and 1% of harbour seals (*Phoca vitulina*) in the East Coast of

273 Scotland Seal Management Area and has recently been proposed as a new designated haul-out site
274 to protect the seals (Marine Scotland 2015). The whale and dolphin watching density map (Fig. 6)
275 also reveals the emergence of a dolphin watching hotspot: Aberdeen harbour. The hotspot only
276 appears in 2013, after the launch by the Royal Society for the Protection of Birds (RSPB) of dolphin
277 watching events from Aberdeen harbour as part of the “Dates with nature” projects
278 (<http://www.rspb.org.uk/discoverandenjoynature/seenature/datemithnature/details.aspx?id=34036>
279 [6](#)). This result indicates that such organised events can attract tourists and create a hotspot, offering
280 an opportunity for managers to shift tourists’ attention from destination that are unsustainably used
281 to unexploited ones.

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

296 In conclusion, despite limitations, the number of geotagged pictures uploaded on Flickr can be used
297 as a proxy for wildlife watching activities at spatial and temporal scales that are relevant for

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

308 **Acknowledgements**

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

315 **Data accessibility**

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

318 **References**

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

- 322 Bivand, R., Hauke, J. & Kossowski, T. (2013). Computing the Jacobian in Gaussian spatial
323 autoregressive models: An illustrated comparison of available methods. *Geographical Analysis*,
324 **45**, 150–179.
- 325 Bivand, R. & Piras, G. (2015). Comparing Implementations of Estimation Methods for Spatial
326 Econometrics. *Journal of Statistical Software*, **63**, 1–36.
- 327 Blake, A., Curtin, S., Richards, S., Vaughan, R., Edwards, J. & Fletcher, J. (2010). *The Economic Impact*
328 *of Wildlife Tourism in Scotland. Report by the International Centre for Tourism and Hospitality*
329 *Research Bournemouth University for Scottish Government Social Research.*
- 330 Bryden, D.M., Westbrook, S.R., Burns, B., Taylor, W.A. & Anderson, S. (2010). *Assessing the economic*
331 *impacts of nature based tourism in Scotland. Scottish Natural Heritage Commissioned Report*
332 *No. 398.*
- 333 Cazelles, B., Chavez, M., Berteaux, D., Ménard, F., Vik, J.O., Jenouvrier, S. & Stenseth, N.C. (2008).
334 Wavelet analysis of ecological time series. *Oecologia*, 287–304.
- 335 Curtin, S. (2013). Lessons from Scotland: British wildlife tourism demand, product development and
336 destination management. *Journal of Destination Marketing and Management*, **2**, 196–211.
- 337 Curtin, S. (2003). Whale-Watching in Kaikoura: Sustainable Destination Development? *Journal of*
338 *Ecotourism*, **2**, 173–195.
- 339 ESRI. (2013). ArcGIS Desktop: Release 10.2. Redlands CA: Environmental Systems Research Institute.
- 340 Ferraro, P.J. & Hanauer, M.M. (2014). Quantifying causal mechanisms to determine how protected
341 areas affect poverty through changes in ecosystem services and infrastructure. *Proceedings of*
342 *the National Academy of Sciences of the United States of America*, **111**, 4332–4337.
- 343 Global Tourism Solutions Ltd. (2006). Overview of STEAM, GTS. URL
344 <http://www.globaltourismsolutions.co.uk/steam-model>

- 345 Gossling, S. (1999). Ecotourism: a means to safeguard biodiversity and ecosystem functions?
346 *Ecological Economics*, **29**, 303–320.
- 347 Griffith, D.A. & Peres-Neto, P.R. (2006). Spatial Modeling in Ecology: the Flexibility of Eigenfunction
348 Spatial Analyses. *Ecology*, **87**, 2603–2613.
- 349 Keeler, B.L., Wood, S.A., Polasky, S., Kling, C., Filstrup, C.T. & Downing, J.A. (2015). Recreational
350 demand for clean water: Evidence from geotagged photographs by visitors to lakes. *Frontiers in*
351 *Ecology and the Environment*, **13**, 76–81.
- 352 Land Use Consultants. (2016). Scottish Marine Recreation and Tourism Survey 2015.
- 353 Lang, D.T. & the CRAN team. (2015a). Rcurl: General Network (HTTP/FTP/...) Client Interface for R. R
354 package version 1.95-4.7. URL <https://cran.r-project.org/package=Rcurl>.
- 355 Lang, D.T. & the CRAN team. (2015b). XML: Tools for Parsing and Generating XML Within R and S-
356 Plus. R package version 3.98-1.3. URL <https://cran.r-project.org/package=XML>.
- 357 Leighton, G.R.M., Hugo, P.S., Roulin, A. & Amar, A. (2016). Just Google it: Assessing the use of Google
358 Images to describe geographical variation in visible traits of organisms. *Methods in Ecology and*
359 *Evolution*.
- 360 Levin, N., Kark, S. & Crandall, D. (2015). Where have all the people gone? Enhancing global
361 conservation using night lights and social media. *Ecological Applications*, **25**, 2153–2167.
- 362 Marine Scotland. (2015). *Consultation on Possible Designation of a Seal Haul-out Site. Proposal in*
363 *respect of the Ythan Estuary*.
- 364 McClung, M. R., Seddon, P. J., Massaro, M. & Setiawan, A. N. (2004). Nature-based tourism impacts
365 on yellow-eyed penguins *Megadyptes antipodes*: does unregulated visitor access affect
366 fledging weight and juvenile survival? *Biological Conservation*, **119**, 279-285. Ménard, F.,
367 Marsac, F., Bellier, E. & Cazelles, B. (2007). Climatic oscillations and tuna catch rates in the

- 368 Indian Ocean: A wavelet approach to time series analysis. *Fisheries Oceanography*, **16**, 95–104.
- 369 Morris, C., Duck, C., Lonergan, M., Baxter, J., Middlemas, S. & Walker, I. (2014). *Method used to*
370 *identify key seal haul-out sites in Scotland for designation under the Marine (Scotland) Act*
371 *Section 117*.
- 372 Pebesma, E.J. (2004). Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*,
373 **30**, 683–691.
- 374 Pirotta, E. & Lusseau, D. (2015). Managing the wildlife tourism commons. *Ecological Applications*, **25**,
375 729–741.
- 376 R Core Team. (2015). R: A Language and Environment for Statistical Computing. R Foundation for
377 Statistical Computing, Vienna, Austria. URL <http://www.r-project.org/>.
- 378 Reed, S.E. & Merenlender, A.M. (2008). Quiet, Nonconsumptive Recreation Reduces Protected Area
379 Effectiveness. *Conservation Letters*, **1**, 146–154.
- 380 Roesch, A. & Schmidbauer, H. (2014). WaveletComp: Computational Wavelet Analysis. R package
381 version 1.0. URL <https://cran.r-project.org/package=WaveletComp>.
- 382 Russell, R., Guerry, A.D., Balvanera, P., Gould, R.K., Basurto, X., Chan, K.M. a., Klain, S., Levine, J. &
383 Tam, J. (2013). *Humans and Nature: How Knowing and Experiencing Nature Affect Well-Being*.
- 384 Scott, D.W. (1992). *Multivariate Density Estimation: Theory, Practice and Visualization* (D.W. Scott,
385 Ed.). John Wiley & Sons, Inc., Hoboken, NJ, USA.
- 386 Silva, L. (2013). How ecotourism works at the community-level: the case of whale-watching in the
387 Azores. *Current Issues in Tourism*, **18**, 196–211.
- 388 Stevens, C.J., Fraser, I., Mitchley, J. & Thomas, M.B. (2007). Making ecological science policy-
389 relevant: Issues of scale and disciplinary integration. *Landscape Ecology*, **22**, 799–809.
- 390 Tisdell, C. & Wilson, C. (2002). Ecotourism for the survival of sea turtles and other wildlife.

- 391 *Biodiversity and Conservation*, **11**, 1521–1538.
- 392 Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer International Publishing.
- 393 Wickham, H. (2016). httr: Tools for Working with URLs and HTTP. R package version 1.1.0. URL
- 394 <https://cran.r-project.org/package=httr>.
- 395 Wood, S. a, Guerry, A.D., Silver, J.M. & Lacayo, M. (2013). Using social media to quantify nature-
- 396 based tourism and recreation. *Scientific reports*, **3**, 2976.
- 397 Worthington, J.P., Silvertown, J., Cook, L., Cameron, R., Dodd, M., Greenwood, R.M., McConway, K.
- 398 & Skelton, P. (2012). Evolution MegaLab: a case study in citizen science methods. *Methods in*
- 399 *Ecology and Evolution*, **3**, 303–309.
- 400