



UNIVERSITY OF LEEDS

This is a repository copy of *Landscape aesthetics: Spatial modelling and mapping using social media images and machine learning*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/162748/>

Version: Accepted Version

Article:

Gosal, AS orcid.org/0000-0001-6782-0706 and Ziv, G orcid.org/0000-0002-6776-0763
(2020) Landscape aesthetics: Spatial modelling and mapping using social media images and machine learning. *Ecological Indicators*, 117. 106638. ISSN 1470-160X

<https://doi.org/10.1016/j.ecolind.2020.106638>

© 2020 Elsevier Ltd. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDeriv (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Landscape aesthetics: spatial modelling and mapping using social media images and machine learning

Gosal, A.S.^{a,*} and Ziv, G.^a

^a School of Geography, Faculty of Environment, University of Leeds, Leeds, LS2 9JT, UK.

*Corresponding author. E-mail address: a.gosal@gmx.co.uk.

Declarations of interest: none.

Research highlights:

- Social media images have more aesthetic value than random ones from the same area.
- Probability models can be used to rank-order images for aesthetic value.
- Incomplete paired comparison datasets can be used to rank order of the images.
- Model results were upscaled for 22,615 images and mapped.

Abstract

Cultural ecosystem services such as aesthetic value are highly context-specific and often present difficulties in their assessment. Here we present a case study in the northern English protected area of the Yorkshire Dales National Park. Utilising publicly available images, paired-comparisons survey, machine-learning based text annotations, natural language processing and regression analysis, we developed a spatial model to predict and map landscape aesthetics across the whole site. The predictive model found eighteen significant variables, including the positive role of rural areas, mountains and vegetation for aesthetic value. Finally, we demonstrate the potential of our approach to varying size datasets and partial paired-comparison matrices, finding a very good agreement with only 20% of paired comparisons was found. This study demonstrates the use of freely available data and mostly open source tools to ascertain landscape aesthetic value in a large protected area.

Keywords: Bradley-Terry model, Flickr photos, LUCAS photos, Google Vision, probability models, image content analysis.

Introduction

Landscapes are integral to human welfare and support many human activities, including scientific, education, heritage-based, aesthetic, symbolic, sacred or for entertainment purposes (Haines-Young and Potschin, 2018). These services are integral to human existence, as highlighted by the European Landscape Convention (2000), which recognised their importance for quality of life. Under the Common International Classification of Ecosystem Services (CICES version 5.1), living systems that enable aesthetic experiences are defined as a separate class in the cultural (biotic) service section of the classification (Haines-Young and Potschin, 2018; Oteros-Rozas et al., 2018), providing an aesthetic value to society. The importance of aesthetic services as a cultural service is well accepted, having been enshrined in the Millennium Ecosystem Assessment in 2005 (de Groot and Ramakrishnan, 2005) and the UK National Ecosystem Services Assessment Follow-on reports (Kenter et al., 2014).

Landscape aesthetics has been studied across different disciplines, including psychology, anthropology, evolutionary biology and landscape planning. Whilst this field has grown in socio-ecological research and public interest, it is still missing well developed quantitative and standardised techniques for assessment (Daniel, 2001; Frank et al., 2013; Tenerelli et al., 2017). Biophysical approaches have included spatial modelling (e.g. Dramstad et al., 2006), using stated preference methods such as participatory GIS (Fagerholm and Käyhkö, 2009; Gosal et al., 2018),

44 interviews and questionnaires (Casado-Arzuaga et al., 2013; Kienast et al., 2015), and photographic
45 surveys (Cheng et al., 2019; Palmer, 2004; Schirpke et al., 2013) with numeric Likert scales to allow
46 participants to score photos. Several studies have shown that people aesthetically prefer natural
47 over urban environments (de Groot et al., 2005; Kaplan and Kaplan, 1989).

48 Gobster et al. (2007) suggests that the history of ancient farming and forestry systems in Europe
49 contribute to the attachment by, and identity of its people, leading to the landscape as being
50 perceived as attractive. Van Zanten et al. (2016) utilised Flickr, Instagram and Panaromio social
51 media photos to quantify aesthetic and recreation values at a continental European scale, finding
52 preference for more mountainous areas. Peña et al. (2015) used aesthetic preference as a proxy for
53 recreation in the Basque Country, Spain, finding mountains and water bodies in landscapes were
54 preferred over homogenous landscapes. Casado-Arzuaga et al. (2014) investigated recreation and
55 aesthetic services in peri-urban environments using a GIS approach finding coastal areas, mountain
56 summits, forests and rural areas corresponded with high aesthetic value. The linkages between
57 ecology and aesthetics have been investigated by Gobster et al. (2007) highlighting that aesthetic
58 experiences are a result of interaction with the 'perceptible realm' (the scale at which humans
59 engage with their landscape surroundings) and are affected by context (landscape or personal-
60 social) . Figueroa-Alfaro and Tang (2017) undertook spatial analysis of geo-tagged photos in
61 Nebraska to identify areas of aesthetic value, and were able to identify clusters of 'new' areas of
62 aesthetic value. Casalegno et al. (2013) quantified online geo-tagged images to evaluate the
63 perceived aesthetic value of an ecosystem finding hotspots of aesthetics value were in coastal areas
64 and a negative correlation with population density. Tenerelli et al. (2017) used social media to
65 investigate the scenic beauty of mountain landscapes in the French Alps, using a technique which
66 combined the images with visual indicators of scenic beauty, with results showing that naturalness,
67 ephemera and visual scale attract foreign visitors, whilst local visits were more attracted by
68 historicity, imageability and complexity. Other studies investigating the features that contribute to
69 aesthetics have included analysis of bundles of landscape features (Oteros-Rozas et al., 2018), within
70 agricultural landscapes (van Zanten et al., 2016b), and Dutch river landscapes (Tieskens et al., 2018).

71 The usual method of eliciting preference in landscape aesthetics research utilise a Likert scale to rate
72 the images (Hägerhäll et al., 2018; Kaplan and Kaplan, 1989). For example, Masuda et al. (2008) used
73 a Likert scale-based survey to investigate aesthetic preference differences between Americans and
74 East Asians of portrait photos in varying contexts. Seresinhe et al. (2015) demonstrated how
75 crowdsourced ratings for the scenic quality of geotagged photos could be coupled with health data,
76 finding that inhabitants of scenic areas report better health. Research has also shown that social
77 media data combined with OpenStreetMap data (to remove photos taken within buildings) increases
78 the accuracy of scenic quality estimates compared to models using census data alone (Seresinhe et
79 al., 2017a), with the application of neural networks and deep learning allowing extraction of scenic
80 features from images (Seresinhe et al., 2017b).

81 While being simple to use, the use of Likert scales has known inherent problems. Heine et al. (2002)
82 suggested that cross-cultural comparisons using subjective Likert scales are compromised as they
83 capture a response related to a shared norm, not to the participants' absolute standing. The
84 'endpoints' of the Likert scale are often set to an individual's expectation of the dimension being
85 measured (Heine et al., 2002; Volkmann, 1951) which varies from one individual to another.
86 Alternatives to the Likert approach are choice experiments and paired comparisons (e.g. Hägerhäll et
87 al. (2018)) to mitigate against the 'endpoints', or internal scale, of the Likert issue.

88 In a paired-comparison survey, interviewees are faced with two options at each step and are asked
89 to pick the one they prefer. Each comparison is typically done in a second or less. Indeed, literature

90 suggests that people make ‘better’ choices if they do not engage in conscious thought, with post-
91 choice satisfaction being reduced with introspection (Dijksterhuis and van Olden, 2006) and the
92 greater the number of choices made available (Iyengar and Lepper, 2000). Dijksterhuis and Nordgren
93 (2008) suggest that whereas conscious thought is better for simple decisions, complex decisions are
94 better suited with unconscious thought. Dijksterhuis (2004) suggests that the unconscious thought
95 allows more polarized, clear and integrated representations in memory. Photo paired-comparisons
96 have been used in several studies to understand aesthetics and scenic beauty (DeLucio and Múgica,
97 1994; Schirpke et al., 2019; Tahvanainen et al., 2001; Tyrväinen et al., 2003). Tyrväinen et al. (2003)
98 used paired-comparisons between images of forests under different management types in.
99 Tahvanainen et al. (2001) examined scenic beauty in Ruissalo Island, Finland, using paired
100 comparison. DeLucio and Múgica (1994) used the paired comparison method to explore landscape
101 preferences from four Spanish national parks. More recently, aesthetic landscape preferences have
102 been investigated by Schirpke et al. (2019) from a paired photo-based survey and other landscape
103 indicators.

104 In contrast to Likert scale, analysis of paired-comparisons data is slightly more complicated and
105 requires a statistical approach. The Bradley Terry (BT) model is a predictive probability logistic model
106 first studied in the 1920s by Zermelo, 1929) and later by Bradley and Terry (1952) after whom the
107 model is named. The BT model is built on the concept of ‘contests’ between two alternatives, be it
108 players, scenarios or images. The model is based on the probability of one item being chosen over
109 another, allowing a full ranking of items to be derived from a sample of paired comparisons (Agresti,
110 2002; Turner and Firth, 2012; Zucco et al., 2019). The potential for using the BT model with natural
111 landscapes has been highlighted by Hägerhäll et al. (2018) in the field of environmental psychology,
112 who used 9 images, and 36 total comparisons, in a study to investigate human preference for natural
113 landscapes. A second challenge with pair-comparisons data is in upscaling – the number of pair
114 combinations increases as N^2 where N is the number of options (e.g. images). However, there are
115 ways to use the BT model with partial paired-comparison data sets as shown in (Zucco et al., 2019).

116 Many of the previously mentioned studies used social media data. In their review, Ghermandi and
117 Sinclair (2019) counted 15 studies that used social media to assess aesthetic value of landscapes.
118 Though the use geotagged images and/or photo-user days alone to assess aesthetic value of an area
119 can often be conflated with recreation, making it difficult to understand how much of each service is
120 really being assessed. Innovative approaches have included Lovato et al. (2013) whom attempted to
121 extract aesthetic preference by analysing aspects of Flickr user ‘favourite’ images, including scenes
122 and colours, which allowed different users to be identified from preference. Yoshimura and Hiura
123 (2017) mapped aesthetic values of landscapes in Hokkaido using social media images and calculation
124 of viewsheds for demand and MaxEnt for supply.

125 Whereas paired-comparisons based probability models have been previously used in a more limited
126 scope (e.g. Hägerhäll et al. (2018)), here we look at how it can be upscaled across an entire
127 protected area landscape using social media images and machine learning. Our study considers
128 several pertinent issues in the field that guide our research questions:

129 (a) As the number of paired comparisons increases by a power of 2, how well can this technique be
130 applied to larger datasets with partial coverage and the BT model?

131 (b) Given that social media photos are likely to be ‘aesthetic’ (at least to the Flickr user who took and
132 uploaded them), would including publicly available and systematically stratified European Land Use
133 and Coverage Area frame Survey (LUCAS) survey photos (Ballin et al., 2018) give a reference set with
134 a wider ‘aesthetic gradient’ on which we can train a predictive model?

135 (c) Can machine learning be used to create a useable set of linguistic predictors for a regression
136 analysis, allowing to extrapolate a BT model based on a subset of images to predict landscape
137 aesthetics at the landscape scale in a given cultural context?

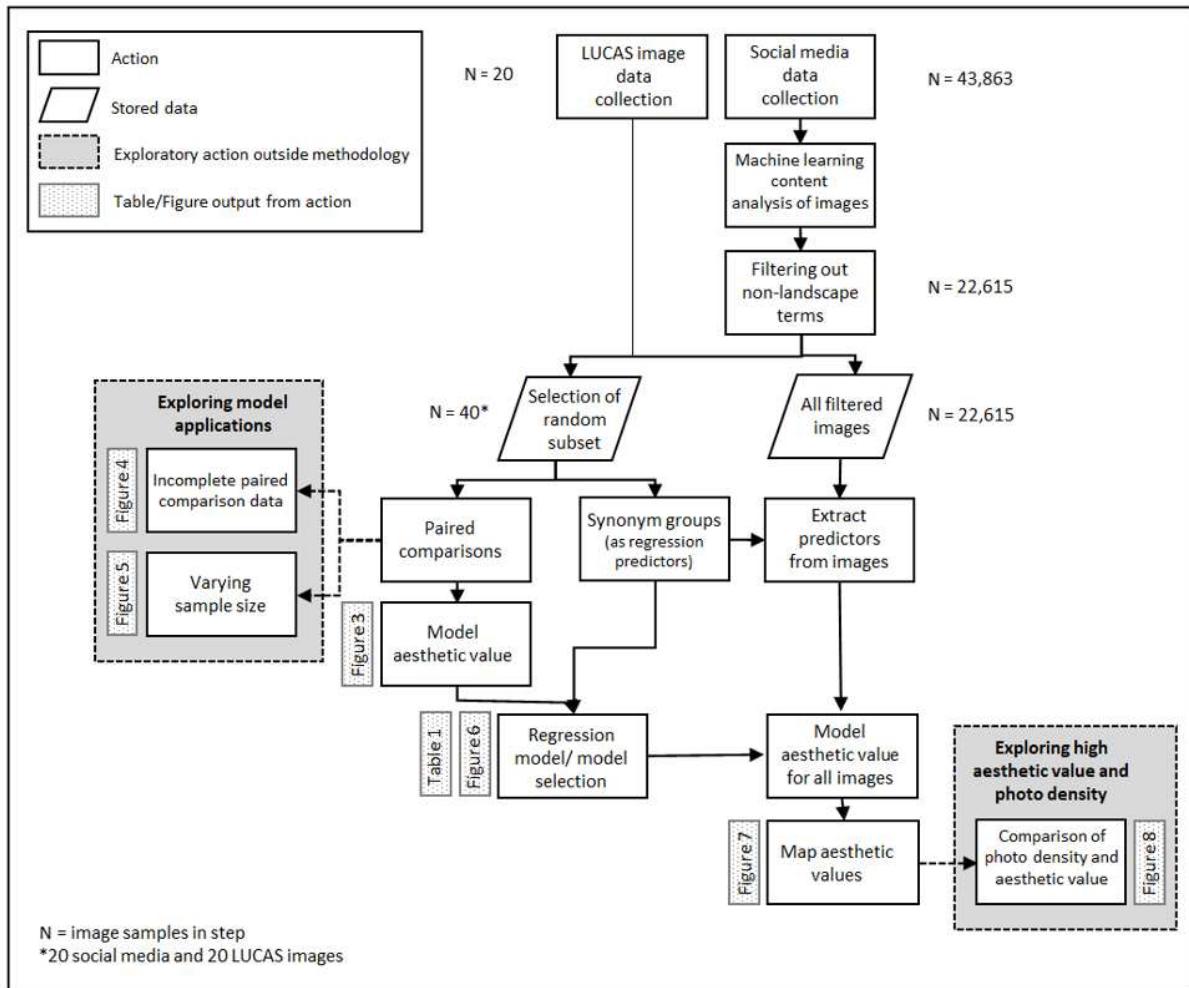
138 **Methods**

139 We adopt multiple techniques, including photo retrieval and content analysis of images using
140 machine-learned APIs and Python scripts, text mining to filter for landscape only images, using
141 natural language processing for synonym term group predictor variables, probability modelling for
142 aesthetic value and finally linear regressions, model selection and mapping. The various steps of the
143 analysis are shown in Figure 1 and explained in detail in the following sections.

144 **Study site**

145 The Yorkshire Dales National Park (YDNP), situated in the north of England (Figure 2) was the site
146 used for this study. The park is known for its natural beauty, with fells over 700m in height, grassy
147 rounded hills with deep ravines, glacial and post-glacial landforms, waterfalls and pastoral landscape
148 (Yorkshire Dales National Park Authority, 2018a). The park is nestled in a highly cultural landscape,
149 with several Areas of Outstanding Natural Beauty (AONB) designated adjacent to the national park
150 including North Pennines AONB, Nidderdale AONB and Forest of Bowland AONB. Designated in
151 1954, and further expanded in 2016, the YDNP habitats are predominantly enclosed pasture/ grass
152 crop farmland and unenclosed uplands used for grazing (Wilson et al., 2018). The YDNP is seen as a
153 prime example of limestone scenery in the UK, with the largest exposure of Carboniferous limestone
154 in England and having over 30% of English limestone pavements, with three quarters classed as
155 being in good condition (Lee, 2015; Yorkshire Dales National Park Authority, 2018a, 2017). Just
156 under half of the UK's upland calcareous grassland is found in the park, with its blanket bog and
157 upland heathland being important at an international level (Yorkshire Dales National Park Authority,
158 2017). The park is also home to many important bird species including Curlew, Lapwing and Black
159 Grouse, and provides high-value areas for breeding waders, with important invertebrate species
160 (e.g. the Northern Brown Argus butterfly), Atlantic white-clawed crayfish and mammals including the
161 red squirrel (Wilson et al., 2018; Yorkshire Dales National Park Authority, 2017).

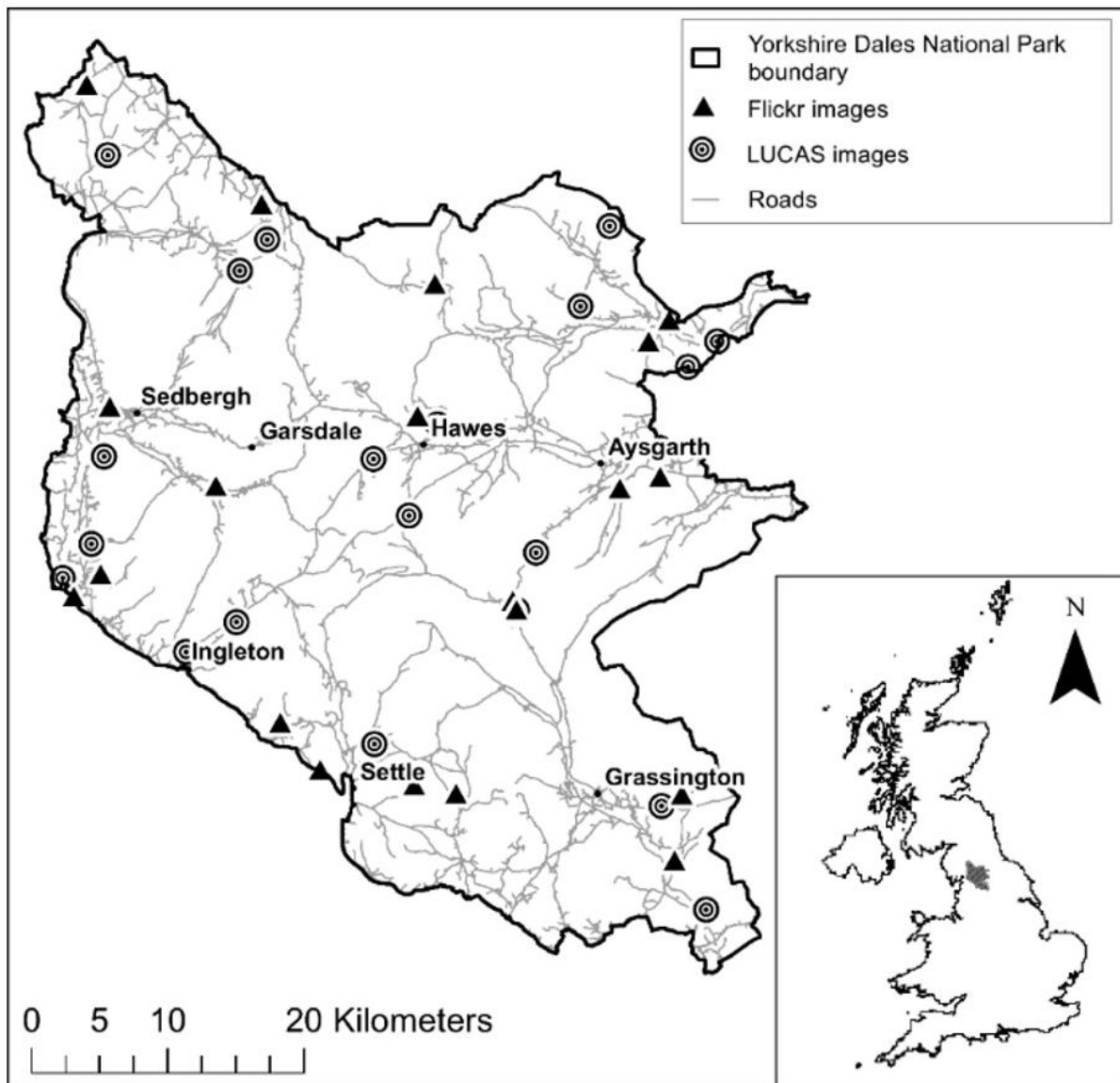
162 Over 24,000 people work and live in the park (Yorkshire Dales National Park Authority, 2018a). It has
163 been predicted that over the next five years, the resident population in the park will decrease, with a
164 skewed age structure due to younger people moving out of the area (Yorkshire Dales National Park
165 Authority, 2018a). Tourism is important for the national park with 3.85 million visitors (including
166 0.52 million overnight visitors) totalling 5.06 million tourist days in 2017 alone, bringing in £263
167 million into the regional economy (Yorkshire Dales National Park Authority, 2018b).



168

169 Figure 1. Flow chart of methodology, including additional exploratory processes and relation to table and other
 170 figures in this study. The numbers of images (N) at various stages have been included for clarity.

171



172

173 Figure 2. Location of the study site, the Yorkshire Dales National Park, situated in the north of England, with
 174 twenty social media (Flickr) and twenty images from the LUCAS database used in this study highlighted. British
 175 A and B roads have been included to show proximity to accessible locations. The inset shows the location of
 176 the national park within the United Kingdom. © Crown Copyright and database right 2019.

177

178 **Photo retrieval**

179 Flickr images with associated metadata (including geographic coordinates, date and unique
 180 photographic ID) were downloaded for the park with bounding box coordinates and Flickr's
 181 Application Programming Interface (API) using the Python 'Scrapy' package. Images were selected
 182 for the years 2009-2018 and geographically cropped to the boundary of the Yorkshire Dales National
 183 Park resulting in 43,863 images. In addition, we used LUCAS data, which is from points that are
 184 sampled from the intersections of a 2 km grid that includes around 1 million points all over the EU
 185 (Ballin et al., 2018). Cardinal photos taken by surveyors can be ordered online (Eurostat, 2020). As
 186 the location of LUCAS points is pre-determined as a desk-based exercise before field visits, there is
 187 no a-priori reason for those photos to be 'beautiful', in contrast to social media photos, which
 188 represent the outcome of cognitive decision to take a photo and share it publicly by users. We will

189 hereafter refer to LUCAS points as being 'random' in space, from a landscape aesthetics perspective.
190 Of the combined LUCAS photos surveyed in 2009, 2012 and 2015 (n=302), we randomly kept twenty
191 LUCAS images through a stratified sample using an overlaid 2 by 5 grid (Supplementary material
192 Figure A1).

193 **Content analysis and predictor variables**

194 Images were analysed for content using Google's Cloud Vision API and Python scripts. Google Vision
195 provides labels, hereafter referred to as terms, and confidence scores (between 0 - no confidence
196 and 1 - high confidence) using their pre-trained machine learning algorithm(Google Cloud Vision,
197 2019), for example, an image may result in being tagged with the term 'highland' with a score of
198 '0.948'. The results were returned as JSON files, which were parsed to extract the terms and scores,
199 above 0.8, in RStudio (RStudio Team, 2015). All terms were manually checked, with 1,393 terms
200 related to human activity or infrastructure, for example, 'terrier', 'garden gnome', or 'coffee house',
201 used to filter the images, resulting in 22,615 images related to landscape with minimal human
202 infrastructure. The term 'farm' was kept, as though it is related to human infrastructure, the UK
203 landscape is heavily influenced with a history of farming which has shaped our landscapes. The final
204 set of photos in predominantly natural landscape type images. Of these, twenty images were
205 selected at random for the paired-comparisons survey (Supplementary material Table A1).

206 Techniques from text mining were used to create a term-document matrix of all terms from all forty
207 images (twenty each from Flickr and LUCAS) using the 'tm' R package (Feinerer et al., 2008; Feinerer
208 and Hornik, 2018), with infrequent partial terms greater than a threshold of 0.95 removed, resulting
209 in 36 terms (such as 'plant', 'highland' and 'stream'). These remaining terms were passed into
210 'Wordnet' lexical English database, where words are grouped by sets of cognitive synonyms, using
211 the 'wordnet' R package for nouns and verbs, and the synonyms for each term extracted (Feinerer
212 and Hornik, 2017; Fellbaum, 1998). The resulting term synonym groups (e.g. highland had 'highland'
213 and 'upland') were used to filter terms for all images (all landscape filtered Flickr images and LUCAS
214 images), and dummy variables created, allocating '0' for non-presence and '1' for presence of the
215 term synonym group. The term synonym groups 'grass' and 'pasture' were removed, as all images
216 had presence of these groups, resulting in 34 dummy variables which were later used as predictors.

217 **Paired-comparison survey**

218 An online experiment interface was built using Construct 3 (Scirra Ltd, 2019), a system designed
219 predominantly as a game editor. The experiment was designed to present in a random order all
220 paired comparisons (780 in total for 40 images), with the user asked to choose the image they found
221 most aesthetically pleasing. Physical participant information sheets were read, and consent forms
222 were signed before participants began. The user clicked the image they chose, and the system
223 presented the next pair until all comparisons were made. Each user was allocated a unique ID and
224 image choice and image pair order presented to the user stored for analysis. Ten participants were
225 recruited from University of Leeds PhD students, with an equal number of males and females, 80%
226 were aged between 26-35 with the remaining 20% between 36-45, and 70% had master's degrees or
227 equivalent (with the remainder having PhDs).

228 **Modelling aesthetic value**

229 The data from the paired-comparisons survey was modelled using the BT model using the
230 'BradleyTerry2' R package (Turner and Firth, 2012) with the value set relative to the baseline
231 category (set a posteriori to be the highest-ranked image) (Zucco et al., 2019) and normalised
232 between 0 and 1. The model predicts 'ability (α)' (also called the 'estimate' or 'worth' value), for

233 clarity and ease of understanding in our study, the modelled value will henceforth be referred to as
234 the 'aesthetic value'. The statistical significance of differences in value between each image and the
235 baseline image (highest ranked image) is also included. Following Zucco et al. (2019), as it is not
236 feasible for all p -values to be reported between all paired comparisons, bootstrapped confidence
237 intervals for the aesthetic value and rank order of each image are reported. This involved drawing
238 1000 simulations for each image's position from a multivariate normal distribution with mean and
239 variance-covariance matrix set to their empirical values, with each image's 5th and 95th percentile
240 rank across all simulations reported.

241 **Partial BT modelling**

242 To investigate the potential of partial or incomplete paired comparisons (where participants only
243 rate a subset of paired image comparisons) to model aesthetic value, the paired-comparisons data
244 was modelled in parallel with differing numbers of responses (truncated at 104, 200, 300, 400, 500,
245 600 and 700 responses, in order of paired comparison, out of the total 780 comparisons) using the
246 BT model. In this study we used a fully connected undirected 'graph' (the set of vertices (images)
247 connected by edges (paired comparisons)). This meant the lowest truncation was 104, as lower than
248 this would have left the graph unconnected. This value is not theoretical and specific to our
249 experiment and will vary with the same value for N due to randomization of the paired comparisons
250 used in a partial dataset. Although methods exist to analyse partial matrices with unconnected
251 graphs with variants of the BT model, we use the standard model which required a fully connected
252 graph to allow the model to rank the images based on maximum likelihood estimates. To investigate
253 the applicability to varying image datasets, three random subsets of 10 (25%), 20 (75%) and 30 (75%)
254 images each were taken, and BT model analysis repeated, once again using truncated paired-
255 comparison matrices for each. The total number of comparisons for each subset was 45 comparisons
256 for 10 images, 190 comparisons for 20 images, 435 comparisons for the 30 images and 780
257 comparisons for the full set of 40 images. Kendall's tau-b was calculated to investigate correlation,
258 due to its ability to handle tied ranks.

259 **Regression analysis and mapping aesthetic values**

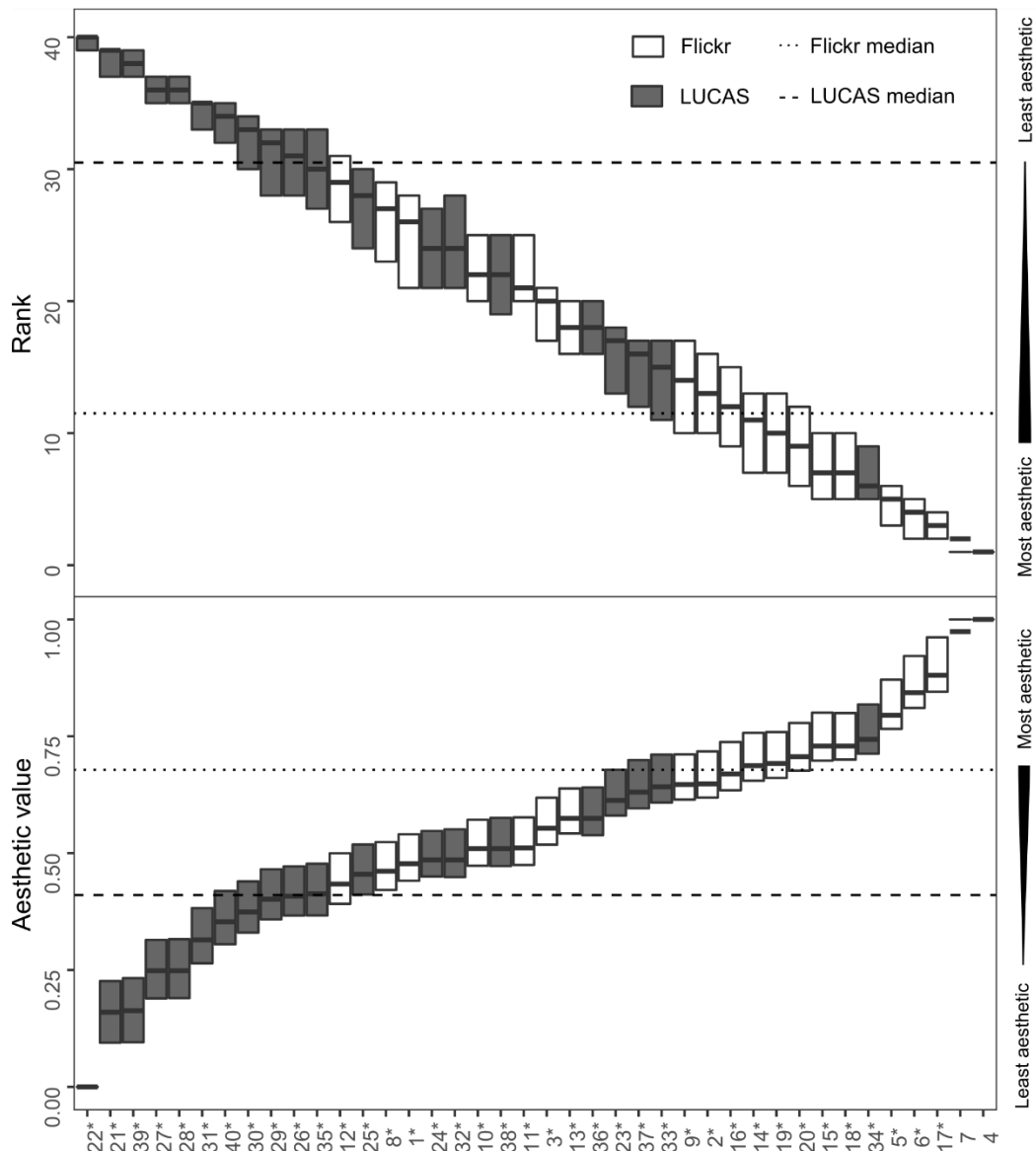
260 Linear regression modelling was used, with the aesthetic value (from the full paired-comparison BT
261 model) as the outcome variable. The 'waterway', 'watercourse', 'riparian zone' and 'body of water'
262 predictors had the same presence distribution across 40 images, hence were grouped into a single
263 collective variable. This resulted in 31 term synonym groups as predictor variables. Variance Inflation
264 Factors (VIFs) were calculated to investigate multicollinearity, with the highest VIFs removed until all
265 VIFs were under 10. Some collinearity is to be expected, due to the nature of the landscape type
266 predictors and the use of synonym groups. Stepwise selection (both directions) was used, and non-
267 significant predictors were successively removed until the model consisted of only significant
268 predictors. Cross-validation was used to quantify uncertainty in the model. The final model was used
269 to predict the aesthetic score for accessible areas within Yorkshire Dales, using for all 22,615
270 landscape-only Flickr images in the park and mapped at a 1 km² resolution, using the mean aesthetic
271 estimate of all images within a cell. Mapping was undertaken in ESRI ArcMap (v.10.6), with a
272 standard deviation ($n=2.5$) stretch applied to all visualised maps.

273

274 **Results**

275 **Paired comparison survey**

276 Our analysis revealed that perceived aesthetic value can be discerned through paired-comparisons
277 and Bradley-Terry models, with LUCAS images having both a lower mean and rank for aesthetic value
278 compared to Flickr images (Figure 3). Images were ranked by perceived aesthetic value, see
279 Supplementary Information Appendix C. The estimates are provided relative to the baseline category
280 (the highest-ranked image) set arbitrarily at zero, and then normalised between 0-1 for ease of
281 interpretation. The statistical significances reported refer to the difference to each image and the
282 baseline. The top two images are not significantly different from each other; hence there is some noise
283 in the estimated values, though both are Flickr images. Remaining images are significantly
284 distinguishable from each other ($P < 0.01$), compared to the baseline (highest scoring) image.
285 Following Zucco et al. (2019), we computed bootstrapped confidence intervals for the rank order for
286 each image in lieu of reporting p-values for all paired comparisons. The average standard error of the
287 aesthetic value was 0.23, with the Flickr images having slightly less at 0.22 and LUCAS images having
288 0.23. The median estimate and rank were higher for Flickr images with 0.68 and 11.5 respectively,
289 with LUCAS images having lower value of 0.41 and 30.5. Flickr images also dominated the higher the
290 ranks, with 75% of the most aesthetic images being of Flickr origin, in agreement with our postulate
291 that aesthetic preference is higher in non-random landscape images from social media.

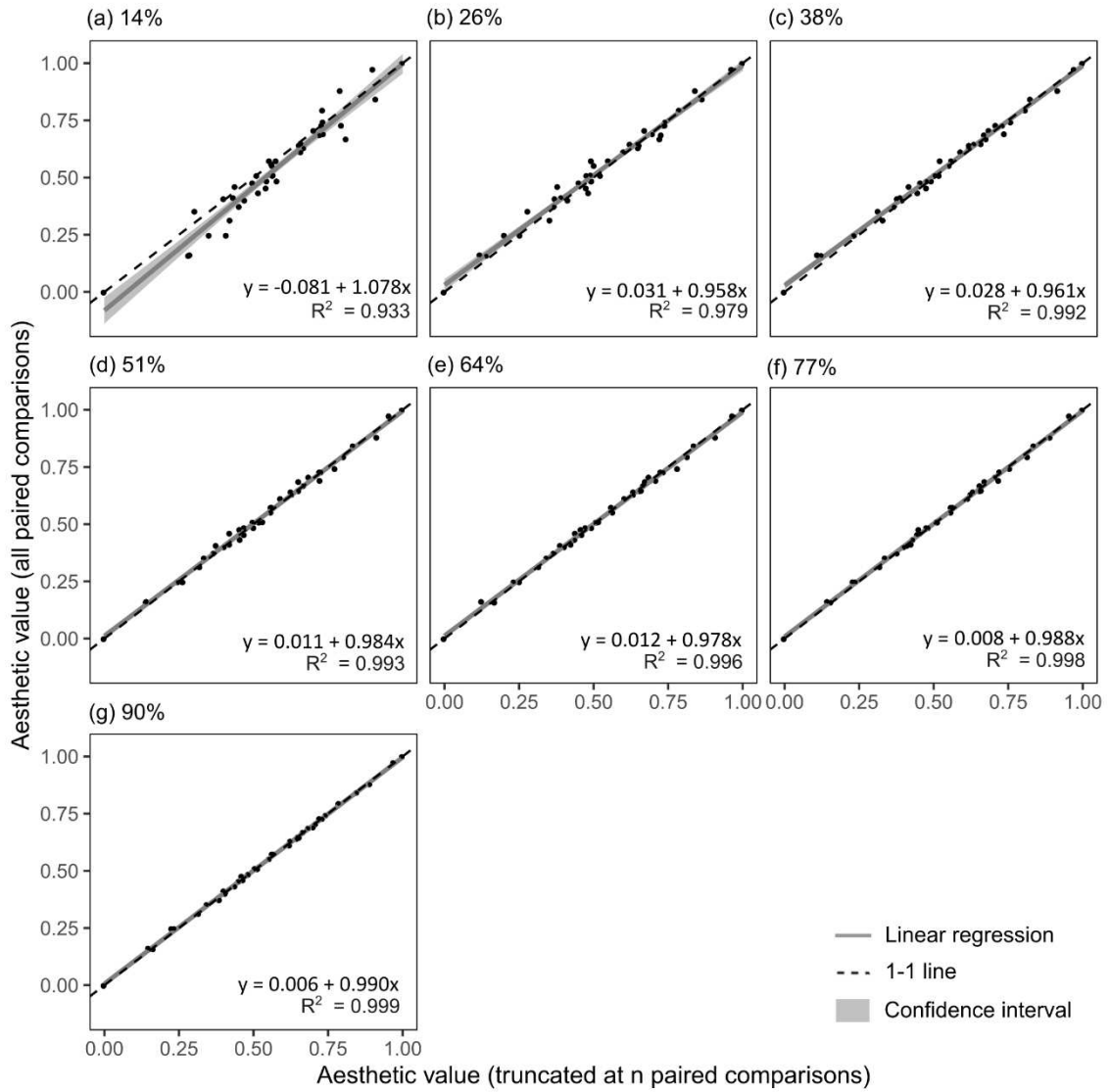


292

293 Figure 3. Estimated aesthetic value (Bradley-Terry model estimates) for forty samples images using complete
 294 paired comparisons from the Yorkshire Dales National Park. P-values compared to the baseline (highest value)
 295 image (04) (*P < 0.01). The CI's reports the 90% confidence intervals of aesthetic value and rank order of each
 296 image, computed using a 1000 simulation parametric bootstrap.

297 **Partial paired comparison data**

298 A set of seven partial (or incomplete) paired comparison subsets of the data at 16-90% of the total
 299 780 comparisons, were modelled and compared to the full 100% model shown in Figure 4. Results
 300 show that even at 14% of the full set of pair-comparisons a strong correlation (R^2 of 0.93) between BT
 301 model estimates for full and partial set exists, increasing to 0.99 from 38%. The standard error
 302 increases with less responses, between 0.225 to 0.616 with the maximum width of the rank-order
 303 confidence interval increasing from 7 to 14 with the least responses BT-model (Supplementary
 304 Information Appendix C, Table C1). Differences are to be expected between varying response models,
 305 though correlation between all paired comparison models to the truncated paired comparisons
 306 models show a high level of similarities.



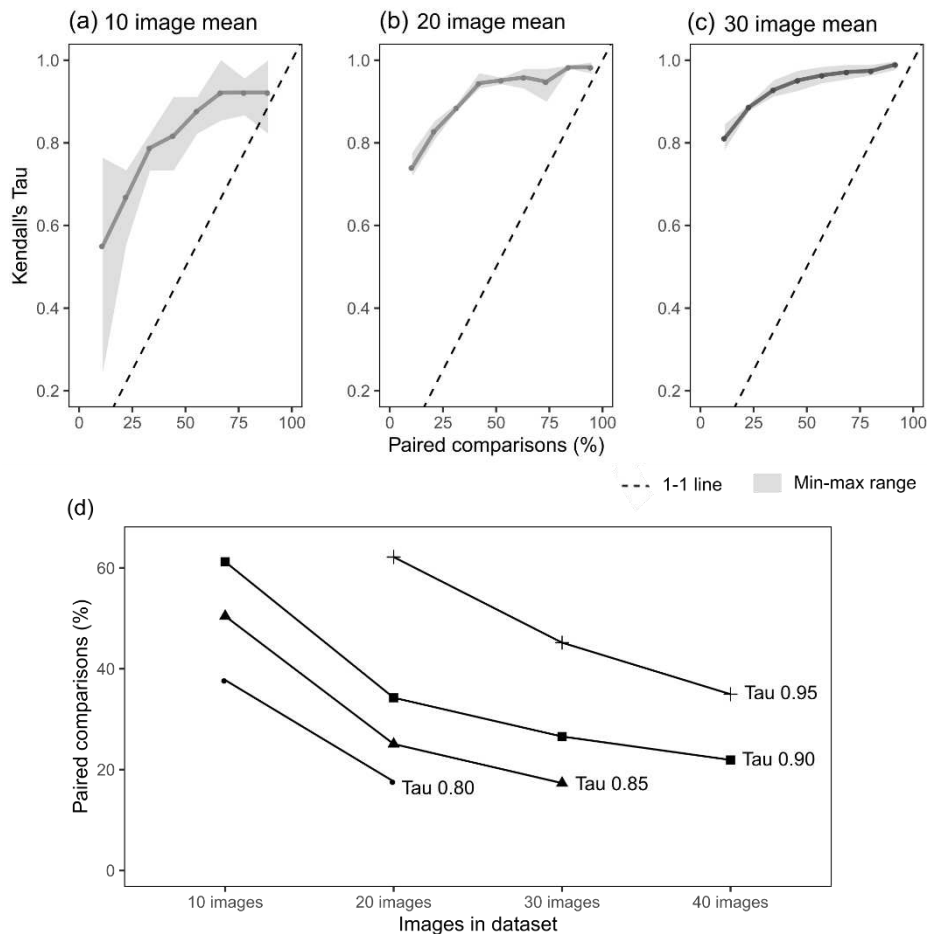
307

308 Figure 4. Complete paired comparison Bradley-Terry model estimates vs varying BT models using different
 309 number of paired comparisons (partial matrices) for forty sample images. Adjusted R² shown.

310

311 To demonstrate the transferability of the method, random subsets of 10 (25%), 20 (75%) and 30
 312 (75%) images were modelled using the Bradley-Terry model (each subset modelled in parallel with
 313 differing levels of partial paired-comparison data). Kendall's Tau-b was used to indicate the
 314 correlation between the partial and full paired comparison data within each subset independently)
 315 (Figure 5). It can be seen that Kendall's tau-b reaches higher values at lower paired comparison
 316 truncations the higher the sample size, though a tau of over 0.80 can be achieved having at least
 317 20% paired comparisons in any of the subsets.

318



319

320 Figure 5. Plot showing the correlation between complete paired comparison BT model estimates and varying
 321 BT models using different number of paired comparisons (partial matrices) for three subsets of the dataset: 10,
 322 20 and 30 images, with three replicates (n=3). Panel (d) shows multiple Kendall's Tau-b values for partial
 323 paired-comparisons (in % of total number of possible pairs) by images used. Tau-b values used in (d) are
 324 intercept values from a linear model for respective (a), (b) and (c) BT models.

325

326 **Predictive modelling**

327 The BT estimates for the original complete BT model (Figure 3) were regressed onto synonym groups
 328 derived from content in landscape images from the Yorkshire Dales National Park. The final model
 329 has an adjusted R² of 0.62 and an RMSE of 0.17 (based on 8-fold cross-validation) with 8 negative
 330 predictors and 11 positive predictors (Table 1 and Figure 6), with the latter having an overall higher
 331 level of significance. The most important of the positive predictors were ‘rural areas’ and ‘mountain’
 332 with estimates of 0.671 and 0.60, with ‘vegetation’, ‘ecoregion’, ‘plain’, ‘stream’, ‘prairie’ and ‘fell’ all
 333 having estimates between 0.29 and 0.44. The strongest negative predictor was ‘field’ with a value of
 334 -0.75, followed by ‘land lot’, ‘steppe’, and ‘highland’, with estimates ranging from -0.33 to -0.46.
 335 Whereas the model explains a significant percentage of the variance using regression based on
 336 machine learning (using reduced term synonym groups), the results suggest that there are factors
 337 outside of the synonym groups that impact aesthetic preference, as would be expected.

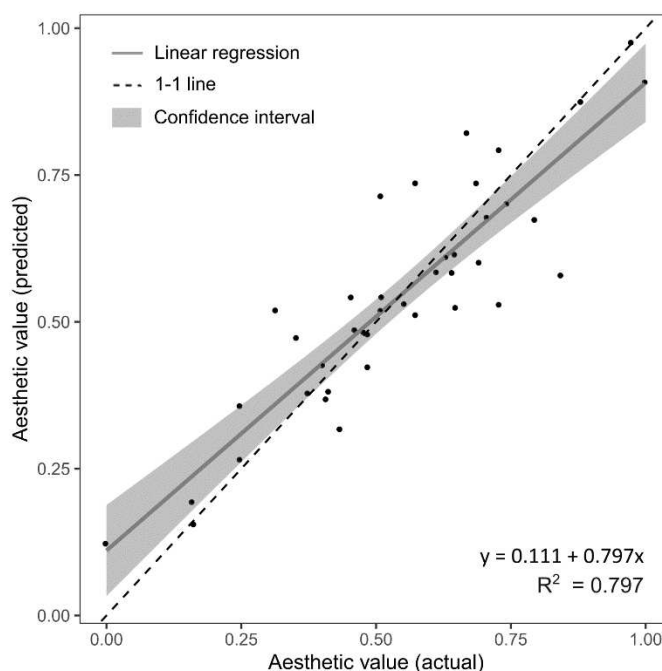
338

339 Table 1. Linear regression model (using forty sample images) after stepwise and Variance Inflation Factor
 340 selection (VIF<10) and removal of non-significant variables. This resulted in R² of 0.797 (adjusted R² = 0.622)
 341 with model p-value <0.001. Variables p-values are indicated as *** <0.001, ** <0.01, * <0.05.

Aesthetic value ability		
<i>Predictors</i>	<i>Estimates</i>	<i>CI</i>
(Intercept)	0.05	-0.17 – 0.27
Rural area	0.71 ***	0.43 – 0.99
Mountain	0.60 ***	0.35 – 0.84
Vegetation	0.44 ***	0.29 – 0.60
Ecoregion	0.35 **	0.17 – 0.52
Plain	0.33 **	0.11 – 0.55
Stream	0.32 ***	0.17 – 0.46
Prairie	0.30 **	0.10 – 0.50
Fell	0.29 **	0.10 – 0.48
Mountainous landforms	0.20 *	0.05 – 0.35
Green	0.16 **	0.05 – 0.27
Farm	-0.12 *	-0.24 – -0.01
Tundra	-0.19 *	-0.36 – -0.01
Hill	-0.24 *	-0.41 – -0.07
Water	-0.24 *	-0.42 – -0.06
Highland	-0.33 **	-0.53 – -0.14
Steppe	-0.34 **	-0.54 – -0.14
Land lot	-0.46 **	-0.70 – -0.23
Field	-0.75 ***	-1.04 – -0.45

342

343



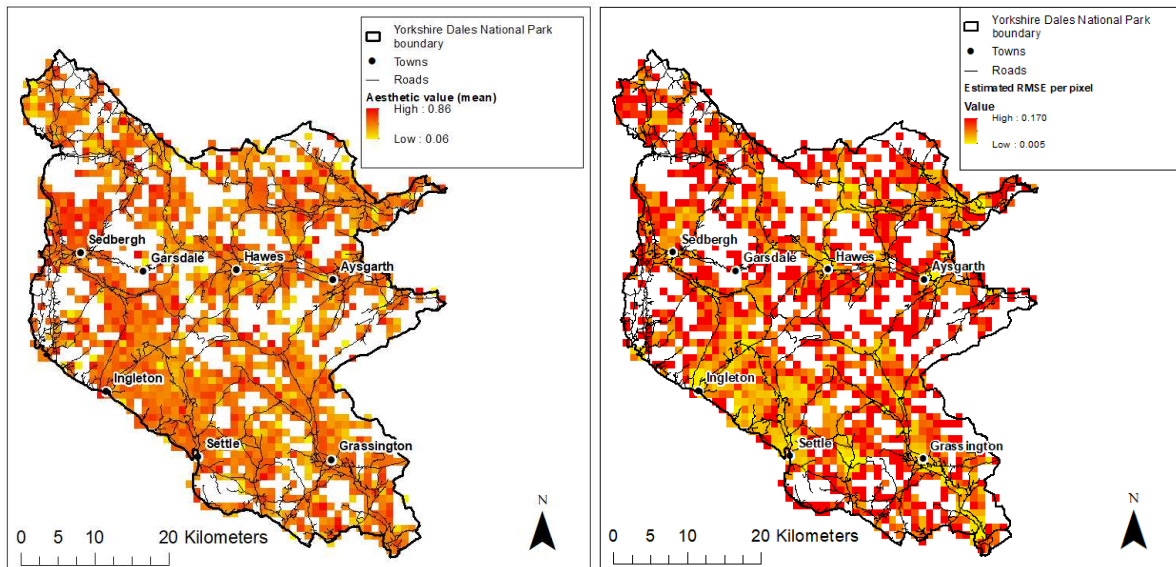
345

346 Figure 6. Predicted aesthetic estimates vs actual BT model estimates using the original full model with all
 347 images and paired comparisons. Note: all values normalised between 0 and 1.

348 **Mapping aesthetic value**

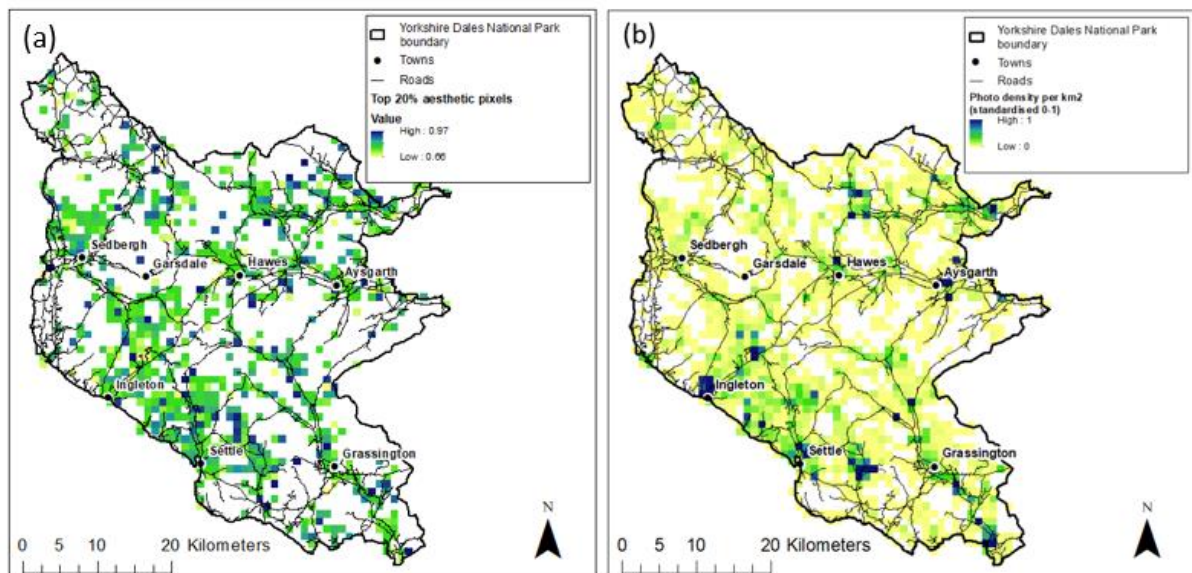
349 The regression model was used to predict aesthetic value for all landscape for the entirety of the
 350 Yorkshire Dales National Park (Figure 7). High values can be seen across the National Park. Highlights
 351 include the region north of Sedbergh, a hilly area with some including the Howgill Fells which is used
 352 by climbers and known locally for its picturesque views. White spaces in the map are 34% (752 km²)
 353 of all pixels, where no Flickr landscape photos were found and are noticeably away from easy access
 354 (roads). An area with many roads and the River Swale, to the north of Hawes and Aysgarth, also
 355 exhibits high aesthetic value. Malham Moor, north-east of Settle, has high valued pixels, as does the
 356 village Ingleton, and Bolton Priory (Augustinian priory ruins set in riparian rural landscape). To
 357 explore the visual correlation between the areas with high aesthetic value and number of photos
 358 taken, the top 20% and photo density were mapped (Figure 8) and showed visual similarity in spatial
 359 pattern.

360



361

362 Figure 7. (a) Predicted aesthetic values for the Yorkshire Dales National Park using 22,615 images (prefiltered
 363 to only landscape images), (b) estimated RMSE per pixel. White space between pixels indicate areas where no
 364 landscape images were present. © Crown Copyright and database right 2019. [COLOUR PRINT]



365

366 Figure 8. (a) Top 20% of predicted aesthetic values, (b) density of photos taken, per km² in the Yorkshire Dales
 367 National Park. White space between pixels indicate areas where no landscape images were present. © Crown
 368 Copyright and database right 2019.

369 Lastly, images with high and low aesthetic value from the National Park were extracted and visually
 370 inspected (Figure 9). High scoring images can be seen to exhibit more ‘complexity’, having a ‘hilly’
 371 character, with the horizon not being complexly ‘horizontal’. There also appears to be a variety of
 372 components in each image, including vegetation (both low, i.e. grass, and high such as trees) and
 373 water. There are also images of snow, though the underlying landscape can be seen at least in part.
 374 The low scoring images often show a flatter horizon, with large swathes of low green vegetation, i.e.
 375 fields and grassland. Visual analysis shows that the techniques culminating in the final mapping of
 376 aesthetic value suggest the method was robust.



377

378 Figure 9. The top five (a1 – a5) and bottom five (b1 – b5) images with predicted standardised (0-1) aesthetic
 379 value estimates from 2,428 available Creative Commons images within the dataset (from a total of 22,615
 380 images that were identified as landscape images). Note, estimates are rounded to two decimal places. Where
 381 multiple images had the same aesthetic value estimates, the images were selected based on numerical
 382 descending ID number of the image. Grey upper and lower bounding bars and black outlines were added to
 383 images for this figure for clarity and are not present in the originals. Full Creative Commons attributions are
 384 given in Supplementary Information Appendix E. **[COLOUR PRINT]**

385 **Discussion**

386 The use of paired-comparisons, while remediating the ‘endpoints’ of inter-personal variability of
 387 Likert scale surveys, is difficult to upscale to larger number of photos. This study showed that a
 388 partial matrix of paired-comparison data can be used to provide results that are not dissimilar from
 389 using complete paired-comparison data to extract aesthetic preference from social media images.
 390 This highlights the potential of the techniques demonstrated to combat the N^2 challenge from using
 391 paired comparison methods whilst still maintaining the ability to get reliable R^2 values. We found
 392 very good agreement with only 20% of the paired comparisons. Additionally, similar patterns in the
 393 data can also be seen in different values of N. Results were expected, as other studies such as
 394 Zucco et al. (2019) have successfully used partial paired comparison datasets, having used 275 out of
 395 666 potential comparison sets (or 41%).

396 Our final model using regression based on machine learning and content tags reduced by synonym
 397 grouping explained a large amount of the variance in the BT model output. When used at the
 398 landscape scale, the final maps showed areas with high aesthetic value were often those where
 399 many photos were taken (Figure 8). This affirms the expectation that landscapes with high aesthetic
 400 value attract more recreational users, and thus have more photos taken (and uploaded to Flickr) (as
 401 this would translate into a high photograph user-days (PUD) value (e.g. Gosal et al., (2019)). The
 402 resulting high-value aesthetic images from the upscaled data showed a preference for landscape
 403 complexity, which has been shown to be a driver for preference (Ode et al., 2010). The models also
 404 showed a positive impact of mountains, mountainous landforms and streams. This partially follows
 405 van Zanten et al. (2016) who found that the strongest predictors for high aesthetic value were hills
 406 and mountains, while ‘hills’ in our study show a moderate negative effect, perhaps a result of the
 407 differentiation in the Google Vision tags for these terms. Though our results are confirmed by Peña
 408 et al. (2015) whom found that personal aesthetic assessments were predominantly based on
 409 naturalness of an area and how land use management impacted it. They found that diverse,
 410 mountainous and those landscapes with water bodies were preferred over flat, homogenous, water-

411 less landscapes. Casado-Arzuaga et al. (2014) found that high aesthetic value areas were
412 corresponded with more natural landscapes, especially coastal ecosystems, rough summits,
413 forest/pasture mosaic areas and rural settlements. They found that the human influence of farming,
414 in terms of farmhouse, livestock or crops) added aesthetic value in landscape compared to those
415 landscape that no longer had farming associations. This helps to explain our own result of rural area
416 being the strongest positive predictor in our model.

417 Areas of high landscape coherence are often considered of higher aesthetic value, including those
418 physiognomic classes with agricultural areas, forests, water bodies and areas of transitional
419 vegetation, with low coherence seen in areas with urban fabric and bare rocks (Karasov et al., 2020).
420 This high coherence can be seen in our results (i.e. positive predictors including rural area, streams,
421 ecoregion, prairie and 'green'. Field, land lot, steppe and tundra, all negative predictors in our study
422 are indicative of low coherence areas.

423 It has also been found that high landscape aesthetic quality in mountain ranges in Germany, riverine
424 areas, coastal areas and lower in urban areas (Hermes et al., 2018). Our results showing
425 mountainous landscape, and mountains having positive effects reinforce Hermes' et al. (2018) study.
426 It has been found that viewpoints, agricultural land and cultural/historical features contributed
427 highly landscape aesthetic capacity (Langemeyer et al., 2018). Interestingly, Langemeyer et al. (2018)
428 found that context impacted landscape aesthetic capacity. For example, in littoral-mountainous
429 landscapes, it was the presence of sea and forest elements that increased value, in mountainous
430 landscapes the crest line (scenic-ness in the background) and in urban landscapes, it was rivers and
431 hills.

432 It has been suggested that the attractiveness of landscapes are linked to its physical attributes
433 (Casado-Arzuaga et al., 2014; De Vries et al., 2007), and our results would confirm that features of
434 landscapes identified from photographs are indeed strong predictors of aesthetic value, and
435 generally is consistent with other studies that have investigated the impact of landscape
436 characteristics on landscape aesthetics. Our study is coherent with Tieskens et al. (2018), whom also
437 found that grass and water were positively predictive of high aesthetic value the close the relative
438 feature was to the geolocated photo. Though in contact, forest was also predictive, though is not
439 one of our significant predictors in our model.

440 Seresinhe et al. (2017b) used a neural network approach to extract features from images and used
441 scenic scores from an online platform. They found that mountain-related terms such as 'mountain
442 snowy', 'mountain, 'mountain' paths were positive predictors of scenic-ness, consistent with our
443 results. The term 'green' was also weakly positive, and the term 'farm' weakly negative, for both our
444 studies. Though on some terms, such as tundra. Hence it is suggested that use of automated
445 processes can create more consistency and aid in the removal of human error for feature extraction.

446 Whereas the regression model explained a large proposition of the variance, other factors in the
447 photographs must have contributed to aesthetic preference. Non-symmetrical landscape images are
448 preferred over symmetrical images, which was not analysed in our study (Bertamini et al., 2019).
449 This may explain why more unsymmetrical images ranked higher in the examples given in Figure 9.
450 Seasonal stages have also been found to influence preference of landscape elements, with flowering
451 stages most liked (Junge et al., 2015), and indeed the use of photographs from all seasons in this
452 study was a limitation. Coloured flowers, especially blue ones, have been found to be preferred
453 (Hůla and Flegr, 2016). Colours generally in the images may have impacted aesthetic preference. For
454 example, Huang and Lin (2019) have found that in mountainous landscapes, magenta-green colour
455 diversity is preferred.

456 By using only ten people in this study, similar in age and education level, with an interest in
457 geography as a subject, there was a potential limitation in not being representation of the general
458 population. Though all participants lived and work within 30km of the study area, hence were
459 assumed to have some familiarity of the area. While some studies have shown that landscape
460 preference can be impacted by socio-demographic factors (i.e. age, gender and residential
461 experience (Luo et al., 2019; Lyons, 1983)), Kalivoda et al. (2014) argues that positively rated
462 landscapes share broad consensus across socio-demographic characteristics. Indeed, while the
463 results in Figure 7 may be culturally context-dependent (because of our respondents' sample) the
464 similarity to PUDs (Figure 8) supports the argument in the latter study.

465 The upscaling and transferability of the method is dependent on the availability of photos for a given
466 landscape. LUCAS photos are available across the EU states, with social media data, including Flickr,
467 being available for 61% (of 5401) national parks globally (Tenkanen et al., 2017). The method
468 employed is restricted in that it only generates aesthetic values for those areas that are accessible,
469 or at least 'visually' accessible. Future work should address larger and more diverse photos and
470 respondents sample size, other photos' qualities such as symmetry and seasonality.

471 Being able to map aesthetic preferences for any given protected area important for management
472 decision making. van Zanten et al. (2016b) found that within agricultural landscapes hedgerows and
473 tree lines were valued for aesthetics and recreation value in the Netherlands, in Germany, it was
474 trees and crop diversity. Hence the values of aesthetic services can vary geographically, and hence
475 approaches such as ours that are highly context-relevant are important for advising management
476 decisions. Section 61 of the Environment Act 1995 assert that English National Parks should
477 'promote opportunities for the understanding and enjoyment of the special qualities [of the national
478 parks] by the public' (Yorkshire Dales National Park Authority, 2018a). With UK's exit to the Common
479 Agricultural Policy, a new land management system for farming system will have unknown
480 consequences for the parks landscape (Yorkshire Dales National Park Authority, 2018a). Hence it
481 becomes even more important to assess the current aesthetic assets that are held. This study
482 demonstrates an up-scalable approach that only uses free and publicly available photos and using
483 opens-sourced tools (mostly in R) for analysis.

484 **Conclusion**

485 Landscape aesthetics are nuanced. Our methods help to elucidate the subtle nature of landscape
486 aesthetics through combining techniques that have been used in other fields such as BT models and
487 machine learning. Aesthetics are both spatially and temporally explicit and subject to a myriad of
488 factors, but the use of public data and advanced algorithms opens a potential to successfully value
489 landscape aesthetics at large extents – which is key to prevent further degradation of our natural
490 heritage.

491 **Acknowledgements**

492 We would like to thank the Flickr users who contribute their images to the public sphere; without
493 whom this study would not be possible. We would also like to extend our thanks to our volunteer
494 participants, for taking both the time and effort to help the ideas behind our paper become reality.

495 **Funding:** This work was supported by ECOPotential under the European Union's Horizon 2020
496 research and innovation programme (grant agreement No. 641726). This research is a contribution
497 to the GEO BON working group on Ecosystem Services.

498

499 **References**

- 500 Agresti, A., 2002. An introduction to categorical data analysis, Second edi. ed. John Wiley & Sons,
501 Hoboken, New Jersey.
- 502 Ballin, M., Barcaroli, G., Masselli, M., Scarno, M., 2018. Redesign sample for Land Use/Cover Area
503 frame Survey (LUCAS) 2018.
- 504 Bertamini, M., Rampone, G., Makin, A.D.J., Jessop, A., 2019. Symmetry preference in shapes, faces,
505 flowers and landscapes. *PeerJ* 2019, 1–26. <https://doi.org/10.7717/peerj.7078>
- 506 Bradley, R.A., Terry, M.E., 1952. Rank analysis of incomplete block designs I: The method of paired
507 comparisons 39, 324–345.
- 508 Casado-Arzuaga, I., Madariaga, I., Onaindia, M., 2013. Perception, demand and user contribution to
509 ecosystem services in the Bilbao Metropolitan Greenbelt. *J. Environ. Manage.* 129, 33–43.
510 <https://doi.org/10.1016/j.jenvman.2013.05.059>
- 511 Casado-Arzuaga, I., Onaindia, M., Madariaga, I., Verbarg, P.H., 2014. Mapping recreation and
512 aesthetic value of ecosystems in the Bilbao Metropolitan Greenbelt (northern Spain) to support
513 landscape planning. *Landsc. Ecol.* 29, 1393–1405. <https://doi.org/10.1007/s10980-013-9945-2>
- 514 Casalegno, S., Inger, R., DeSilvey, C., Gaston, K.J., 2013. Spatial Covariance between Aesthetic Value
515 & Other Ecosystem Services. *PLoS One* 8, 6–10. <https://doi.org/10.1371/journal.pone.0068437>
- 516 Cheng, X., Van Damme, S., Li, L., Uyttenhove, P., 2019. Evaluation of cultural ecosystem services: A
517 review of methods. *Ecosyst. Serv.* 37, 100925. <https://doi.org/10.1016/j.ecoser.2019.100925>
- 518 Daniel, T.C., 2001. Whither scenic beauty? Visual landscape quality assessment in the 21st century.
519 *Landsc. Urban Plan.* 54, 267–281. [https://doi.org/10.1016/S0169-2046\(01\)00141-4](https://doi.org/10.1016/S0169-2046(01)00141-4)
- 520 de Groot, R., Ramakrishnan, P.S., 2005. Cultural and Amenity Services. *Ecosystems and Human Well-*
521 *Being: Current State and Trends. Findings of the Condition and Trends Working Group.*
522 *Millennium Ecosystem Assessment.*
- 523 de Groot, R.S., Ramakrishnan, P.S., Berg, A. van de, Kulenthran, T., Muller, S., Pitt, D., Wascher, D.,
524 Wijesuriya, G., Groot, R. de, Ramakrishnan, P.S., Berg, A. van de, Kulenthran, T., Muller, S., Pitt,
525 D., Wascher, D., Wijesuriya, G., 2005. Chapter 17, Cultural and Amenity Services. *Ecosyst. Hum.*
526 *well-being Curr. State Trends* 457–474. <https://doi.org/10.1016/j.fm.2010.10.016>
- 527 De Vries, S., Lankhorst, J.R.K., Buijs, A., 2007. Mapping the attractiveness of the Dutch countryside: a
528 GIS-based landscape appreciation model. *For. Snow Landsc. Res.* 43–57.
- 529 DeLucio, J. V., Múgica, M., 1994. Landscape preferences and behaviour of visitors to Spanish national
530 parks. *Landsc. Urban Plan.* 29, 145–160. [https://doi.org/10.1016/0169-2046\(94\)90024-8](https://doi.org/10.1016/0169-2046(94)90024-8)
- 531 Dijksterhuis, A., 2004. Think different: The merits of unconscious thought in preference development
532 and decision making. *J. Pers. Soc. Psychol.* 87, 586–598. [https://doi.org/10.1037/0022-](https://doi.org/10.1037/0022-3514.87.5.586)
533 [3514.87.5.586](https://doi.org/10.1037/0022-3514.87.5.586)
- 534 Dijksterhuis, A., Nordgren, L., 2008. A Theory of Unconscious Thought. *Perspect. Psychol. Sci.* 1, 95–
535 109. <https://doi.org/10.1111/j.1745-6916.2006.00007.x>
- 536 Dijksterhuis, A., van Olden, Z., 2006. On the benefits of thinking unconsciously: Unconscious thought
537 can increase post-choice satisfaction. *J. Exp. Soc. Psychol.* 42, 627–631.
538 <https://doi.org/10.1016/j.jesp.2005.10.008>
- 539 Dramstad, W.E., Tveit, M.S., Fjellstad, W.J., Fry, G.L.A., 2006. Relationships between visual landscape

540 preferences and map-based indicators of landscape structure. *Landsc. Urban Plan.* 78, 465–
541 474. <https://doi.org/10.1016/j.landurbplan.2005.12.006>

542 Eurostat, 2020. Lucas Photos Order Form [WWW Document]. *Comm. Eur.* URL
543 <https://ec.europa.eu/eurostat/web/lucas/data/primary-data/order-form>

544 Fagerholm, N., Käyhkö, N., 2009. Participatory mapping and geographical patterns of the social
545 landscape values of rural communities in Zanzibar , Tanzania.

546 Feinerer, I., Hornik, K., 2018. tm: Text Mining Package.

547 Feinerer, I., Hornik, K., 2017. wordnet: WordNet Interface. R package version 0.1-14.

548 Feinerer, I., Hornik, K., Meyer, D., 2008. Text mining infrastructure in R. *J. Stat. Softw.* 25, 1–54.
549 <https://doi.org/10.18637/jss.v025.i05>

550 Fellbaum, C., 1998. WordNet: An Electronic Lexical Database. The MIT Press, Cambridge, MA.

551 Figueroa-Alfaro, R.W., Tang, Z., 2017. Evaluating the aesthetic value of cultural ecosystem services
552 by mapping geo-tagged photographs from social media data on Panoramio and Flickr. *J.*
553 *Environ. Plan. Manag.* 60, 266–281. <https://doi.org/10.1080/09640568.2016.1151772>

554 Frank, S., Fürst, C., Koschke, L., Witt, A., Makeschin, F., 2013. Assessment of landscape aesthetics -
555 Validation of a landscape metrics-based assessment by visual estimation of the scenic beauty.
556 *Ecol. Indic.* 32, 222–231. <https://doi.org/10.1016/j.ecolind.2013.03.026>

557 Ghermandi, A., Sinclair, M., 2019. Passive crowdsourcing of social media in environmental research:
558 A systematic map. *Glob. Environ. Chang.* 55, 36–47.
559 <https://doi.org/10.1016/j.gloenvcha.2019.02.003>

560 Gobster, P.H., Nassauer, J.I., Daniel, T.C., Fry, G., 2007. The shared landscape: What does aesthetics
561 have to do with ecology? *Landsc. Ecol.* 22, 959–972. <https://doi.org/10.1007/s10980-007-9110->
562 [x](https://doi.org/10.1007/s10980-007-9110-x)

563 Google Cloud Vision, 2019. Documentation for the Google Cloud Vision API [WWW Document]. URL
564 www.cloud.google.com/vision/ (accessed 7.27.19).

565 Gosal, A.S., Geijzendorffer, I.R., Václavík, T., Poulin, B., Ziv, G., Republic, C., Valat, T., 2019. Using
566 social media, machine learning and natural language processing to map multiple recreational
567 beneficiaries Arjan S. Gosal. *Ecosyst. Serv.* 38, 100958.
568 <https://doi.org/10.1016/j.ecoser.2019.100958>

569 Gosal, A.S., Newton, A.C., Gillingham, P.K., 2018. Comparison of methods for a landscape-scale
570 assessment of the cultural ecosystem services associated with different habitats. *Int. J.*
571 *Biodivers. Sci. Ecosyst. Serv. Manag.* 14, 91–104.
572 <https://doi.org/10.1080/21513732.2018.1447016>

573 Hägerhäll, C.M., Sang, Å.O., Englund, J.E., Ahlner, F., Rybka, K., Huber, J., Burenhult, N., 2018. Do
574 humans really prefer semi-open natural landscapes? A cross-cultural reappraisal. *Front.*
575 *Psychol.* 9, 1–14. <https://doi.org/10.3389/fpsyg.2018.00822>

576 Haines-Young, R., Potschin, M., 2018. CICES V5. 1. Guidance on the Application of the Revised
577 Structure. *Fabis Consult.* 53.

578 Heine, S.J., Lehman, D.R., Peng, K., Greenholtz, J., 2002. What’s wrong with cross-cultural
579 comparisons of subjective Likert scales?: The reference-group effect. *J. Pers. Soc. Psychol.* 82,
580 903–918. <https://doi.org/10.1037/0022-3514.82.6.903>

- 581 Hermes, J., Albert, C., von Haaren, C., 2018. Assessing the aesthetic quality of landscapes in
582 Germany. *Ecosyst. Serv.* 31, 296–307. <https://doi.org/10.1016/j.ecoser.2018.02.015>
- 583 Huang, A.S.H., Lin, Y.J., 2019. The effect of landscape colour, complexity and preference on viewing
584 behaviour. *Landsc. Res.* 00, 1–14. <https://doi.org/10.1080/01426397.2019.1593336>
- 585 Hůla, M., Flegr, J., 2016. What flowers do we like? The influence of shape and color on the rating of
586 flower beauty. *PeerJ* 2016. <https://doi.org/10.7717/peerj.2106>
- 587 Iyengar, S.S., Lepper, M.R., 2000. When choice is demotivating: Can one desire too much of a good
588 thing? *J. Pers. Soc. Psychol.* 79, 995–1006. <https://doi.org/10.1037/0022-3514.79.6.995>
- 589 Junge, X., Schüpbach, B., Walter, T., Schmid, B., Lindemann-Matthies, P., 2015. Aesthetic quality of
590 agricultural landscape elements in different seasonal stages in Switzerland. *Landsc. Urban Plan.*
591 133, 67–77. <https://doi.org/10.1016/j.landurbplan.2014.09.010>
- 592 Kalivoda, O., Vojar, J., Skřivanová, Z., Zahradník, D., 2014. Consensus in landscape preference
593 judgments: The effects of landscape visual aesthetic quality and respondents' characteristics. *J.*
594 *Environ. Manage.* 137, 36–44. <https://doi.org/10.1016/j.jenvman.2014.02.009>
- 595 Kaplan, R., Kaplan, S., 1989. *The Experience of Nature: A Psychological Perspective*. Cambridge
596 University Press, Cambridge, UK.
- 597 Karasov, O., Vieira, A.A.B., Külvik, M., Chervanyov, I., 2020. Landscape coherence revisited: GIS-
598 based mapping in relation to scenic values and preferences estimated with geolocated social
599 media data. *Ecol. Indic.* 111, 105973. <https://doi.org/10.1016/j.ecolind.2019.105973>
- 600 Kenter, J., Reed, M.S., Irvine, K.N., O'Brien, L., 2014. National Ecosystem Assessment Follow-on.
601 Work Package 6: Shared, plural and cultural values of ecosystems.
- 602 Kienast, F., Frick, J., Strien, M.J. Van, Hunziker, M., 2015. The Swiss Landscape Monitoring Program –
603 A comprehensive indicator set to measure landscape change. *Ecol. Modell.* 295, 136–150.
604 <https://doi.org/10.1016/j.ecolmodel.2014.08.008>
- 605 Langemeyer, J., Calcagni, F., Baró, F., 2018. Mapping the intangible: Using geolocated social media
606 data to examine landscape aesthetics. *Land use policy* 77, 542–552.
607 <https://doi.org/10.1016/j.landusepol.2018.05.049>
- 608 Lee, J., 2015. *Yorkshire Dales*. William Collins, London, UK.
- 609 Lovato, P., Perina, A., Cheng, D.S., Segalin, C., Sebe, N., Cristani, M., 2013. We like it! Mapping image
610 preferences on the counting grid, in: 2013 IEEE International Conference on Image Processing,
611 ICIP 2013 - Proceedings. pp. 2892–2896. <https://doi.org/10.1109/ICIP.2013.6738595>
- 612 Luo, T., Xu, M., Wang, Z., Yu, Z., 2019. A comparative study on the performance of three color
613 schemes in landscape preference tests. *J. Environ. Eng. Landsc. Manag.* 27, 114–125.
614 <https://doi.org/10.3846/jeelm.2019.9805>
- 615 Lyons, E., 1983. Demographic correlates of landscape preference. *Environ. Behav.* 15, 487–511.
- 616 Masuda, T., Gonzalez, R., Kwan, L., Nisbett, R.E., 2008. Culture and aesthetic preference: Comparing
617 the attention to context of East Asians and Americans. *Personal. Soc. Psychol. Bull.* 34, 1260–
618 1275. <https://doi.org/10.1177/0146167208320555>
- 619 Ode, Å., Hagerhall, C.M., Sang, N., 2010. Analysing visual landscape complexity: Theory and
620 application. *Landsc. Res.* 35, 111–131. <https://doi.org/10.1080/01426390903414935>
- 621 Oteros-Rozas, E., Martín-López, B., Fagerholm, N., Bieling, C., Plieninger, T., 2018. Using social media

622 photos to explore the relation between cultural ecosystem services and landscape features
623 across five European sites. *Ecol. Indic.* 94, 74–86.
624 <https://doi.org/10.1016/j.ecolind.2017.02.009>

625 Palmer, J.F., 2004. Using spatial metrics to predict scenic perception in a changing landscape : Dennis
626 , Massachusetts 69.

627 Peña, L., Casado-Arzuaga, I., Onaindia, M., 2015. Mapping recreation supply and demand using an
628 ecological and a social evaluation approach. *Ecosyst. Serv.* 13, 108–118.
629 <https://doi.org/10.1016/j.ecoser.2014.12.008>

630 RStudio Team, 2015. RStudio: Integrated Development for R.

631 Schirpke, U., Tappeiner, G., Tasser, E., Tappeiner, U., 2019. Using conjoint analysis to gain deeper
632 insights into aesthetic landscape preferences. *Ecol. Indic.* 96, 202–212.
633 <https://doi.org/10.1016/j.ecolind.2018.09.001>

634 Schirpke, U., Tasser, E., Tappeiner, U., 2013. Landscape and Urban Planning Predicting scenic beauty
635 of mountain regions. *Landsc. Urban Plan.* 111, 1–12.
636 <https://doi.org/10.1016/j.landurbplan.2012.11.010>

637 Seresinhe, C.I., Moat, H.S., Preis, T., 2017a. Quantifying scenic areas using crowdsourced data.
638 *Environ. Plan. B Urban Anal. City Sci.* 0265813516687302.
639 <https://doi.org/10.1177/0265813516687302>

640 Seresinhe, C.I., Preis, T., Moat, H.S., 2017b. Using deep learning to quantify the beauty of outdoor
641 places. *R. Soc. Open Sci.* 4, 170170. <https://doi.org/10.1098/rsos.170170>

642 Seresinhe, C.I., Preis, T., Moat, H.S., 2015. Quantifying the impact of scenic environments on health.
643 *Sci. Rep.* 5, 16899. <https://doi.org/10.1038/srep16899>

644 Tahvanainen, L., Tyrväinen, L., Ihalainen, M., Vuorela, N., Kolehmainen, O., 2001. Forest
645 management and public perceptions - Visual versus verbal information. *Landsc. Urban Plan.* 53,
646 53–70. [https://doi.org/10.1016/S0169-2046\(00\)00137-7](https://doi.org/10.1016/S0169-2046(00)00137-7)

647 Tenerelli, P., Püffel, C., Luque, S., 2017. Spatial assessment of aesthetic services in a complex
648 mountain region: combining visual landscape properties with crowdsourced geographic
649 information. *Landsc. Ecol.* 32, 1097–1115. <https://doi.org/10.1007/s10980-017-0498-7>

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

653 Tieskens, K.F., Zanten, B.T. Van, Schulp, C.J.E., Verburg, P.H., 2018. Landscape and Urban Planning
654 Aesthetic appreciation of the cultural landscape through social media : An analysis of revealed
655 preference in the Dutch river landscape. *Landsc. Urban Plan.* 177, 128–137.
656 <https://doi.org/10.1016/j.landurbplan.2018.05.002>

657 Turner, H., Firth, D., 2012. Bradley-terry models in R: The bradleyterry2 package. *J. Stat. Softw.* 48.

658 Tyrväinen, L., Silvennoinen, H., Kolehmainen, O., 2003. Ecological and aesthetic values in urban
659 forest management. *Urban For. Urban Green.* 1, 135–149. <https://doi.org/10.1078/1618-8667-00014>

661 van Zanten, B.T., Van Berkel, D.B., Meentemeyer, R.K., Smith, J.W., Tieskens, K.F., Verburg, P.H.,
662 2016a. Continental-scale quantification of landscape values using social media data. *Proc. Natl.*
663 *Acad. Sci.* 113, 12974–12979. <https://doi.org/10.1073/pnas.1614158113>

664 van Zanten, B.T., Zasada, I., Koetse, M.J., Ungaro, F., Häfner, K., Verburg, P.H., 2016b. A comparative
665 approach to assess the contribution of landscape features to aesthetic and recreational values
666 in agricultural landscapes. *Ecosyst. Serv.* 17, 87–98.
667 <https://doi.org/10.1016/j.ecoser.2015.11.011>

668 Volkmann, J., 1951. Scales of judgment and their implications for social psychology, in: Rohrer, J.H.,
669 Sherif, M. (Eds.), *Social Psychology at the Crossroads*. Harper, New York.

670 Wilson, M., Jarrett, D., Calladine, J., 2018. Review of data and monitoring options for assessing the
671 status of breeding wader populations in the Yorkshire Dales National Park. Thetford, UK.

672 Yorkshire Dales National Park Authority, 2018a. Yorkshire Dales Management Plan 2019-24.

673 Yorkshire Dales National Park Authority, 2018b. Trends in tourism in the Yorkshire (2016 to 2017).
674 Leyburn, UK.

675 Yorkshire Dales National Park Authority, 2017. Local Biodiversity Action Plan: Trends and Status
676 Review 2016.

677 Yoshimura, N., Hiura, T., 2017. Demand and supply of cultural ecosystem services: Use of geotagged
678 photos to map the aesthetic value of landscapes in Hokkaido. *Ecosyst. Serv.* 24, 68–78.
679 <https://doi.org/10.1016/j.ecoser.2017.02.009>

680 Zermelo, E., 1929. Die berechnung der turnier-ergebnisse als ein maximumproblem der
681 wahrscheinlichkeitsrechnung. *Math. Zeitschrift* 29.

682 Zucco, C., Batista, M., Power, T.J., 2019. Measuring portfolio salience using the Bradley–Terry model:
683 An illustration with data from Brazil. *Res. Polit.* 6. <https://doi.org/10.1177/2053168019832089>

684

685