

Article

Identifying and Mapping Groups of Protected Area Visitors by Environmental Awareness

Arjan S. Gosal ^{1,*} , Janine A. McMahon ¹, Katharine M. Bowgen ² , Catherine H. Hoppe ³ and Guy Ziv ¹ 

¹ School of Geography, Faculty of Environment, University of Leeds, Leeds LS2 9JT, UK; janine@ecologyresources.co.uk (J.A.M.); G.Ziv@leeds.ac.uk (G.Z.)

² British Trust for Ornithology, The Nunnery, Thetford IP24 2PU, UK; katharine.bowgen@bto.org

³ Independent Researcher, 31515 Wunstorf, Germany; hoppec82@gmail.com

* Correspondence: a.gosal@gmx.co.uk

Abstract: Protected areas worldwide receive billions of visitors annually. The positive impact of nature on health and wellbeing, in addition to providing opportunities for cultural activities such as recreation and aesthetic appreciation, is well documented. Management to reduce negative impacts to biodiversity and conservation aims whilst providing amenities and access to visitors is important. Understanding environmental awareness of visitors and their on-site spatial patterns can assist in making effective management decisions within often constrained resources. However, there is a lack of strategies for site-specific identification and predictive mapping of visitors by environmental awareness. Here, we demonstrate a method to map on-site visitation by latent groups of visitors based on their environmental awareness of on-site issues. On-site surveys and participatory mapping were used to collect data on environmental awareness on bird nesting and spatial visitation patterns in an upland moor in northern England. Latent class analysis (LCA), a structural equation model, was used to discover underlying groups of environmental awareness, with random forest (RF) modelling, a machine learning technique, using a range of on-site predictors (bioclimatic, land cover, elevation, viewshed, and proximity to paths and freshwater) to predict and map visitation across the site by each group. Visitors were segmented into ‘aware’ and ‘ambiguous’ groups and their potential spatial visitation patterns mapped. Our results demonstrate the ability to uncover groups of users by environmental awareness and map their potential visitation across a site using a variety of on-site predictors. Spatial understanding of the movement patterns of differently environmentally aware groups of visitors can assist in efficient targeting of conservation education endeavours (i.e., signage, positioning of staff, monitoring programmes, etc.), therefore maximising their efficacy. Furthermore, we anticipate this method will be of importance to environmental managers and educators when deploying limited resources.

Keywords: environmental awareness; machine learning; random forest model; structural equation modelling; latent class analysis; visitor mapping



Citation: Gosal, A.S.; McMahon, J.A.; Bowgen, K.M.; Hoppe, C.H.; Ziv, G. Identifying and Mapping Groups of Protected Area Visitors by Environmental Awareness. *Land* **2021**, *10*, 560. <https://doi.org/10.3390/land10060560>

Academic Editors: Juan F. Beltrán, Pedro Abellán and John Litvaitis

Received: 24 April 2021

Accepted: 22 May 2021

Published: 27 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Much of the environmental deterioration evidenced around the world, including the alarming rate of biodiversity loss, can be attributed to humans [1,2]. Indeed, Schultz [2] argues that only through changing human behaviour can conservation have a chance of success and that identifying those behaviours that need to change, rather than broad education and awareness-raising campaigns, should be the priority. With visitation to protected areas expected to increase [3], already stretched resources for improving visitors' awareness of key site objectives, such as protecting bird species, will only need to go further.

The impact of human disturbance on birds can manifest in multiple ways, through a combination of changing distributions, behaviour, demography or population size [4]. For example, during sensitive periods of the year, such as territory establishment, even low levels of human disturbance can alter bird behaviour [5]. Bird disturbance caused by

humans has been studied in various habitats, including forests [6,7], coastal areas [8], and upland habitats such as moorland [9,10]. The risk of negative impacts to birds is considered serious enough to be written into international law prohibiting deliberate disturbance (EU Birds Directive 2009).

Separating nature area visitors into groups has been explored for several scenarios in literature, including segmentation by motivation for visiting a site, to best respond to visitor needs [11,12]. Several studies have used a market segmentation approach to group visitors [13,14], with Halpenny [15] investigating place attachments as a predictor of pro-environmental behaviours. Conservation social sciences can be very useful in guiding conservation actions and outcomes that are effective and robust [16]. For example, Booth et al. [17] found substantial variation in protected status awareness in site visitors, with only conservation organisation membership being of predictive significance for being more informed.

Context-specific maps can improve and/or enable replication of strategic environmental assessment mapping to aid conservation decision-making [18], and knowledge of visitor preferences offers insights to inform destination development and promotion of ecotourism within a site [19]. In particular, mapping to support environmental decision-making provides an approach to understand where policy interventions could be most effective [20] and is a valid mechanism to translate scientific findings into tangible products that can be used by local on-site practitioners. Conservation education is the mechanism through which awareness and concern for the environment are raised and can take several forms, from free-choice learning to structured initiatives [16,21,22]. Understanding environmental awareness on-site can help in personal delivery of on-site minimum impact education strategies, as messages to park visitors have been found to minimise off-trail behaviour, whilst physical signs were found to be ineffective in preventing this behaviour [23].

Identifying visitation patterns of nature site users lacking in awareness of on-site issues is important. As lower awareness may result in unwanted behaviour, such as disturbance to birds or other wildlife, predicting areas these users may visit on-site is vital for both management and mitigation strategies (e.g., limiting access to sensitive areas or targeted education campaigns). In this study, we aim to provide a replicable method that can be used to both define and spatially map different types of protected area visitors, based on their environmental awareness of on-site issues. We use latent class analysis (LCA), a type of structural equation model, to group visitors based on environmental awareness and random forest (RF) modelling, a machine learning method, with participatory mapping data to produce predictive maps of visitation. This is demonstrated for a protected English upland moor that has importance for protected bird populations.

2. Materials and Methods

2.1. Study Site and Context

Ilkley Moor, an upland moor in northern England, UK (53°54′11.1″ N, 1°49′36.9″ W), Figure 1, is dominated by heather moorland over acidic soils from the Millstone Grit underlying it [22]. It is 402 m above sea level, which exposes it to harsh winds in the winter. Its nature conservation importance is recognised through international and national designations, including Site of Special Scientific Interest, Special Area of Protection, and Special Area of Conservation. In addition to being of high ecological importance, the site provides cultural ecosystem services such as recreation and sense of place, as well as providing regulating and supporting services including carbon storage in peat, water and air purification, and floodwater retention.

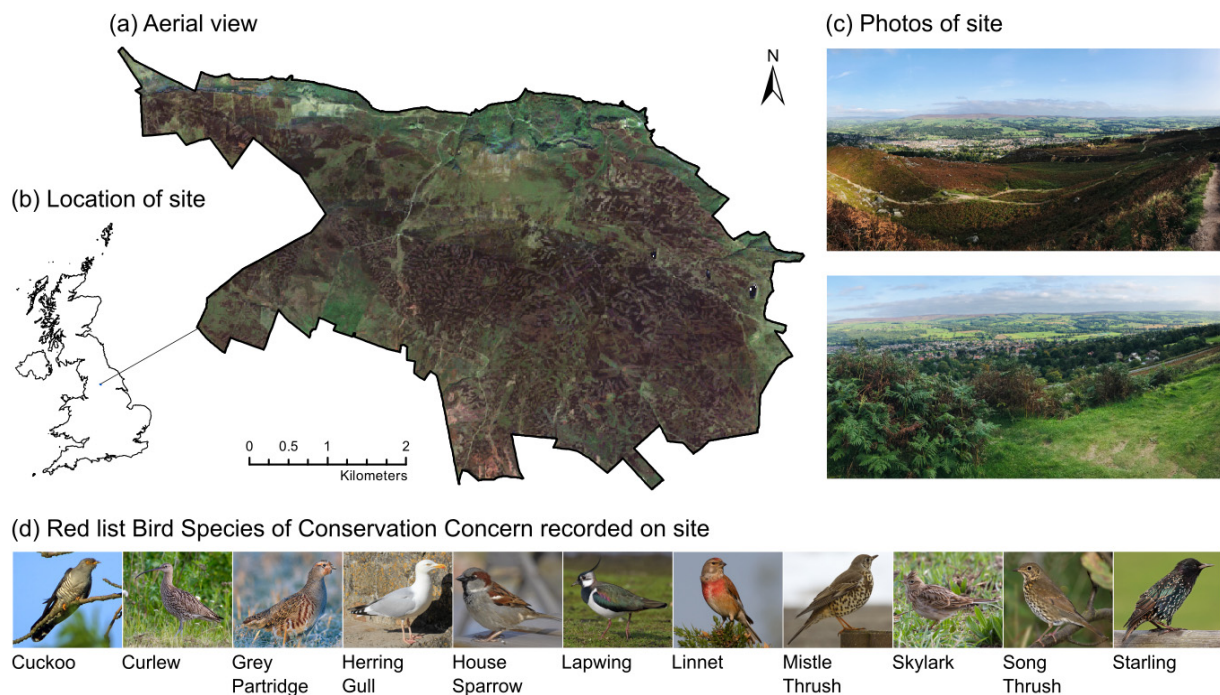


Figure 1. The study site in multiple dimensions: (a) aerial view using Sentinel 2 natural colour (bands 432) mosaic imagery [23], (b) location of site in relation to the British Isles, (c) example photographs taken on-site, with surrounding landscape views, and (d) Red List Species of Conservation Concern [24,25] recorded on the site between 2015 and 2019 during the British Trust for Ornithology (BTO) British Breeding Bird Surveys [26]. The BTO/JNCC/RSPB Breeding Bird Survey is a partnership jointly funded by the BTO, RSPB, and JNCC, with fieldwork conducted by volunteers. All bird images are published on www.wikimedia.org and licensed under Creative Commons; see Table S1 for full attributions. © contains Copernicus data (2019) and JNCC/NE/NRW/SNH/NIEA data, © copyright and database right 2019.

In this study, we specifically consider environmental and ecological awareness regarding nesting birds on Ilkley Moor, which experiences large numbers of visitors. During the bird breeding season, the implications of disturbance for outcomes of nesting attempts range across multiple impacts, including nest failure [24], impaired nestling growth [7], reduction in the areas suitable for breeding [25], and immunosuppression of fledglings [26], which all put pressure on individual birds and future recruitment into the local population [27]. The strength of these impacts can vary widely with species and type of disturbance; they may only have mild effects on fitness or cause total breeding season failure [28,29]. Hence, it is important for disturbance at this time to be kept to a minimum.

The BTO's Breeding Bird Survey [30] has recorded 70 bird species on Ilkley Moor in the last 10 years, declining to 63 in the last 5 years. The seven lost species include the chiffchaff (*Phylloscopus collybita*), coal tit (*Periparus ater*), collared dove (*Streptopelia decaocto*), and hobby (*Falco subbuteo*); two Red listed species, the grasshopper warbler (*Locustella naevia*) and tree sparrow (*Passer montanus*); and one Amber listed species, the house martin (*Delichon urbicum*) [31]. In the last five years, 16 Amber listed species have been recorded (see Table S2 for full lists of bird species that can be found on-site).

2.2. Independent Variables; Survey and Participatory Mapping

A total of 124 surveys were conducted in the summer of 2019 with visitors to the study site. The survey contained questions covering demographics and questions related to participants' experiences on-site. Here we focus on five environmental awareness questions (Table 1) structured in the five-point Likert format [32,33] and a participatory mapping exercise, where the participants were asked to mark on a map the areas of the site they had visited on that specific visit. Through permission of the site owner (Bradford Council), 11 access points were used for surveying, where 10 out of the 11 had car parks. Stratified

sampling was used to identify a minimum of one access point per grid square by draping a 4×4 grid over an Ordnance Survey map of the site. Surveys took place between 09:00 and 17:00 BST, with people interviewed by a surveyor transcribing the responses to the questionnaire.

Table 1. Environmental awareness Likert questions used in the surveys and the abbreviated form referred to in this study.

Environmental Awareness Question	Abbreviated Form
Are you aware that Ilkley Moor is a South Pennine Moors Site of Special Scientific Interest (SSSI), a Special Protection Area (SPA), and a Special Area of Conservation (SAC)?	Site designations
Are you aware that the nesting season for birds is between March and July every year (5 months)?	Nesting season
Are you aware of the ecological importance of the nesting season in March–July every year (5 months) for birds such as the golden plover, red grouse, lapwing, and short eared owls?	Eco. bird nesting
Are you aware of the issues that off-lead dogs can cause to the nesting birds?	Issues of off-leading
Were you previously aware of the request to keep dogs on a lead during the period of March–July (5 months)?	Requested on-leading

The responses to the environmental questions (Table 1) were dichotomously recoded (with Likert ratings 1, 2, and 3 being “unaware” and 4 and 5 being “aware”). As a rating of 3 could be perceived as a neutral response, it was assumed this indicated a lack of awareness of the issue, rather than awareness. During the participatory mapping exercise, participants marked points, lines, or polygons with a pen on a map to indicate the areas they visited. Maps were digitised in ESRI ArcMap v10.6. Points and lines were converted to polygons using a 250 m buffer before all polygons for all participants were converted to a binary raster (25 m per pixel) to indicate visitation per participant. These were summed later by the allocated LCA group to produce visitation maps.

2.3. Secondary Variables; Independent Secondary Data

To upscale from the participatory survey data to a map covering the entire region, we used an RF algorithm with a set of climatic and physical factors as predictors. Bioclimatic variables were sourced from WorldClim version 2.1 [34], in the form of 19 variables derived from monthly rainfall and temperature (1970–2000) at the finest resolution available of 30 s as GeoTiff files. Variables were extracted as values from the centroid point per 25 m pixel across the site.

Land cover (such as forest) was used as the basis for the 25 m pixel size for this study and sourced from CEH LCM2019 [35] based on 21 land cover types at a 25 m by 25 m resolution raster map. Dummy binary variables were created for each pixel for the classes present over the site (deciduous woodland, coniferous woodland, improved grassland, acid grassland, heather, heather grassland, bog, freshwater, and suburban). Elevation was calculated from Ordnance Survey 50 m digital terrain model [36]. Proximity to paths and freshwater was calculated per pixel in meters in QGIS v3.16.0 [37] using OpenStreetMap data retrieved from Geofabrik.de [38].

Viewshed analysis (delineation of the area visible from any given location on a map) was calculated in QGIS v3.16.0 [37] using the “Advanced Viewshed Analysis” v1.4 plugin [39]. The site visibility index was calculated as a cumulative binary viewshed utilising only visibility points from within the site and the Ordnance Survey 50 m digital terrain model [36]. Visibility points were created at the centroid location of each 25 m pixel, with the default values of observer height set at 1.6 m and a 5 km radius used. This allowed consideration of whether on-site visibility of an area contributed to visitation by differently environmentally aware users.

2.4. Statistic Analysis

All statistical analysis was conducted in R (R Core Team, 2020). See Figure 2 for a flowchart of the study methodology.

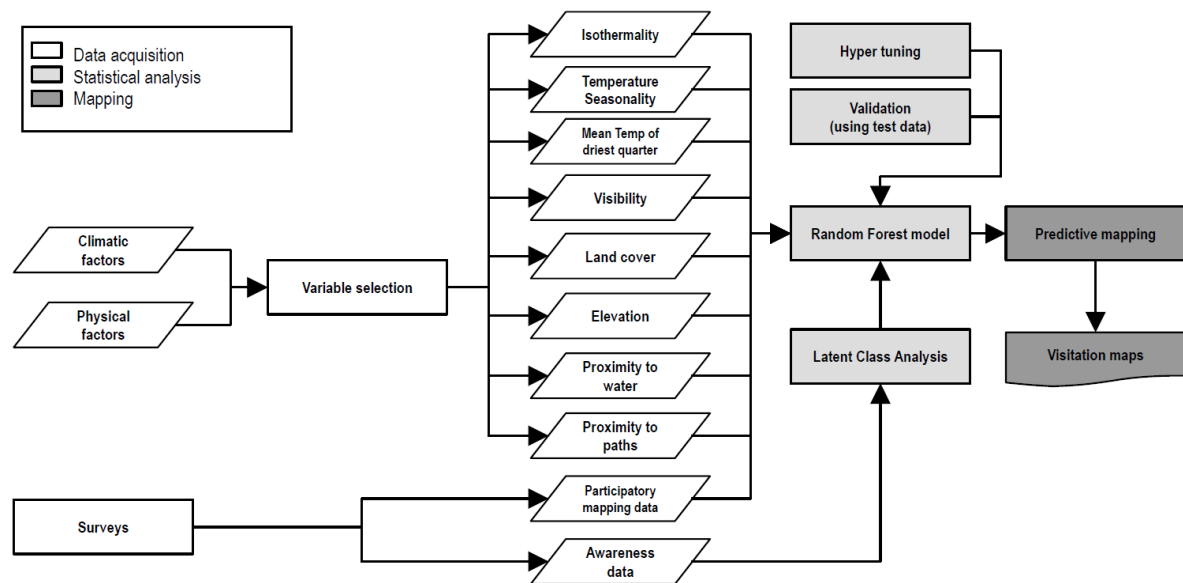


Figure 2. Flowchart of study methodology, showing the progression from data acquisition and selection through to the latent class analysis (LCA) and random forest (RF) model and finishing in mapping. White boxes refer to data acquisition and selection, light grey indicates statistical analysis, and dark grey highlights the mapping stages.

2.4.1. Variable Selection

While RF models suffer less from collinearity than other traditional statistical methods, multiple predictors that are weakly correlated to the response, and strongly with each other, can cause unstable results, with it being unlikely that the averaging that occurs across all trees in an RF being able to overcome this [40,41]. We dealt with collinearity through several steps. Firstly, the *Caret* R package [42] was used to identify variables that were correlated at 0.75 or higher with each other, and the variable with the largest mean absolute correlation amongst all variables was removed. Secondly, variance inflation factors (VIFs) were calculated, and the variable with the highest VIF was removed until the cohort all had values under 10, resulting in fifteen predictors. Thirdly, the *Boruta* R package [43] was used to check for “variable importance” (the utilisation by the RF model to use a given variable to make accurate predictions; higher usage to make predictions implies a greater importance for the model). *Boruta* does this through iterative removal of variables found to be statistically less relevant than random probes using RF [43].

2.4.2. Latent Class Analysis

Latent class analysis (LCA) has been used in the study of environmental issues, including environmentally sustainable food choices [44], environmental concern of the UK population [45], and recreational demand based on attitudes towards water resources [46]. Here we use LCA to segment visitors into different environmental awareness groups. The *poLCA* R package [47] was used for the LCA analysis. Multiple classes were explored, with the lowest BIC dictating the class size of two (see Figure S1). LCA posterior probabilities were used to segment the surveyed individuals into the two classes (‘aware’ and ‘ambiguous’). Classes were named on the resultant probability of binary response to any given environmental question.

2.4.3. Random Forest Modelling

Since its first introduction by Breiman in 2001 random forest (RF) modelling, a machine learning technique, has been used extensively in the literature, e.g., in using semantic information to classify urban buildings [48], forecasting power consumption using hybrid models including RF [49], and modelling travel mode choice behaviour [50]. Braun, Cottrell, and Dierkes [51] used RF to investigate the effect of outdoor education programmes in

school children across multiple countries. The predictive accuracy of RF lends the method to producing potentiality or susceptibility maps. Naghibi, Pourghasemi, and Dixon [52] used RF with boosted, classification, and regression trees to produce groundwater potential maps. Elsewhere, RF has been used for landslide susceptibility mapping [53–55], soil parent material and carbon mapping [56,57], and vegetation mapping and land cover classification [58–60]. In this study, we use RF to produce potential maps of visitation for different environmental awareness groups.

The ranger R package [61] was used to run separate RF models for ‘all visitors’ and both LCA classes, using the sum of the binary visitation maps per pixel as the response variable. Only pixels visited by at least one survey respondent were included. Each model was hyper-tuned to find the model with the best predictive accuracy running a grid search for the number of variables sampled from at each split (1–15), node size (3, 5, 7, and 9), and sample size (0.550, 0.632, 0.700, and 0.800), resulting in 240 variants for each model. All models were run with 500 trees and showed stabilisation of out-of-bag (OOB) error within this number (see Table S3). The models with the lowest OOB error were chosen.

2.4.4. Validation and Mapping

RF models are robust in providing out-of-bag estimation of error, though as an additional layer of validation, data for the RF models were randomly split into testing (20%) and training datasets (80%). Predicted values from the RF models were mapped at 25 m pixel resolution using ESRI ArcGIS Pro 2.7.0. Due to heavy-tail distribution of the data, an m-out-of-n bootstrap was used to discern statistical difference between the maps using the R package distillery [62] with Pearson correlation tests for a random set of 1000 subsampled points, bootstrapped over 1000 iterations with resampling.

3. Results

The LCA analysis resulted in two classes, as this presented the lowest AIC/BIC (Figure S1); these were named ‘aware’ and ‘ambiguous’ following the probability of the first group being likely to be aware across all questions, with the latter having more ambiguity in probability of awareness (Figure 3). Posterior probabilities showed a split of 63.5% “aware” and 35.5% “ambiguous” amongst the surveyed visitors.

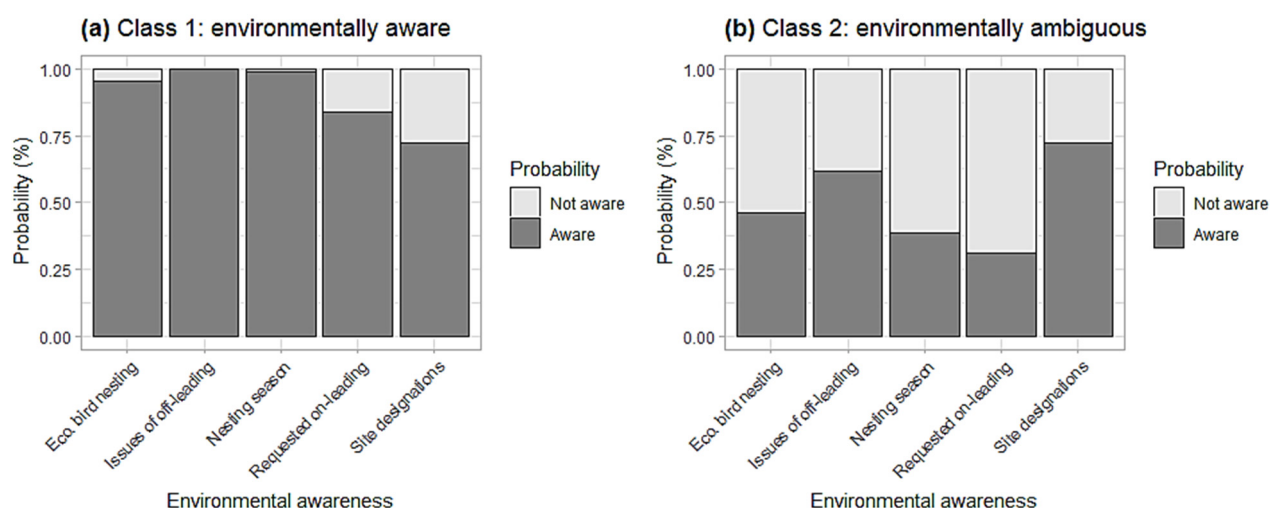


Figure 3. Two latent classes of environmental awareness were identified amongst the visitors: (a) environmentally aware and (b) environmentally ambiguous.

Demographic and behavioural responses in full can be found in Table 2. The RF model was performed for all visitors, and the ‘aware’ and ‘ambiguous’ groups were identified from the LCA. All three models had high accuracy (all visitors and ‘ambiguous’ groups had an OOB R^2 of 0.97, with ‘aware’ having an OOB R^2 of 0.96). OOB prediction error

(MSE) was lowest for the ‘ambiguous’ group (0.36), followed by ‘aware’ (1.20) and the ‘all visitors’ group (2.56), with the error showing stabilisation within 500 RF trees (see Figure S2 and Table S3). Root mean square error (RMSE) was calculated on the 20% holdback validation dataset for all models: ‘all visitors’ (1.61), ‘aware’ (1.10), and ‘ambiguous’ (0.60).

Variable importance was computed for all RF models; variables with higher importance levels contribute to explaining the outcome the most (Figure 4). Mean temperature in the driest quarter was the most important for all models. ‘All visitors’ and ‘aware’ visitors shared similar levels and rank of variable importance, with temperature seasonality being second most important, whereas this was moderately important for the other group. For the ‘ambiguous’ group the second most important variable was isothermality. Isothermality quantifies the extent to which day-to-night temperatures oscillate relative to summer-to-winter annual oscillations [63]. Proximity to water, elevation, and viewshed were also moderately important for all groups. Heather grassland and proximity to paths were important to a lesser extent, and the remaining variables (deciduous woodland, coniferous woodland, suburban, improved grassland, freshwater, bog, and arid grassland) showed very low importance.

Table 2. Demographic and behavioural responses within environmentally aware and ambiguous visitors in percentages. Percentages within each variable group for each environmental awareness group sum to 100.

Variable	Environmentally Aware (%)	Environmentally Ambiguous (%)
Age		
18–30	5.00	22.70
31–40	7.50	6.80
41–50	17.50	15.90
51–60	27.50	20.50
61–70	18.80	25.00
70+	23.70	9.10
Gender		
Female	47.50	29.55
Male	52.50	70.45
Travel distance		
Less than 1 mile	31.25	25.00
1 to 5 miles	28.75	31.82
More than 5 miles	40.00	43.18
Frequency of visit		
Daily	26.25	9.09
2–3 times a week	20.00	20.45
Once a week	10.00	15.91
Twice a month	2.50	2.27
Once a month	7.50	13.64
1–2 times a year	5.00	9.09
2–3 times a year	8.75	4.55
Once every 2–3 years	1.25	2.27
Sporadically	16.25	9.09
First visit	2.50	13.64

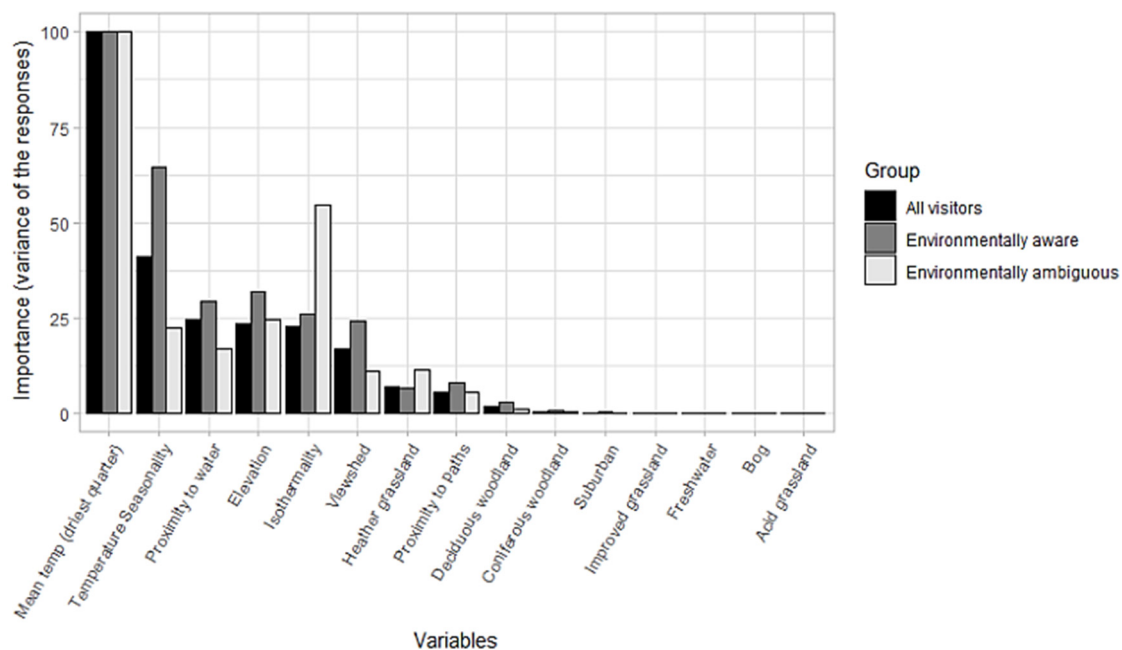


Figure 4. Random forest variable importance (values standardised between 0 and 100 within each group for comparability, ordered by descending ‘all visitors’ values). ‘All visitors’ group is shown in black, ‘environmentally aware’ is shown in dark grey, and ‘environmentally ambiguous’ is shown in light grey.

Mapped values showed distinct spatial patterns across all groups (Figure 5). Potential visitation was predicted, at least for low levels, across the entirety of the site for the ‘all visitors’ and ‘aware’ groups, with activity concentrated in the north to north-east areas of the sites, where popular access points are situated, with north–south and east–west areas of higher visitation clearly visible. In the ‘ambiguous’ group, there are distinct areas where no visitation was predicted, with activity once again highest in the north-east of the site, though only clear north–south channels could be identified, rather than the east–west channels also visible in the ‘aware’ group map.

Bootstrapped Pearson correlation was significant for all iterations for all maps at the $p < 0.001$ threshold; ‘aware’ vs. ‘ambiguous’ (correlation mean: 0.893, 95% CI: 0.892, 0.894), ‘all visitors’ vs. ‘aware’ (correlation mean: 0.984, 95% CI: 0.984, 0.984), and ‘all visitors’ vs. ‘ambiguous’ (correlation mean: 0.955, 95% CI: 0.954, 0.955).

4. Discussion

Visitors to the protected area studied here can be segmented into ‘aware’ and ‘ambiguous’ in their on-site environmental awareness from survey data using LCA. Information provided by each group on their use of the site can be upscaled using RF models and mapped as shown in Figure 5. Over two-thirds of the visitors surveyed could be classed as ‘aware’, similarly to the study of Beh and Bruyere [11], where visitors were segmented by their motivation for site visitation, and most were found to have a high awareness of the environment. Our study spatially maps visitation patterns and provides a means to target ‘ambiguous’ groups.

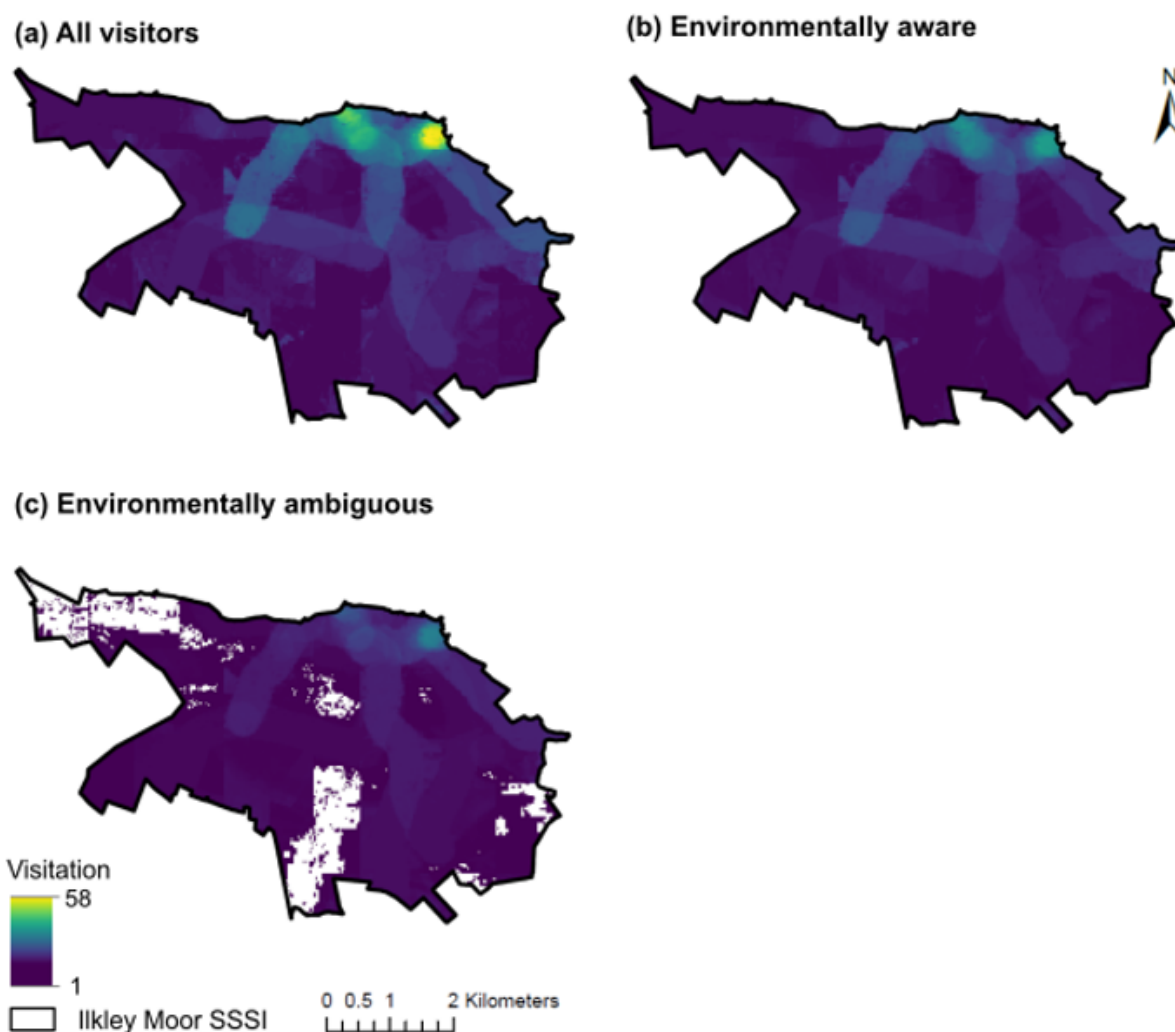


Figure 5. Maps of potential visitation by different groups to Ilkley Moor: (a) all visitors, (b) environmentally aware, and (c) environmentally ambiguous. Zero values predicted by the random forest models are set as white background values for clarity, allowing identification of non-visitation areas for the environmentally ambiguous group. The scale begins at 1 as purple and moves through green to the maximum value of 58 as yellow. The scale is calibrated the same across all three maps, making them comparable.

The LCA showed that awareness levels were higher in the ‘aware’ group, apart from awareness of site designations being similar for both groups. Surprisingly, awareness of site designations had the lowest probability of all the questions in the ‘aware’ group and the highest in the ‘ambiguous’ group (Figure 3). This may infer that being generally ‘aware’ may not necessarily translate to being educated on all environmental aspects of a site; for example, Booth et al. [17] showed a variation of 8–43% in understanding what an SSSI was amongst site visitors.

Variable importance in the random forest models showed that land cover predictors were of lower importance than other factors. The ‘aware’ group had a high importance related to mean temperature and seasonality, which could be attributed to people potentially visiting the site when the weather is favourable. The ‘ambiguous’ group had a relatively low importance for all factors apart from mean temperature and isothermality. Outdoor sensory experiences attract all types of visitors [64], which helps to explain the importance of these factors; hence, when the weather has been constant, visitors could potentially feel more confident in outside pursuits.

A larger proportion of the members of the ‘aware’ group were over 50 years old. Older generations have grown up in the increased presence of nature, and it has been

suggested that they have been sensitised by those experiences they had when they were younger [65], whereas technological advancements may have distanced younger generations from nature [66]. More 'aware' visitors identified as visiting daily. This is supported by Halpenny [15] and Maguire et al. [67], who suggest that a sense of place promotes pro-environmental behaviour and stewardship. We found that within the 'ambiguous' group most visited at a frequency of 2–3/week, but altogether visited more at the lower frequencies. Ballantyne et al. [64] found that infrequent visitors were more likely to be motivated by learning, and thus the more infrequent visitors in the ambiguous group could be more receptive to conservation education delivery. In the aware group, 31.25% travelled less than from one mile away, with 25% in the ambiguous group, yet the highest proportion for both groups was more than five miles away (40% compared to 43.18% within each group respectively). When nature is nearby, it has been suggested that visitation increases, which fosters place dependence [68]. Nevertheless, close proximity can be linked with convenience, being part of an individual's residential environment, which they regularly use [69]. Lower visitation from those nearer the site could have been attributed to the timing of the survey periods in this study, with locals visiting earlier or later in the day. Alcock et al. [70] found that frequency of visitation was lower in individuals living in green-space-abundant areas compared to those living in areas that lack green spaces, who were potentially compensating for the lack of nature, which may also explain the higher number of visitors from further away in this study.

Visitors from the 'ambiguous' group were found to venture less from the northern access points on-site, and then in a linear north–south pattern, as opposed to 'aware' visitors who used these channels and also adopted east–west patterns of movement. This may be due to 'aware' visitors having a larger proportion of daily visitors (indicating familiarity with the site) and included many dog walkers, who tend to spend more time on-site and therefore take longer and more exploratory routes [71]. Bootstrapped Pearson correlation revealed relatively high similarity between the 'aware' and 'ambiguous' potential visitation maps when comparing the entire site (correlation mean: 0.893), which highlights the overlap in visitation across the sites between both groups. Though as can be seen from Figure 5, specific areas, and differences in where visitors from both groups are more likely to visit, can readily be discerned.

By understanding which areas are most visited by different groups, site management can be implemented in association with the results of areas of conservation priority within the site, for example alongside bird survey maps to reduce potentially negative impacts on breeding birds. Kim and Weiler [14] recommend differing on-site communication strategies for visitors with low and high environmental awareness (in relation to fossil collecting). This can be taken further through differentiating strategies for different age cohorts. Personal delivery has been found to be more successful in educating and changing behaviours [23], and hence different events and inclusive educational activities can be targeted for different ages. Habitat management could be focussed on areas with fewer visitors and/or areas with better-informed visitors to attract sensitive species into areas where they are less likely to experience the negative impacts of disturbance. Signage has been shown to work in multiple conservation scenarios [72], though the style of such signs needs to be carefully considered for greatest impact in terms of "behaviour change" outcomes [8,73]. An understanding of where visitors are most concentrated, especially those that may require more conservation education, will allow more targeted education efforts. This could be in the form of increased signage or targeted personal delivery as demonstrated by Kidd et al. [23]. Weaver and Lawton [74] suggest that adhering to a completely biocentric approach that sees visitors as an inherent threat to protected areas can lead to suboptimal sustainability outcomes, whereas seeing visitors as an opportunity rather than a threat for the park can allow visitor mobilisation towards park enhancement, such as pro-environmental activities and citizen science (e.g., bird surveys). Therefore, conservation education combined with these activities could provide a multitude of benefits.

To further assess the robustness of this study, the behaviour of the ‘ambiguous’ visitors would need validation through replication of this study in other protected areas. Other sites may have site-specific characteristics, e.g., coastal sites, woodland, or different biomes, that will need to be assessed for relevant predictor variables, though the same approach elucidated in this study could be used. The variables chosen in this study were limited to bioclimatic and physical, and they could be expanded to include a wider range, both within and outside these categories. Utilisation for machine learning technologies requires specialised skills; thus, the techniques demonstrated in this study are reliant on access to these resources. However, the resources saved in other areas, e.g., wider-scale conservation education, may provide greater savings in the long term. This study used stated responses as part of the survey and participatory GIS exercise. This could be improved using revealed behaviour methodologies, such as voluntary GPS tracking using independent sensors and data loggers or mobile telephones, as seen in Wolf et al. [75]. Alternatively, the accuracy of the participatory mapping could be improved, for example, by studies using online participatory exercise which can capture additional information such as aesthetic appeal of visited areas, as demonstrated by Gosal et al. [76].

5. Conclusions

The methodology elucidated in this study can be readily applied to other areas for which suitable spatial data are available, allowing the development of highly site-specific maps of visitors with differing environmental awareness. Predictors used in this study, e.g., the digital terrain model, calculated viewshed analysis, and bioclimatic variables, are easily found for many regions of the world, albeit at varying resolutions. There is a paramount need for conservation resources to be channelled into the most effective management strategies, and this study demonstrates a method to spatially define those areas that attract visitors with lower environmental and ecological awareness so that on-site resources can be efficiently targeted to where they are needed the most.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/land10060560/s1>, Table S1: Full attribution for bird images in Figure 1 under Creative Commons licenses; Table S2: Birds recorded on Ilkley Moor with Birds of Conservation Concern (BoCC) designations; Table S3: Parameters giving the best out-of-bag (OOB) root mean square error (RMSE) for all RF models using hyper-tuning; Figure S1: Latent class analysis (LCA) AIC and BIC for multiple group (class) sizes; Figure S2: Random forest tree out-of-bag (OOB) stabilisation for random forest models.

Author Contributions: Conceptualization, A.S.G. and G.Z.; methodology, A.S.G., J.A.M., and G.Z.; writing—original draft preparation, review, and editing, A.S.G., J.A.M., K.M.B., C.H.H., and G.Z. All authors have read and agreed to the published version of the manuscript.

Funding: A.S.G. and G.Z. were supported by the Natural Environment Research Council [grant number NE/T002115/1] under the ADVANCES (ADVancing Analysis of Natural Capital in Landscape DecisionS) project. J.A.M. was supported by Q-Step programme (University of Leeds), which is funded by the Nuffield Foundation, the Economic and Social Research Council (ESRC), and the Higher Education Funding Council for England (HEFCE).

Institutional Review Board Statement: The study was approved by the Social Sciences, Environment and Leeds University Business School Faculty Research Ethics Committee of the University of Leeds on 20 May 2019.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Acknowledgments: The authors are grateful for the support of City of Bradford Metropolitan District Council, England, for allowing surveying at Ilkley Moor, Bradford. The authors thank the photographers of the bird species licensed under Creative Commons in Figure 1d (see Table S1) and Kulvinder Kaur for photographs of Ilkley Moor used in Figure 1c. The authors are grateful to the visitors to the study site that participated in the surveys.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Mascia, M.B.; Brosius, J.P.; Dobson, T.A.; Forbes, B.C.; Horowitz, L.; McKean, M.A.; Turner, N.J. Conservation and the social sciences. *Conserv. Biol.* **2003**, *17*, 649–650. [[CrossRef](#)]
- Schultz, P.W. Conservation Means Behavior. *Conserv. Biol.* **2011**, *25*, 1080–1083. [[CrossRef](#)]
- Fisichelli, N.A.; Schuurman, G.W.; Monahan, W.B.; Ziesler, P.S. Protected Area Tourism in a Changing Climate: Will Visitation at US National Parks Warm Up or Overheat? *PLoS ONE* **2015**, *10*, e0128226. [[CrossRef](#)]
- Gill, J.A. Approaches to measuring the effects of human disturbance on birds. *Ibis* **2007**, *149*, 9–14. [[CrossRef](#)]
- Bötsch, Y.; Tablado, Z.; Jenni, L. Experimental evidence of human recreational disturbance effects on bird-territory establishment. *Proc. R. Soc. B Biol. Sci.* **2017**, *284*. [[CrossRef](#)] [[PubMed](#)]
- Alwis, N.S.; Perera, P.; Dayawansa, N.P. Response of tropical avifauna to visitor recreational disturbances: A case study from the Sinharaja World Heritage Forest, Sri Lanka. *Avian Res.* **2016**, *7*, 15. [[CrossRef](#)]
- Remacha, C.; Delgado, J.A.; Bulaic, M.; Pérez-Tris, J. Human Disturbance during Early Life Impairs Nestling Growth in Birds Inhabiting a Nature Recreation Area. *PLoS ONE* **2016**, *11*, e0166748. [[CrossRef](#)]
- Allbrook, D.L.; Quinn, J.L. The effectiveness of regulatory signs in controlling human behaviour and Northern gannet (*Morus bassanus*) disturbance during breeding: An experimental test. *J. Nat. Conserv.* **2020**, *58*, 125915. [[CrossRef](#)] [[PubMed](#)]
- Finney, S.K.; Pearce-Higgins, J.W.; Yalden, D.W. The effect of recreational disturbance on an upland breeding bird, the golden plover *Pluvialis apricaria*. *Biol. Conserv.* **2005**, *121*, 53–63. [[CrossRef](#)]
- Calladine, J.; Critchley, C.N.R.; Baker, D.; Towers, J.; Thiel, A. Conservation management of moorland: A case study of the effectiveness of a combined suite of management prescriptions which aim to enhance breeding bird populations. *Bird Study* **2014**, *61*, 56–72. [[CrossRef](#)]
- Beh, A.; Bruyere, B.L. Segmentation by visitor motivation in three Kenyan national reserves. *Tour. Manag.* **2007**, *28*, 1464–1471. [[CrossRef](#)]
- Fung, C.K.W.; Jim, C.Y. Segmentation by motivation of Hong Kong Global Geopark visitors in relation to sustainable nature-based tourism. *Int. J. Sustain. Dev. World Ecol.* **2015**, *22*, 76–88. [[CrossRef](#)]
- Formica, S.; Uysal, M. Segmentation of travelers based on environmental attitudes. *J. Hosp. Leis. Mark.* **2001**, *9*, 35–49. [[CrossRef](#)]
- Kim, A.K.; Weiler, B. Visitors' attitudes towards responsible fossil collecting behaviour: An environmental attitude-based segmentation approach. *Tour. Manag.* **2013**, *36*, 602–612. [[CrossRef](#)]
- Halpenny, E.A. Pro-environmental behaviours and park visitors: The effect of place attachment. *J. Environ. Psychol.* **2010**, *30*, 409–421. [[CrossRef](#)]
- Bennett, N.J.; Roth, R.; Klain, S.C.; Chan, K.; Christie, P.; Clark, D.A.; Cullman, G.; Curran, D.; Durbin, T.J.; Epstein, G.; et al. Conservation social science: Understanding and integrating human dimensions to improve conservation. *Biol. Conserv.* **2017**, *205*, 93–108. [[CrossRef](#)]
- Booth, J.E.; Gaston, K.J.; Armsworth, P.R. Public understanding of protected area designation. *Biol. Conserv.* **2009**, *142*, 3196–3200. [[CrossRef](#)]
- González Del Campo, A. Mapping environmental sensitivity: A systematic online approach to support environmental assessment and planning. *Environ. Impact Assess. Rev.* **2017**, *66*, 86–98. [[CrossRef](#)]
- Dhami, I.; Deng, J.; Burns, R.C.; Pierskalla, C. Identifying and mapping forest-based ecotourism areas in West Virginia—Incorporating visitors' preferences. *Tour. Manag.* **2014**, *42*, 165–176. [[CrossRef](#)]
- Rieb, J.T.; Chaplin-Kramer, R.; Daily, G.C.; Armsworth, P.R.; Böhning-Gaese, K.; Bonn, A.; Cumming, G.S.; Eigenbrod, F.; Grimm, V.; Jackson, B.M.; et al. When, Where, and How Nature Matters for Ecosystem Services: Challenges for the Next Generation of Ecosystem Service Models. *Bioscience* **2017**, *67*, 820–833. [[CrossRef](#)]
- Hammit, W.E.; Cole, D.N.; Monz, C.A. *Free-Choice Environmental Learning: Framing the Discussion*; Taylor and Francis Ltd.: New York, NY, USA, 2005; Volume 11.
- Salata, T.L.; Ostergren, D.M. Evaluating Forestry Camps with National Standards in Environmental Education: A Case Study of the Junior Forester Academy, Northern Arizona University. *Appl. Environ. Educ. Commun.* **2010**, *9*, 50–57. [[CrossRef](#)]
- Kidd, A.M.; Monz, C.; D'Antonio, A.; Manning, R.E.; Reigner, N.; Goonan, K.A.; Jacobi, C. The effect of minimum impact education on visitor spatial behavior in parks and protected areas: An experimental investigation using GPS-based tracking. *J. Environ. Manag.* **2015**, *162*, 53–62. [[CrossRef](#)] [[PubMed](#)]
- Langston, R.H.W.; Liley, D.; Murison, G.; Woodfield, E.; Clarke, R.T. What effects do walkers and dogs have on the distribution and productivity of breeding European Nightjar *Caprimulgus europaeus*? *Ibis* **2007**, *149*, 27–36. [[CrossRef](#)]
- Liley, D.; Sutherland, W.J. Predicting the population consequences of human disturbance for Ringed Plovers *Charadrius hiaticula*: A game theory approach. *Ibis* **2007**, *149*, 82–94. [[CrossRef](#)]
- Bötsch, Y.; Tablado, Z.; Almasi, B.; Jenni, L. Human recreation decreases antibody titre in bird nestlings: An overlooked transgenerational effect of disturbance. *J. Exp. Biol.* **2020**, *223*. [[CrossRef](#)] [[PubMed](#)]
- Steven, R.; Pickering, C.; Guy Castley, J. A review of the impacts of nature based recreation on birds. *J. Environ. Manag.* **2011**, *92*, 2287–2294. [[CrossRef](#)]
- Mayer-Gross, H.; Crick, H.Q.P.; Greenwood, J.J.D. Bird Study The effect of observers visiting the nests of passerines: An experimental study. *Bird Study* **2010**, *44*, 53–65. [[CrossRef](#)]
- Burger, J.; Leonard, J. Conflict resolution in coastal waters: The case of personal watercraft. *Mar. Policy* **2000**, *24*, 61–67. [[CrossRef](#)]

30. Harris, S.J.; Massimino, D.; Balmer, D.E.; Eaton, M.A.; Noble, D.G.; Pearce-Higgins, J.W.; Woodcock, P.; Gillings, S. *The Breeding Bird Survey 2019*; Thetford: Ann Arbor, MI, USA, 2020.
31. Eaton, M.; Aebischer, N.; Brown, A.; Hearn, R.; Lock, L.; Musgrove, A.; Noble, D.; Stroud, D.; Gregory, R.; Powell, R. Birds of Conservation Concern 4: The Population Status of Birds in the UK, Channel Islands and Isle of Man. *Br. Birds* **2015**, *108*, 708–746.
32. Likert, R. A technique for the measurement of attitudes. *Arch. Psychol.* **1932**, *140*, 44–53.
33. Joshi, A.; Kale, S.; Chandel, S.; Pal, D. Likert Scale: Explored and Explained. *Br. J. Appl. Sci. Technol.* **2015**, *7*, 396–403. [\[CrossRef\]](#)
34. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [\[CrossRef\]](#)
35. Centre for Ecology and Hydrology the UKCEH Land Cover Maps for 2017, 2018 and 2019. Available online: <https://www.ceh.ac.uk/services/lcm2019-lcm2018-and-lcm2017> (accessed on 15 December 2020).
36. Ordnance Survey OS Terrain@50. Available online: <https://osdatahub.os.uk/downloads/open/Terrain50> (accessed on 13 November 2020).
37. QGIS. Development Team QGIS Geographic Information System. Available online: <https://qgis.org/> (accessed on 4 January 2021).
38. Geofabrik OpenStreetMap Data Extracts. Available online: <https://download.geofabrik.de/> (accessed on 15 December 2020).
39. Cuckovic, Z. Advanced viewshed analysis: A Quantum GIS plug-in for the analysis of visual landscapes. *J. Open Source Softw.* **2016**, *1*, 32. [\[CrossRef\]](#)
40. Berk, R.A. An introduction to ensemble methods for data analysis. *Sociol. Methods Res.* **2006**, *34*, 263–295. [\[CrossRef\]](#)
41. Segal, M.R. Machine Learning Benchmarks and Random Forest Regression. UCSF: Center for Bioinformatics and Molecular Biostatistics. 2004. Available online: <https://escholarship.org/uc/item/35x3v9t4> (accessed on 4 January 2021).
42. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **2008**, *28*, 1–26. [\[CrossRef\]](#)
43. Kursa, M.B.; Rudnicki, W.R. Feature selection with the boruta package. *J. Stat. Softw.* **2010**, *36*, 1–13. [\[CrossRef\]](#)
44. Peschel, A.O.; Grebitus, C.; Steiner, B.; Veeman, M. How does consumer knowledge affect environmentally sustainable choices? Evidence from a cross-country latent class analysis of food labels. *Appetite* **2016**, *106*, 78–91. [\[CrossRef\]](#)
45. Rhead, R.; Elliot, M.; Upham, P. Using latent class analysis to produce a typology of environmental concern in the UK. *Soc. Sci. Res.* **2018**, *74*, 210–222. [\[CrossRef\]](#)
46. Ehrlich, O.; Bi, X.; Borisova, T.; Larkin, S. A latent class analysis of public attitudes toward water resources with implications for recreational demand. *Ecosyst. Serv.* **2017**, *28*, 124–132. [\[CrossRef\]](#)
47. Linzer, D.A.; Lewis, J.B. polCA: An R package for polytomous variable latent class analysis. *J. Stat. Softw.* **2011**, *42*, 1–29. [\[CrossRef\]](#)
48. Du, S.; Zhang, F.; Zhang, X. Semantic classification of urban buildings combining VHR image and GIS data: An improved random forest approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *105*, 107–119. [\[CrossRef\]](#)
49. Moon, J.; Kim, Y.; Son, M.; Hwang, E. Hybrid Short-Term Load Forecasting Scheme Using Random Forest and Multilayer Perceptron. *Energies* **2018**, *11*, 3283. [\[CrossRef\]](#)
50. Cheng, L.; Chen, X.; De Vos, J.; Lai, X.; Witlox, F. Applying a random forest method approach to model travel mode choice behavior. *Travel Behav. Soc.* **2019**, *14*, 1–10. [\[CrossRef\]](#)
51. Braun, T.; Cottrell, R.; Dierkes, P. Fostering changes in attitude, knowledge and behavior: Demographic variation in environmental education effects. *Environ. Educ. Res.* **2018**, *24*, 899–920. [\[CrossRef\]](#)
52. Naghibi, S.A.; Pourghasemi, H.R.; Dixon, B. GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. *Environ. Monit. Assess.* **2016**, *188*, 1–27. [\[CrossRef\]](#)
53. Dou, J.; Yunus, A.P.; Tien Bui, D.; Merghadi, A.; Sahana, M.; Zhu, Z.; Chen, C.W.; Khosravi, K.; Yang, Y.; Pham, B.T. Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. *Sci. Total Environ.* **2019**, *662*, 332–346. [\[CrossRef\]](#)
54. Kim, J.-C.; Lee, S.; Jung, H.-S.; Lee, S. Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-Chang, Korea. *Geocarto Int.* **2018**, *33*, 1000–1015. [\[CrossRef\]](#)
55. Park, S.; Kim, J. Landslide Susceptibility Mapping Based on Random Forest and Boosted Regression Tree Models, and a Comparison of Their Performance. *Appl. Sci.* **2019**, *9*, 942. [\[CrossRef\]](#)
56. Heung, B.; Bulmer, C.E.; Schmidt, M.G. Predictive soil parent material mapping at a regional-scale: A Random Forest approach. *Geoderma* **2014**, *214–215*, 141–154. [\[CrossRef\]](#)
57. Mascaro, J.; Asner, G.P.; Knapp, D.E.; Kennedy-Bowdoin, T.; Martin, R.E.; Anderson, C.; Higgins, M.; Chadwick, K.D. A Tale of Two “Forests”: Random Forest Machine Learning Aids Tropical Forest Carbon Mapping. *PLoS ONE* **2014**, *9*, e85993. [\[CrossRef\]](#)
58. Feng, Q.; Liu, J.; Gong, J. UAV Remote sensing for urban vegetation mapping using random forest and texture analysis. *Remote Sens.* **2015**, *7*, 1074–1094. [\[CrossRef\]](#)
59. Hayes, M.M.; Miller, S.N.; Murphy, M.A. High-resolution landcover classification using Random Forest. *Remote Sens. Lett.* **2014**, *5*, 112–121. [\[CrossRef\]](#)
60. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104. [\[CrossRef\]](#)

61. Wright, M.N.; Ziegler, A. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *JSS J. Stat. Softw.* **2017**, *77*. [\[CrossRef\]](#)
62. Gilleland, E. Bootstrap methods for statistical inference. Part i: Comparative forecast verification for continuous variables. *J. Atmos. Ocean. Technol.* **2020**, *37*, 2117–2134. [\[CrossRef\]](#)
63. O'Donnell, M.S.; Ignizio, D.A. Bioclimatic Predictors for Supporting Ecological Applications in the Conterminous United States. *US Geol. Surv. Data Ser.* **2012**, *691*, 4–9.
64. Ballantyne, R.; Packer, J.; Hughes, K. Environmental awareness, interests and motives of botanic gardens visitors: Implications for interpretive practice. *Tour. Manag.* **2008**, *29*, 439–444. [\[CrossRef\]](#)
65. Finger, M. From Knowledge to Action? Exploring the Relationships Between Environmental Experiences, Learning, and Behavior. *J. Soc. Issues* **1994**, *50*, 141–160. [\[CrossRef\]](#)
66. Schultz, P.W. Inclusion with nature: The psychology of human-nature relations. In *Psychology of Sustainable Development*; Springer: Cham, Switzerland, 2002; pp. 61–78.
67. Maguire, G.; Rimmer, J.; Weston, M. Stakeholder Perceptions of Threatened Species and Their Management on Urban Beaches. *Animals* **2013**, *3*, 1002–1020. [\[CrossRef\]](#)
68. Halpenny, E.A. *Environmental Behaviour, Place Attachment and Park Visitation: A Case Study of Visitors to Point Pelee National Park*; University of Waterloo: Waterloo, ON, Canada, 2006.
69. Sterl, P.; Brandenburg, C.; Arnberger, A. Visitors' awareness and assessment of recreational disturbance of wildlife in the Donau-Auen National Park. *J. Nat. Conserv.* **2008**, *16*, 135–145. [\[CrossRef\]](#)
70. Alcock, I.; White, M.P.; Pahl, S.; Duarte-Davidson, R.; Fleming, L.E. Associations between pro-environmental behaviour and neighbourhood nature, nature visit frequency and nature appreciation: Evidence from a nationally representative survey in England. *Environ. Int.* **2020**, *136*, 105441. [\[CrossRef\]](#)
71. Reeves, M.J.; Rafferty, A.P.; Miller, C.E.; Lyon-Callo, S.K. The Impact of Dog Walking on Leisure-Time Physical Activity: Results From a Population-Based Survey of Michigan Adults. *J. Phys. Act. Health* **2011**, *8*, 436–444. [\[CrossRef\]](#)
72. Williams, D.R.; Child, M.F.; Dicks, L.V.; Ockendon, N.; Pople, R.G.; Showler, D.A.; Walsh, J.C.; zu Ermgassen, E.K.H.J.; Sutherland, W.J. Bird Conservation. In *What Works in Conservation 2020*; Sutherland, W.J., Dicks, L.V., Petrovan, S.O., Smith, R.K., Eds.; Open Book Publishers: Cambridge, UK, 2020; pp. 137–281.
73. Ballantyne, R.; Hughes, K. Using front-end and formative evaluation to design and test persuasive bird feeding warning signs. *Tour. Manag.* **2006**, *27*, 235–246. [\[CrossRef\]](#)
74. Weaver, D.B.; Lawton, L.J. A new visitation paradigm for protected areas. *Tour. Manag.* **2017**, *60*, 140–146. [\[CrossRef\]](#)
75. Wolf, I.D.; Hagenloh, G.; Croft, D.B. Visitor monitoring along roads and hiking trails: How to determine usage levels in tourist sites. *Tour. Manag.* **2012**, *33*, 16–28. [\[CrossRef\]](#)
76. Gosal, A.S.; Newton, A.C.; Gillingham, P.K. Comparison of methods for a landscape-scale assessment of the cultural ecosystem services associated with different habitats. *Int. J. Biodivers. Sci. Ecosyst. Serv. Manag.* **2018**, *14*, 91–104. [\[CrossRef\]](#)