



HAL
open science

Understanding the sentiment associated with cultural ecosystem services using images and text from social media

Ilan Havinga, Diego Marcos, Patrick Bogaart, Devis Tuia, Lars Hein

► **To cite this version:**

Ilan Havinga, Diego Marcos, Patrick Bogaart, Devis Tuia, Lars Hein. Understanding the sentiment associated with cultural ecosystem services using images and text from social media. *Ecosystem Services*, 2024, 65, pp.101581. 10.1016/j.ecoser.2023.101581 . hal-04661874

HAL Id: hal-04661874

<https://hal.science/hal-04661874>

Submitted on 25 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Full Length Article

Understanding the sentiment associated with cultural ecosystem services using images and text from social media

Ilan Havinga^{a,*}, Diego Marcos^b, Patrick Bogaart^c, Devis Tuia^d, Lars Hein^a

^a Environmental Systems Analysis Group, Wageningen University, PO Box 47, 6700 AA Wageningen, the Netherlands

^b Inria, University of Montpellier, Montpellier 34090, France

^c National Accounts Department, Statistics Netherlands, Henri Faasdreef 312, 2492 JP The Hague, the Netherlands

^d Environmental Computational Science and Earth Observation Laboratory, Ecole Polytechnique Fédérale de Lausanne, Industrie 17, Sion, Switzerland



ARTICLE INFO

Keywords:

Ecosystem services

Cultural ecosystem services

Natural language processing

Machine learning

Social media

Big data

ABSTRACT

Social media is increasingly being employed to develop Cultural Ecosystem Services (CES) indicators. The image-sharing platform Flickr has been one of the most popular sources of data. Most large-scale studies, however, tend to only use the number of images as a proxy for CES due to the challenges associated with processing large amounts of this data but this does not fully represent the benefit generated by ecosystems in terms of the positive experiences expressed by users in the associated text. To address this gap, we apply several Computer Vision (CV) and natural language processing (NLP) models to link CES estimates for Great Britain based on the content of images to sentiment measures using the accompanying text, and compare our results to a national, geo-referenced survey of recreational well-being in England. We find that the aesthetic quality of the landscape and the presence of particular wildlife results in more positive sentiment. However, we also find that different physical settings correlate with this sentiment and that sentiment is sometimes more strongly related to social activities than many natural factors. Still, we find significant associations between these CES measures, sentiment and survey data. Our findings illustrate that integrating sentiment analysis with CES measurement can capture some of the positive benefits associated with CES using social media. The additional detail provided by these novel techniques can help to develop more meaningful CES indicators for recreational land use management.

1. Introduction

The experience of nature generates a great amount of human well-being (Russell et al., 2013; Chang et al., 2020). People's interactions with wildlife and ecosystems at the landscape level contribute to a number of benefits including better mental and physical health (Sandifer et al., 2015). These contributions can be broadly defined as Cultural Ecosystem Services (CES) which capture the contributions of ecosystems to the non-material benefits arising from human-nature interactions (Chan et al., 2012). For example, human-ecosystem interactions with individual flora and fauna through casual observation can produce a level of cognitive enjoyment related to biodiversity (Keniger et al., 2013) while also generating feelings of aesthetic appreciation at the landscape level (Bratman et al., 2012).

The benefits derived from CES are closely tied to the positive experiences generated by human-nature interactions (Havinga et al., 2020). Such experiences are currently under threat as opportunities to

experience nature and develop positive emotional attachments have decreased with increasing urbanisation and the rise of alternative multisensory experiences such as electronic media (Soga and Gaston, 2016). Termed the "extinction of experience" (Miller, 2005), this decrease in human-nature interactions, and the positive sentiment attached to these experiences, has motivated calls for careful land use planning and management to reconnect people with the natural environment (Abson et al., 2017).

However, land management policies, especially at large scales, are often restricted by the amount of quantitative, spatially-explicit information available (Gould et al., 2020; Edens et al., 2022). As well as employing a number of qualitative methods such as deliberative approaches, CES assessments also seek to provide quantitative information on ecosystem service supply to achieve sustainable policy goals (Raymond et al., 2014). For these purposes, social media has emerged as a promising new source of data due to the scale and detail in which CES measures can be established versus traditional survey methods

* Corresponding author.

E-mail address: ilan.havinga@wur.nl (I. Havinga).

<https://doi.org/10.1016/j.ecoser.2023.101581>

Received 24 September 2022; Received in revised form 2 September 2023; Accepted 26 November 2023

Available online 12 December 2023

2212-0416/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

(Ghermandi et al., 2023). This has led to a number of studies exploring the potential of social media including the image-sharing platform Flickr, which has become one of the most widely utilised due to the accessibility and geographic scope of its data (Richards and Tunçer, 2018; Ghermandi and Sinclair, 2019).

So far, large-scale CES measures based on multimedia data from social media have generally relied on generic indicators, such as the number of Flickr images per site or spatial unit, with an implied uniform value (Wood et al., 2013; van Zanten et al., 2016; Graham and Eigenbrod, 2019). This is because processing such large amounts of data in a systematic way at these spatial scales while still capturing the uniqueness of people's cultural interactions is a challenge (Ghermandi and Sinclair, 2019). However, one human-nature interaction recorded on social media may hold more weight than another in terms of the emotional response it elicits in the person, including observations of wildlife or landscapes recorded through images uploaded to social media platforms (Fox et al., 2021).

Studies conducted at smaller scales, such as in urban and national park areas, have demonstrated the rich variety of experiences reflected in people's posts to social media (Chesnokova et al., 2019; Gosal et al., 2019; Huai and Van de Voorde, 2022). More advanced measures that take account this variation in experience are especially relevant when considering CES as contributing factors to non-material benefits because this relies on people having a positive experience in nature, which may not always be the case, or vary considerably between individuals (Havinga et al., 2020). For example, different physical features, settings or activities contribute to the level of mental or physical enjoyment experienced by an individual and this should also be considered in developing CES measures (Plieninger et al., 2013; Gould et al., 2019).

Here, the Flickr images and text associated with them in the form of titles, tags and descriptions can offer a large amount of relevant information to determine the degree to which CES generate positive experiences (Wartmann et al., 2019). Some recent studies have begun to utilise this information from Flickr to assess people's experiences of nature using sentiment analysis, a form of natural language processing (NLP) used to predict affective states in text (Becken et al., 2017; Brindley et al., 2019; Fox et al., 2021). NLP is the application of computational techniques to the analysis and synthesis of natural language and speech (Eisenstein, 2019). NLP includes a broad range of methods including hand-coded rule sets, dictionary lookups and machine learning techniques including deep learning-based models which use artificial neural networks to extract abstract feature representations from text to assess, for example, its sentiment (Liu et al., 2019). The advantage of these modern NLP models is that these can process large amounts of textual data to generate predictions such as the sentiment reflected in the text associated with social media posts, thereby overcoming many of the data processing challenges which have limited many large-scale studies to simple indicators of CES rather than explicitly including the self-reported experience associated with users' cultural interactions with nature (Gosal et al., 2019).

Nevertheless, despite the availability of these models, the application of sentiment analysis to better understand CES measures based on social media has so far been limited at large scales (Fox et al., 2021), even though text data has shown great potential in studies at smaller scales (Koblet and Purves, 2020). In addition, modern machine learning methods for Computer Vision (CV) have also enabled detailed CES predictions using the image content on Flickr (Egarter Vigl et al., 2021; Lee et al., 2022). CES predictions generated using CV have also relied on deep learning techniques to produce measures including aesthetic quality (Havinga et al., 2021) and observations of wildlife in people's spatial environment (Havinga et al., 2023).

The application of a range of NLP models, including deep learning methods, can in turn enable an assessment of the sentiment associated with these predictions at large scales and, ultimately, the degree to which these are associated with a positive experience of nature. The application of deep learning models to predict scene classes and

attributes in image data can also reveal the different visual factors determining the sentiment associated with CES supply (Cao et al., 2022). In this way, these more advanced methods can be utilised to measure CES and their connections with people's positive experiences in a systematic and cost-effective way versus traditional survey methods (Havinga et al., 2020). However, validating the use of social media and NLP/CV methods for CES assessment using alternative sources, such as surveys, to understand their value, is also important (Englund et al., 2017; Oteros-Rozas et al., 2018).

The availability of large-scale, geo-tagged image/text data coming from platforms such as Flickr, along with the ever more powerful and accessible CV and NLP methods, raises the question of whether these data could be used to develop meaningful CES indicators alongside survey data. Therefore, the objective of our study is to examine the positive experiences associated with CES supply using images and text from social media, along with the required NLP and CV methods, and validate these measures through a comparison with comprehensive survey data. To do this, we utilise Flickr photos and associated text including image descriptions, along with CV and NLP models trained to output visual indicators and sentiment measures. Combined, we aim to understand how these can serve as enhanced measures for CES (Havinga et al., 2021).

We seek to answer the following research questions in our study: (1) What measures for CES-related human-nature interactions can be obtained using text and image data from social media? (2) How do different physical settings activities relate to the sentiment generated by these interactions? (3) Do measures of sentiment and CES match national survey measures of well-being? In doing so, we aim to broadly understand sentiment analysis and social media as an effective tool in measuring CES and, in particular, the positive experiences derived from human-nature interactions including landscape aesthetics and wildlife observations (Langemeyer and Calcagni, 2022).

2. Methods

2.1. Study design

Our study sought to examine the experiences associated with human-nature interactions using social media data and NLP/CV (Fig. 1). We focused on two CES: landscape enjoyment and cultural appreciation of wildlife (see Section 2.2). We conceptualised aesthetic quality as an integral factor in determining landscape enjoyment as a CES and wildlife observations via photography as an indicator of the cultural appreciation of wildlife. CES are, in turn, determined by whether people are having a positive experience and the size of contribution to this experience from nature, in terms of CES supply, can be determined by the setting and type of interaction occurring, alongside other contributing factors such as social context and the effort of the individual themselves (Havinga et al., 2020). We focused on landscape enjoyment and cultural appreciation of wildlife as two key CES through which people value nature (Mace et al., 2012) and are therefore important CES to focus on for policy applications (Daily et al., 2009).

Due to the importance of landscape enjoyment and cultural appreciation of wildlife in people's interactions with nature, key training datasets exist which enable the large-scale analysis which we sought to achieve using CV and NLP methods. For image-based landscape aesthetics estimation we use a model trained on the ScenicOrNot dataset, with landscape images uniformly distributed across Great Britain (ScenicOrNot, 2015; Havinga et al., 2021). This implicitly captures people's landscape aesthetic preferences. For estimating the number of wildlife observations with Flickr images, we applied a model trained on the iNaturalist dataset for species identification on the Flickr dataset (Havinga et al., 2023).

To understand the positive experiences linked to these human-nature interactions, we compared three NLP models tasked with estimating the sentiment expressed in user-generated image text and also looked at the

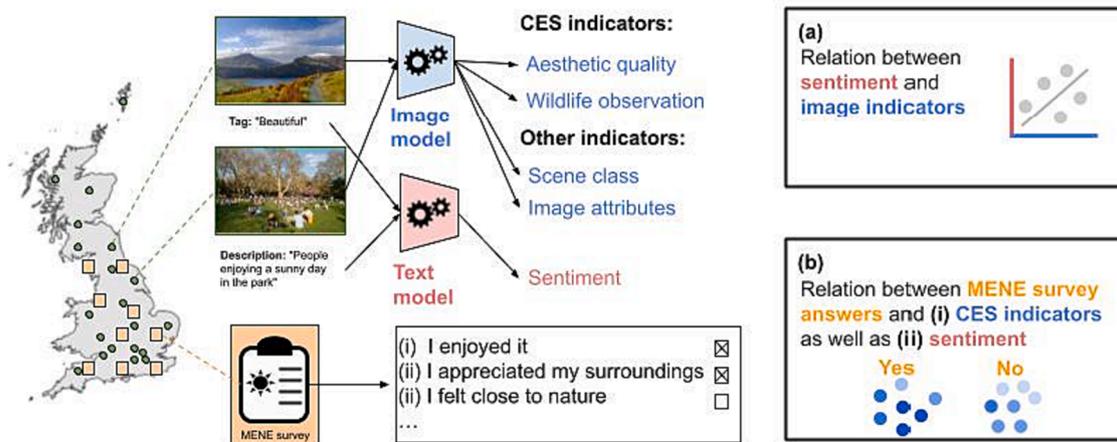


Fig. 1. Study design. (a) CV-based image predictions on Flickr in Great Britain including the aesthetic quality of the landscape and wildlife observations were compared with the sentiment expressed in their accompanying text, measured using text models employing NLP methods. A subset of these predictions were then compared with (b) results of a recreational survey in England, including (i) CES indicators and (ii) sentiment.

adjectives employed by users. We examine the differences in predicted sentiment for aesthetic quality and selected one NLP model based on a correlation analysis with a georeferenced, nationwide survey of nature recreation in England, the Monitor of Engagement with the Natural Environment (MENE), to examine correlations between sentiment and wildlife observations, as well as the sentiment associated with different physical settings and activities. The sentiment associated with physical settings and activities are explored based on CV image classifications generated by the Places365 model (Zhou et al., 2017). Finally, we further validate and compare our Flickr and NLP-based measures of sentiment and CV-based indicators of CES, landscape quality and wildlife observations, with the MENE survey of nature recreation in England.

2.2. Dataset: A collection of outdoor Flickr images in Great Britain

To examine the sentiment associated with landscape aesthetic quality and wildlife observations, we drew on a Flickr image dataset and CES predictions produced in previous research (Havinga et al., 2021; Havinga et al., 2023). This consisted of 9.8 million outdoor images, identified as outdoor using a deep learning model train on Places365 (Zhou et al., 2017). For this research, we also utilised the image scene class and attribute predictions generated by the Places365 model to understand the sentiment associated with different physical settings and activities. A scene class can be defined as the overall semantic description of an image while an image attribute is a specific characteristic within it (e.g. a collection of objects or human activity) (Havinga et al., 2021). To conduct the sentiment analysis in this study, we then also utilised the associated Flickr image metadata in the form of titles and descriptions by accessing the Flickr Application Programming Interface (API) using the ‘flickrapi’ library in Python. Flickr is a highly-curated, image-first social media platform, with a focus on aesthetics (García-Palomares et al., 2015). In contributing photos to the platform, users express what is of value to them, which is an important aspect to CES theory (Havinga et al., 2020). It also provides a wealth of text data in the tags and descriptions associated with user images which, at the scale available, constitutes a unique data source.

2.3. Text data processing and models for sentiment analysis

To conduct part of our analysis, an additional text processing step was needed to better structure the data. This was because the image titles, tags and descriptions generated by users contained a large amount of irrelevant text such as website links, stopwords and duplicated words. For example, without removal, duplicated words recognised by the Hedonometer model would contribute multiple times to the overall

sentiment of a single image. We therefore took a number of text filtering steps to produce a second, processed text dataset to apply the Hedonometer model and look at individual words employed by users through different types of interactions. In contrast, machine-learning based models, such as the Sentiment140 and RoBERTa NLP models, can deal with large, unstructured text datasets and therefore do not require multiple text filtering steps before application (Yan and Liu, 2021).

To filter the dataset, first, we removed all html, numeric characters, punctuation, English stopwords and words less than three characters in length using the ‘tm’ package in R. Second, we removed local toponyms by querying local place names using a 5 km grid overlay and the OS Place Names API.¹ This was because local place names generate a considerable amount of semantic ambiguity. For example, “Lizard” can mean the name of the southernmost peninsula in Britain rather than someone’s observation of this animal species. As a final step, we only used images with at least three words to limit the bias of single or small groups of words on overall sentiment. This ultimately meant working with a dataset of 4.7 million images to conduct our analysis, with the number of images in the unfiltered dataset reduced to the same set of images in the filtered dataset for an equal comparison.

2.3.1. Hedonometer

To estimate the sentiment associated with people’s aesthetic enjoyment and cultural appreciation of wildlife, we applied three different NLP models. The first model, the Hedonometer dataset, was applied using a bag-of-words approach which meant estimating sentiment using the intersection of a dictionary of words with individual sentiment scores. This hedonometer dataset consists of a 10,000 word corpus each with a crowdsourced sentiment rating between 1 and 9 collected using the Amazon Mechanical Turk platform (Alshaabi et al., 2021). The dataset has been successfully applied to Twitter data in the context of national well-being measures in the USA (Mitchell et al., 2013), examining the sentiment of urban green space visitors (Schwartz et al., 2019) and people’s attitudes towards climate change (Cody et al., 2015). The mean sentiment per image was calculated by taking a “bag” of words from the filtered Flickr text dataset and then taking an average of the crowdsourced sentiment assigned to the subset of words that appear in the Hedonometer dataset. We then normalised to a range of –1 to 1 so as to compare with the predictions of the two other models.

¹ <https://www.ordnancesurvey.co.uk/business-government/products/names-api>.

2.3.2. Sentiment140

The second model applied we applied was the Sentiment140 model which uses the machine learning classifier Maximum Entropy (MaxEnt) to generate predictions (Go et al., 2009). MaxEnt models are feature-based models which calculate a conditional probability across a set of classes using optimised feature weights generated after training. The model was trained on a dataset of Tweets with emoticons as training features and a model output of -1 for negative, 0 for neutral, and 1 for positive sentiment. It has been applied in the context of public health monitoring to infer Twitter users' health status over time (Kashyap and Nahapetian, 2014) and to test the relationship between sentiments expressed on Twitter and socio-demographic indicators (Ostermann, 2021). In our study, we applied the model to the image descriptions of the unprocessed Flickr text dataset rather than the filtered dataset in the case of the Hedonometer model, leveraging the Sentiment140 model's ability to retrieve meanings from whole sentences through its Tweet-based training dataset from Twitter. We applied the model using the Sentiment140 API,² before normalising the resulting scores to be between -1 and 1 .

2.3.3. RoBERTa

Finally, the third model we applied to generate sentiment predictions was the RoBERTa model. RoBERTa is a deep learning-based model which uses artificial neural networks to generate feature representations of words and then uses these to produce sentiment estimates (Liu et al., 2019). It has produced state-of-the-art results based on a two-step training scheme, drawing on both large amounts of unlabelled training data and task-specific, labelled data. It has previously been used to estimate the effects of the Covid-19 pandemic on student sentiment using data from the social media site Reddit (Yan and Liu, 2021) and the sentiment associated with common points of interests such as parks in cities globally (Stelzmüller et al., 2021). We used a sentiment model based on RoBERTa which was trained on 124 million tweets from January 2018 to December 2021,³ and finetuned for sentiment analysis using a crowdsourced, annotated training dataset of tweets (Barbieri et al., 2020; Loureiro et al., 2022). The model produced a softmax score across a negative, neutral and positive class. We used the Python libraries 'Transformers' to download the model and apply it to our dataset, taking the positive score and subtracting the negative score before normalising the model output to a $-1,1$ range. As with the Sentiment140 model, we applied the model to only the descriptions of the unprocessed Flickr text dataset as opposed to the filtered dataset to preserve whole sentences, similar to the format of Tweets on Twitter.

2.3.4. Adjective analysis

To gain additional insight into the sentiments expressed by users in the text associated with their images, we also looked at the most common and unique adjectives used across aesthetic quality ratings and for different species. Adjectives are the most emotive elements of language and can therefore provide one of the clearest indications of the affective state of people (Bush, 1973). To do this, we looked at how users employed positive and negative adjectives across aesthetic quality ratings. We looked at a selection of commonly used words that are typically associated with positive or negative sentiment, choosing "beautiful", "enjoy", "like", "calm", "happy", "love" as well as "afraid", "awful", "sad", "angry", "hate" and "ugly". In addition, we also examined the most uniquely employed adjectives for the species orders spiders (*Araneae*), songbirds (*Passeriformes*), butterflies/moths (*Hymenoptera*) and the *Asparagales* plant order, which includes the daffodil flowering plant genus (*Narcissus*).

To identify the most uniquely employed adjectives for the species orders, the term frequency-inverse document frequency (tf-idf) of

adjectives used by Flickr users was calculated for each species order in comparison with each other. Tf-idf is a metric that highlights the importance of a word in a document by comparing how often it occurs in the document (in our case, in a subset of images) compared to its general abundance in the whole corpus of documents (all images). Employing tf-idf meant we could analyse the frequency of an adjective used to describe a particular species in proportion to the frequency of its use for other species. This was done using the 'udpipe' package in R to identify adjectives and the 'tidytext' package to calculate the tf-idf. To gain a more balanced view of the adjectives employed by all users, we took a 10 % sample of images per user and only included adjectives if they occurred at least twice for each species order.

2.4. Image analysis for CES predictions

The Flickr dataset contained both CES measures of landscape enjoyment and cultural appreciation of wildlife. For landscape enjoyment, this consisted of an aesthetic quality rating between 1 and 10 per image. These predictions were generated using a Convolutional Neural Network (CNN) trained on a crowdsourced image dataset of the British landscape, the ScenicOrNot database. The database consists of over 1.5 million ratings of over 200,000 images, with multiple ratings per image (ScenicOrNot, 2015). The model was trained on the average rating per image and the results compared to a measure of aesthetic quality using environmental indicators to confirm the reliability of the spatial predictions (Havinga et al., 2021).

Cultural appreciation of wildlife was captured as a binary prediction representing a generic or wildlife observation. This prediction was also generated using a CNN model, trained to distinguish between wildlife observations and generic interactions with the environment using a dataset of Flickr and iNaturalist images, a citizen science platform (Havinga et al., 2023). We took only the most confident predictions and structured our analysis around the five most frequently captured taxonomic groups: plants, birds, insects, mammals and arachnids. This enabled us to gain large-scale insights into people's cultural interactions with nature. We then looked at the sentiment associated with the most common species (0.75 quantile, with a minimum of 100 interactions) by using the captions associated to their images.

2.5. Sentiment associated with physical settings and activities

To examine the sentiment associated with physical settings and human activities, we utilised the image scene and attribute predictions of the Places365 model. In doing so we sought to understand the physical and human-related factors influencing people's positive experiences such as landscape features, including the built and natural environment, or common activities such as hiking. We used 35 of the most relevant attribute classes out of a total of 102 classes and 30 scene classes out of a total of 365 classes (see Supplementary Tables S1 and S2). Attribute classes drew on the SUN attribute database, a crowdsourced dataset of common image attributes (Patterson et al., 2014). For each image within our Flickr dataset, we took the most confident scene class and image attribute prediction. We then calculated the mean sentiment and the total number of images for each scene class and image attribute.

2.6. MENE survey comparison

To validate our results, we compared our social media and NLP/CV model results with nature trips and well-being measures reported in a national recreation survey of England, setting aside any images in Scotland or Wales. The MENE survey is a demographically representative survey of England conducted throughout the year in which respondents are asked about their trips to nature in the past week. Respondents are then asked more information about one of these trips, such as the location and motivation for their visit, with some then asked

² <https://help.sentiment140.com/api>.

³ <https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest>.

to report measures of well-being as a result of the trip. For our study, we used the survey results collected over a ten year period between 2009 and 2019 (Natural England, 2019).

In the survey, people are given a number of options to choose from to state the motivation for their trip. The options most relevant to our study were the statements “to enjoy the scenery” and “to enjoy the wildlife” to which people could answer either “Yes” or “No”. As a first point of comparison, we compared the mean sentiment, aesthetic quality and wildlife observations captured in Flickr images within a 1 km radius of these visits using the trip coordinates. This subset consisted of 82,950 respondents. This produced two sets of data, one associated with people’s “Yes” responses and the other with their “No” responses. As we were comparing both continuous data and binary data, we checked the significance of these differences using two sample tests chosen after checking for normal distributions using the Shapiro-Wilk test. These were the Welch Two Sample *t*-test as a parametric test and the Wilcoxon rank sum test as a non-parametric test.

A further subset of respondents (11,667 people) were asked to report feelings of well-being associated with their trip. These respondents were asked to what degree they agreed with the statements “I enjoyed it”, “it made me feel calm and relaxed”, “it made me feel refreshed and revitalised”, “I took time to appreciate my surroundings”, “I learnt something new about the natural world” and “I felt close to nature”. We took the responses to these statements as a second point of comparison, also calculating the mean sentiment, aesthetic quality and wildlife observations in Flickr images within a 1 km radius and then comparing the level of agreement with these Flickr-based measures.

3. Results

3.1. Sentiments associated with CES

3.1.1. Aesthetic quality

Fig. 2 shows the relationship between sentiment and aesthetic quality using the predictions of the three NLP models as well as the relative word frequencies of a set of positive and negative adjectives. All three models resulted in significant ($p < .001$) trends in the relationship between aesthetic quality and sentiment using a linear regression analysis, with a clear upward overall trend. The application of all three models showed a similar pattern with the least positive sentiment for the lowest aesthetic quality ratings, a small levelling-off between the 5 and 7 ratings, before increasing again with the highest ratings. A significant linear relationship ($p < .0001$) was also found between all three models. This increase was the most prominent for the RoBERTa sentiment predictions, followed by the Sentiment140 and Hedonometer models, with the Hedonometer predictions barely increasing from the average ratings. The sentiment140 model saw the largest range of uncertainty, indicating a wider range of positive/neutral predictions for each rating. All models were found to have a significant relationship with sentiment based on a linear regression analysis ($p < .001$). Ultimately, although the RoBERTa model showed the strongest relationship with higher aesthetic ratings, the Hedonometer model showed the best overall relationship (Supplementary Table S4), (Pearson’s $R = 0.142$) versus Sentiment140 (Pearson’s $R = 0.053$) and RoBERTa (Pearson’s $R = 0.09$), as well as the highest correlation with the MENE survey responses (Supplementary Table S3).

In terms of the adjectives employed by users across the aesthetic ratings, more positive adjectives and less negative adjectives were used per image relative to the total number of images per rating as aesthetic quality increased (Fig. 1b). The positive adjectives “beautiful”, “calm” and “like” showed the clearest relationship with aesthetic quality, showing substantial increases with higher ratings, even though “like” can also be used as an adverb with no clear positive association. For example, many landscape images were associated with the word “beautiful”. The words “happy”, “enjoy” and “love” showed less of a clear relationship, with a slight increase across the average ratings associated

with activities such as picnicking in the park. Conversely, the use of the word “hate” fell steeply from low to high aesthetic quality. The use of the word “hate” was associated with urban settings such as the road network being described in Fig. 1b (iii). Generally, the words “sad” and “ugly” also fell with increasing aesthetic quality while the words “afraid”, “awful” and “angry” showed less of a clear relationship, sometimes being used to describe the behaviour of animals in images. Notably, the number of positive adjectives employed by users was far greater than negative adjectives. We chose the positive and negative words a priori (i. e. not based on their abundance in the dataset by in the English language), but note that a different set of words could have been chosen for this analysis.

3.1.2. Wildlife observations

We found that, overall, the Hedonometer model produced the strongest correlations between the estimated sentiment and reported well-being statements in the MENE survey with the majority significant ($p < 0.001$) (Supplementary Table S3). We therefore used its predictions to look at the relationship between different species interactions and sentiment, shown in Fig. 3, although, overall, correlations for all sentiment predictions were low, ranging between 0.01 and 0.06. There was also a significant relationship ($p < 0.001$) between the number of species observations predicted by the CV model in the Flickr images and the statement “I felt close to nature” with a 0.06 correlation (Supplementary Table S7).

Plants produced the most positive sentiment with plant orders such as *Asparagales*, which includes daffodils, seeing some of the most positive sentiment. Other flowering orders such as *Ranunculales* which includes the buttercup family, also saw very positive sentiment versus others, such as *Caryophyllales*, which includes cacti, and *Poales*, which includes grasses, with less positive sentiment. After plants, birds and specifically butterflies and moths as insects saw the next greatest level of sentiment associated with user observations. In terms of birds, this included cranes, waterfowl and songbirds, while pelicans and herons produced the lowest levels of sentiment. Mammals also saw moderately positive sentiment in comparison to the other species while arachnids, including spiders and scorpions, saw the least positive sentiment expressed towards them out of all species.

Considering the difference in the words used by Flickr users for a selection of these species, the tf-idf scores in Fig. 2 shows the uniqueness of the adjectives employed by users in their observations of spiders, songbirds, butterflies/moths and the *Asparagales* plant order. Spiders saw very little unique words with only five words gaining a non-zero tf-idf score, the strongest of which were “international” and “globular”. In contrast, songbirds, butterflies/moths and the *Asparagales* order saw a much larger and more positive range of adjectives more uniquely employed by users versus the other species. For example, in the case of butterflies/moths, the adjectives “magnificent”, “perfect” and “beautiful” gained relatively high tf-idf scores. “beautiful” was also often used to describe *Asparagales* versus the other species orders. In the case of songbirds, adjectives such as “young” and “friendly” were more often used in comparison to the other species.

3.2. Sentiment associated with physical settings and activities

Fig. 4 shows the sentiment and aesthetic quality related to different image attributes and scenes based on the prediction of the Hedonometer model. Again, we used the Hedonometer model for this as it showed the strongest correlation with the responses in the MENE survey. Overall, the correlations were small but significant. For example, the statement “I enjoyed it” saw a Pearson’s R of 0.05 with $p < 0.001$ (Supplementary Table S3). The sentiment and aesthetic quality linked to different attributes varied greatly. The attributes “flowers” and “socializing” were linked to the most positive sentiment while “railroad” and “brick” experienced the least positive sentiment. Most of the urban or man-made attributes, including other attributes such as “pavement” and “asphalt”,

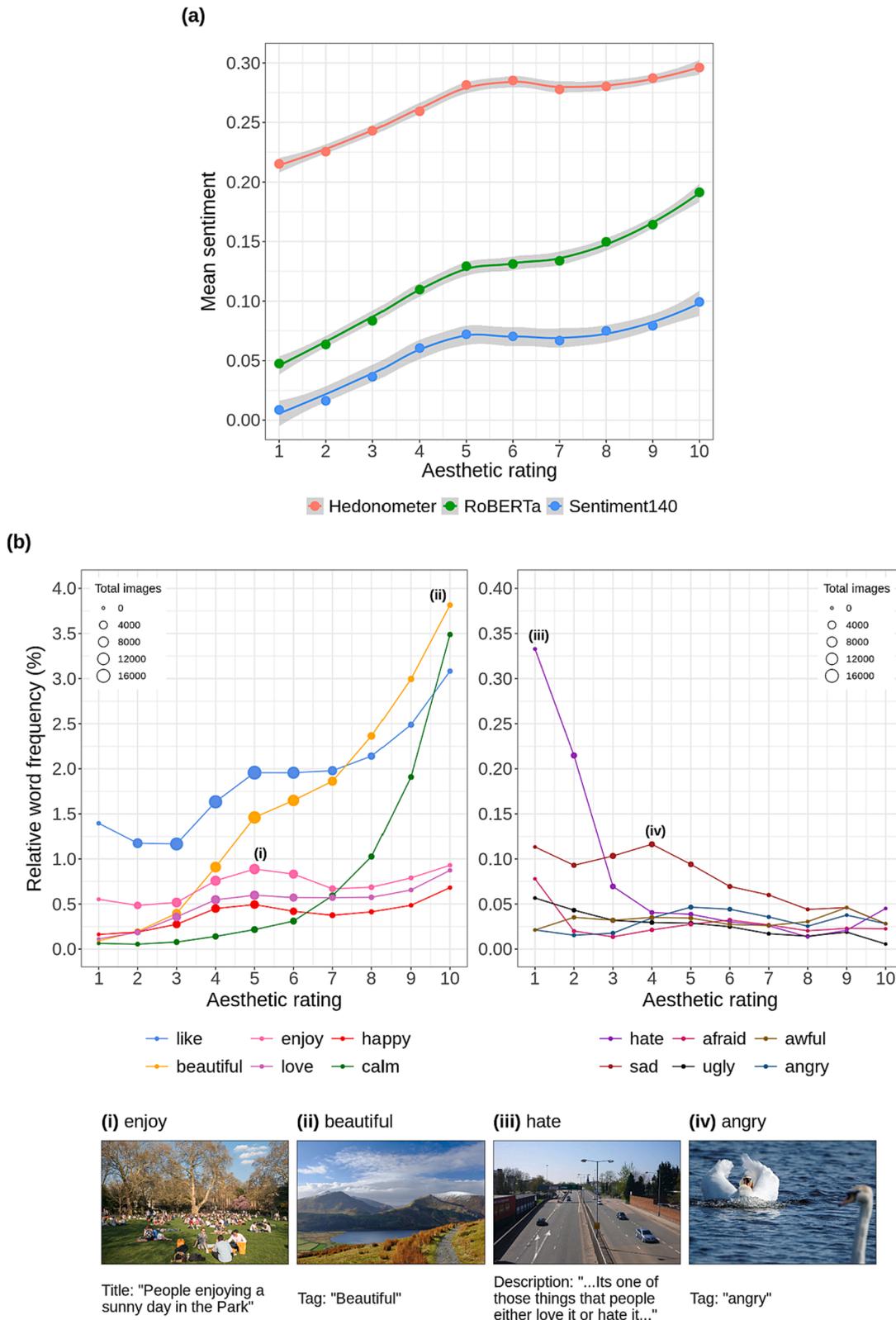


Fig. 2. The relationship between Flickr user sentiment and aesthetic quality. In (a) the sentiment predictions of the three NLP models are shown versus aesthetic quality. Images were grouped into aesthetic quality rating bins (e.g. 1 represents images with a rating between 0 and 1) and the mean sentiment calculated per rating bin with a loess smoothing filter showing the 95 % confidence intervals. Based on a linear regression analysis, all three models were found to produce a significant ($p < .001$) relationship between sentiment and aesthetic quality. Significant agreement between all three models was also found ($p < .0001$). In (b) the relative word frequencies are shown of six positive and six negative adjectives per rating, in addition to four image examples for (i) “enjoy”, (ii) “beautiful”, (iii) “hate” and (iv) “angry”. Photos © Marco Verch, Scott Wylie, Lydia and Stephen Gidley (cc-by/2.0).

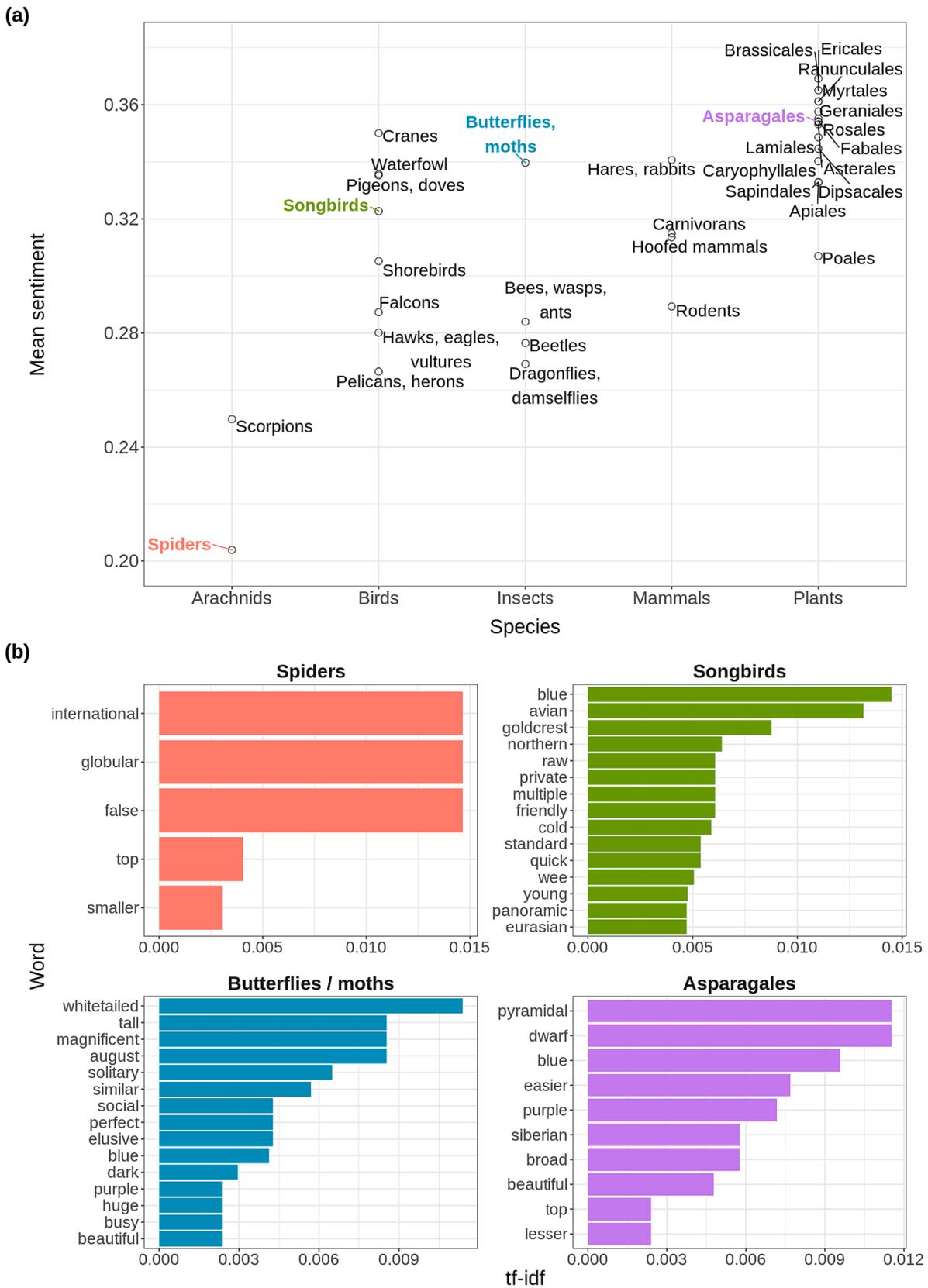


Fig. 3. The sentiment and adjectives associated with species observations using the predictions of the hedonometer NLP model. In (a) mean sentiment for the most common species order observations are shown while (b) shows the adjectives with the highest tf-idf scores employed by Flickr users to describe their observations with spiders, songbirds, butterflies/moths and the asparagales order. Some additional place and camera-related words such as “nationaltrust” and “macro” have been removed.

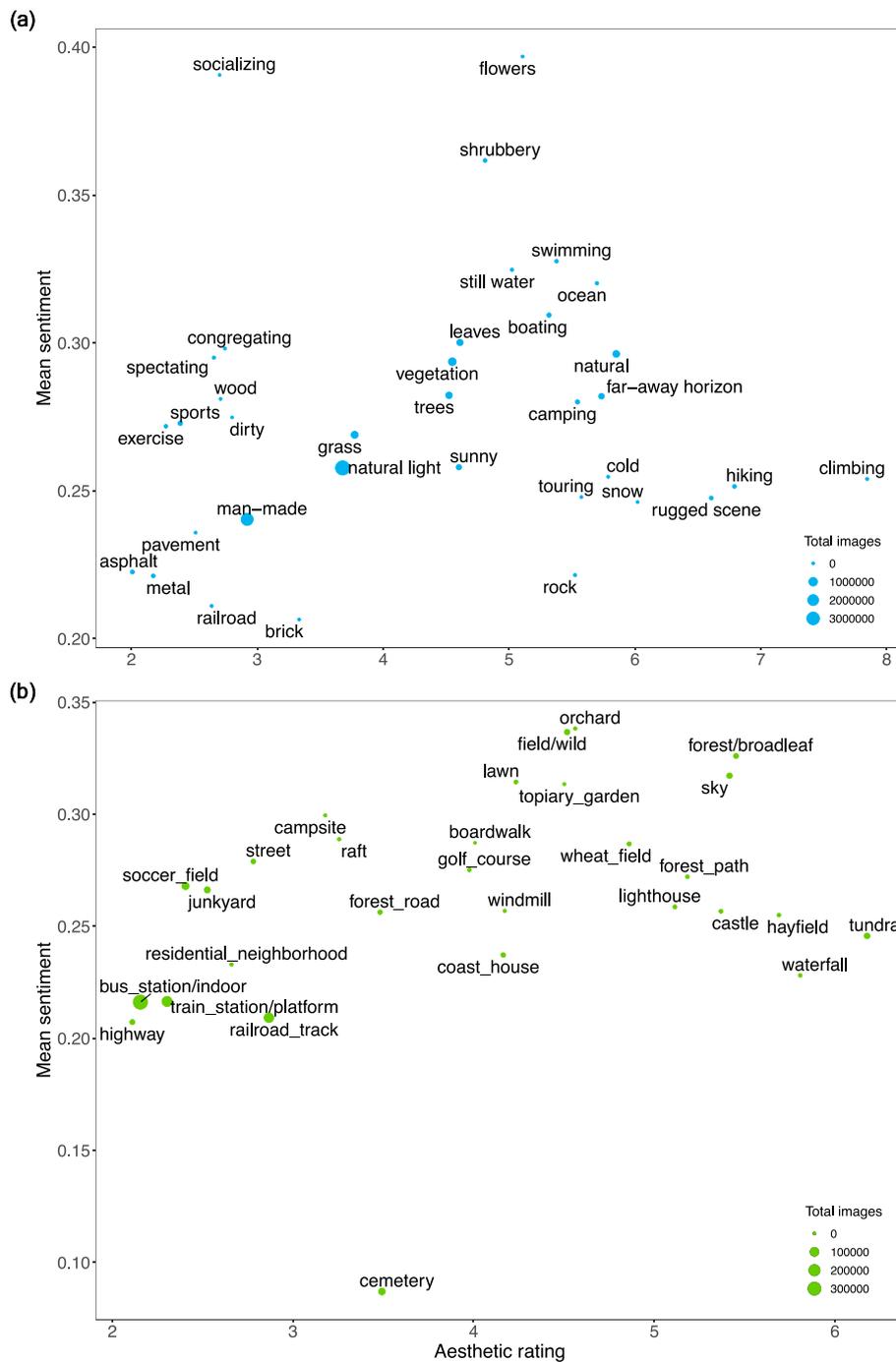


Fig. 4. Sentiment versus aesthetic quality for a selection of (a) image attributes and (b) scene classes generated by the Places365 model representing different physical settings and activities. The size of points for attributes indicates the number of images with a score > 0.6 and for scenes images with a softmax score > 0.5.

saw both the least positive sentiment and low aesthetic quality. Also, notably, other social activity-related attributes such as “congregating” and “spectating” were linked to very positive sentiment but low aesthetic quality. In contrast, “hiking” and “climbing” were associated with high aesthetic quality but less positive sentiment. This was also true of nature-related attributes such as “ocean”, “natural” and “snow”, although water-related activities and features elicited a fair amount of positive sentiment.

Although the overall relationship between sentiment and aesthetic quality again varied greatly for scene classes, a similar pattern of both the least positive sentiment and lowest aesthetic quality associated with man-made and urban-related settings was also detected (Fig. 3). For example, the scenes “train station/platform”, “highway” and “bus

station/indoor”. In particular, “cemetery” was associated with the least positive sentiment. However, it was still associated with a moderate level of aesthetic quality. “orchard” and “field/wild” were linked to positive sentiment, suggesting a link with nature, while “golf course” and “campsite” were also associated with very positive sentiment versus moderate aesthetic quality, an indication that social activities are also important to people’s positive experiences. Scenes linked to natural settings, such as “tundra”, “waterfall” and “hayfield” were linked to very high aesthetic quality but moderate sentiment.

3.3. Comparison with surveyed well-being measures

Table 1 shows Flickr-based measures of CES and sentiment, using the

Table 1

Comparison of Flickr-based sentiment and CES measures within 1 km of trips taken to enjoy the scenery or wildlife as reported in the MENE survey.

Trip Motivation	Response	Flickr-based CES measure	Mean CES	Mean sentiment
To enjoy scenery	Yes	Aesthetic quality	4.29 (± 0.03) ($p < .0001$) [†]	0.14 (± 0.003) ($p < .0001$) [‡]
	No		3.19 (± 0.02) ($p < .0001$) [†]	0.13 (± 0.002) ($p < .0001$) [‡]
To enjoy wildlife	Yes	Wildlife observations	19.3 (± 5.46) ($p = .001$) [‡]	0.145 (± 0.004) ($p < .0001$) [‡]
	No		10.9 (± 0.780) ($p = .001$) [‡]	0.135 (± 0.002) ($p < .0001$) [‡]

[†] based on Welch Two Sample *t*-test, [‡] based on Wilcoxon rank sum test.

mean prediction of the hedonometer model, compared with trips reported in the MENE survey. A clear relationship was observed between the locations visited by survey respondents and the CES measures estimated using Flickr images. Trip locations visited to enjoy the scenery saw a mean aesthetic quality of 4.29 in Flickr images within 1 km versus 3.19 in images near trip locations not taken for this reason ($p < .0001$). The sentiment associated with Flickr images near trip locations undertaken for enjoying the scenery was also significantly higher ($p < .0001$). There was also a significant difference between trips taken to enjoy wildlife and the mean number of wildlife observations in Flickr images, with 19.3 species observations for wildlife-related trips versus 10.9 for trips not taken for this reason ($p = .001$). Again, sentiment was found to be significantly higher in Flickr images in the same areas as wildlife-related trips reported in the MENE survey ($p < .0001$).

Similarities were also found between Flickr-based CES measures and sentiment and the self-reported well-being statements in the MENE survey (Fig. 5). Significant differences were especially visible for the level of agreement with the statement “I took time to appreciate my surroundings” and aesthetic quality. The statements “it made me feel calm and relaxed” and “I felt close to nature” also showed significant differences between the levels of agreement and aesthetic quality. Further analysis looking at the correlations between aesthetic quality and the levels of agreement with each statement confirmed this (Supplementary Table S6). The levels of agreement with the different statements and number of wildlife observations generally showed large standard errors which affected the significance of some of the differences. However, the statement “I felt close to nature” produced significant differences in the number of wildlife observations associated with “strongly agree” and almost all other levels of agreement. Correlation analysis again confirmed this relationship between the statement and people’s level of agreement (Supplementary Table S7).

In terms of sentiment, significant differences were also observed. The most notable were associated with the statement “I enjoyed it” which saw larger sentiment associated with stronger levels of agreement. Less notable but still significant differences were also observed between levels of agreement for the statements “I took time to appreciate my surroundings” and “I felt close to nature”. Levels of agreement with other statements were generally not associated with significant differences. In particular, “strongly agree” was affected by large standard errors. This was because there were only a small number of people (<60) that gave this response in the survey. Further correlations between sentiment, CES measures and the MENE survey results can be found in Supplementary Tables S3, S4 and S5).

4. Discussion

The use of social media to measure CES has seen rapid growth in recent years (Ghermandi and Sinclair, 2019). Nevertheless, so far, its potential has not been fully realised, and key challenges remain regarding the representativeness of the data and its use in producing

policy-relevant indicators (Oteros-Rozas et al., 2018; Havinga et al., 2020). In our study, we have demonstrated that, through the application of NLP methods and machine learning, connections can be made between social media-based CES measures and people’s positive experiences of nature, a key consideration for CES indicator development in the context of public health (Sandifer et al., 2015). At the same time, our findings highlight a number of physical settings and activities that influence people’s positive experiences of nature.

Overall, there was a clear relationship between aesthetic quality and sentiment, with a clear uptrend towards higher aesthetic ratings. This finding is consistent with the health-promoting impacts of outdoor environments outlined in the landscape aesthetics literature (Abraham et al., 2010). For example, landscapes of high aesthetic quality, in terms of their perceived naturalness and diversity, evoke positive emotions in people, resulting in improvements to their general mood and mental health (Bieling et al., 2014; Seresinhe et al., 2019). Notably, there was a levelling-off in sentiment for the images within the average aesthetic rating bins. This may highlight the greater variability in people’s responses to landscapes of ordinary quality and the more consistent responses elicited by either very ugly or attractive environments (Workman et al., 2017). This effect was observable in the relative word frequencies of adjectives employed by users in our study, such as “beautiful” and “calm” versus “ugly” and “sad”, which showed strong relationships with high and low aesthetic quality. This effect was clear in spite of the different ways in which these words may be used in the context of an image caption (for instance polysemia, sarcasm, or the fact that an adjective could refer to the photographer or to a subject in the photo).

In terms of wildlife-related sentiment, some notable differences between species groups were also observed, with plants, especially flowering plants, generating the most positive sentiment, followed by insects and birds. These results are in line with those of survey-based studies in Great Britain (Aerts et al., 2018). For example, in public green spaces around England, higher levels of flower cover were associated with larger restorative effects (Hoyle et al., 2017). Similarly, in a survey of public attitudes towards biodiversity attributes, butterflies were found to be one of the most valued insect species (Austen et al., 2021), confirming our results. In another study of urban households, songbirds were especially appreciated by people out of all garden birds (Cox and Gaston, 2015). It has been argued that more positive overall reactions to plants versus birds or insects may be due to their static and thus more visible presence (Fuller et al., 2007; Dallimer et al., 2012). Some insects can also appear as a threat to people, evoking less positive emotions (Austen et al., 2021). This is also true of arachnids such as spiders which saw the lowest levels of sentiment expressed towards them in our study (Zvaríková et al., 2021).

Previous studies employing survey methods have also shown the importance of particular physical settings and activities on the positive experiences people gain from their interactions with nature. For example, social cohesion has been found to be one of the strongest mediators between green space and health (de Vries et al., 2013; Wolf and Wohlfart, 2014). This was also reflected in the results of our study, with the image attribute and scene classes “socializing” and “campsite” linked to some of the most positive sentiment, and the word “enjoy” being associated with picnicking in an urban park. Similarly, the very positive sentiment associated with the attributes “flowers” and “shrubbery”, as well as the more colourful plant species orders, reflect the significant positive effects of colour diversity and green planting (Carrus et al., 2015; Hoyle et al., 2018). Conversely, the sentiment linked to “tundra”, “hiking” and “climbing” versus aesthetic quality suggests very appealing but less comfortable environments may elicit more reflective responses rather than explicitly positive ones (Baklien et al., 2016; Stevenson and Farrell, 2018).

The comparison with the MENE survey results showed some significant relationships between CES measures, sentiment and respondents’ trip motivations and reported well-being. Although the correlations

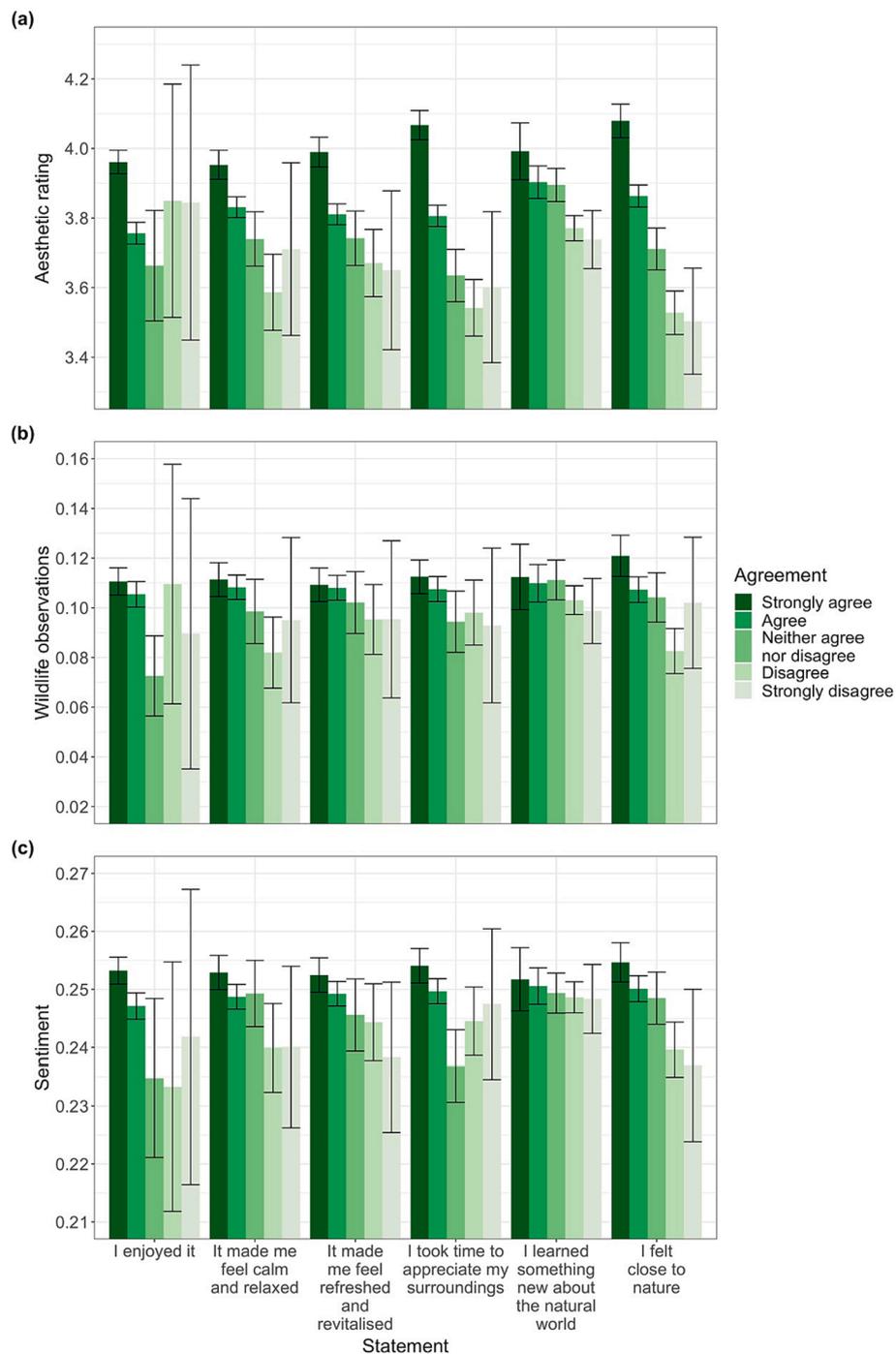


Fig. 5. Comparison of satisfaction expressed in the MENE survey data and the mean (a) aesthetic quality, (b) number of wildlife observations and (c) sentiment expressed in Flickr images within 1 km distance of the geo-located visits.

found were relatively low, their high significance level is noteworthy, given the noisy nature of Flickr data and very different nature of the two data sources. This provides good evidence that Flickr-based measures of CES are consistent with a large sample of self-reported visits across a large geographic and seasonal range, evidence that is generally missing in the CES literature (Oteros-Rozas et al., 2018). The positive experiences recalled by respondents in the survey such as “I enjoyed it” were also consistent with the sentiment expressed by Flickr users, suggesting a common level at which these CES contribute to non-material benefits such as better mental health. Social media sentiment has been linked to national indicators of well-being (Kokil et al., 2020) while higher levels of vegetation cover and bird abundances have been found to be

positively associated with a lower prevalence of depression, anxiety, and stress (Cox et al., 2017; Methorst et al., 2021). Social media-based measures of CES can therefore integrate these casual pathways to better study the effect of nature on people below the ecosystem scale (Botzat et al., 2016).

However, not all statements of agreement in the MENE survey matched sentiment levels on Flickr. This may again be because statements such as “I took time to appreciate my surroundings” are linked to more thoughtful experiences, rather than expressly positive ones. The number of people disagreeing or strongly disagreeing was also very low, reflecting a general human positivity bias in relation to people’s voluntary trips to experience nature (White et al., 2013). Nevertheless,

the significant differences between the stronger levels of agreement and the Flickr-based sentiment measures are encouraging evidence that these novel data and techniques capture representative information on people's interactions with nature. This confirms the results of some more recent studies which have started looking at validating social media-based indicators of CES with survey data (Johnson et al., 2019; Moreno-Llorca et al., 2020).

The more reflective nature of some people's positive experiences during their outdoor recreation also highlights some of the limitations of the approach taken in our study. Not all positive experiences related to nature may be expressed on social media while the full range of positive experiences that are expressed in users' text may not be fully captured through the application of NLP models. For example, we found the image scene class "cemetery" to be associated with the least positive sentiment. However, cemeteries are associated with significant restorative effects (Quinton and Duinker, 2018), produced through feelings of peacefulness and contemplation (Nordh et al., 2017). Such positive but more ambiguous emotions are difficult to detect through sentiment analysis and require additional contextual information to generate accurate predictions (Poria et al., 2020). At the same time, Flickr users may be less eager to share more contemplative experiences such as these due to the heavy positivity bias found on social media platforms (Waterloo et al., 2017). Moreover, the CV and NLP approaches we employed in this paper rely on the existence of appropriate training datasets to detect specific CES. Interactions related to spirituality, such as those experienced in settings such as cemeteries, would be better captured if equivalent, large-scale crowdsourced training datasets were available. Nevertheless, the interconnectedness of people's experiences with ecosystem aesthetics and wildlife as two key CES may already have captured some of these broader positive experiences (Cooper et al., 2016).

The manner in which people use Flickr may also have resulted in the Hedonometer model achieving the highest correlation with the MENE survey, even versus the more advanced RoBERTa model. This is because RoBERTa has been trained to predict sentiments using Twitter data, a platform on which users express much stronger opinions versus Flickr (Samani et al., 2018). Therefore, the way in which Flickr users refer to the natural environment around them, such as a particular flower species, may be more neutral than the model is trained to detect. In contrast, the positive benefit gained from the wildlife observation is better identified by the Hedonometer model just based on the inclusion of these words and their crowdsourced sentiment rating. The application of the Hedonometer model has previously been demonstrated versus other national well-being survey data with the added advantage of a high level of interpretability (Mitchell et al., 2013; Loff et al., 2015). In contrast, understanding the factors influencing the predictions of machine-learning based NLP models is a challenge and these types of models have often been criticised for being "black boxes" (Barredo Arrieta et al., 2020). For example, it is not always clear how RoBERTa is using the words in a sentence, such as verbs or adjectives, to generate a prediction (Clark et al., 2019).

Nevertheless, our findings illustrate the relevancy of these novel data and techniques in the development of CES indicators in recreational land use management. From a CES perspective, it is important to move beyond the use of simple proxies such as the number of images to better incorporate people's positive experiences (Richards and Tunçer, 2018; Zhang et al., 2020). In a broader context, public health measures also typically miss specific but important details of nature exposure such as the quality of the nature (Hartig et al., 2014). In our study, we were able to relate different levels of aesthetic quality to different levels of sentiment using social media and CV/NLP methods. This can support CES measures that incorporate both measures, and those that therefore move beyond simple image counts (Fox et al., 2021). However, the configuration of these different elements into a single indicator per CES requires further research. For example, simple weights using the sentiment and aesthetic quality scores could be used. This would follow similar

weighting approaches taken in survey-based CES assessments using expert opinion and stakeholder views (Nahuelhual et al., 2014; Alvarez-Codoceo et al., 2021). These measures could then be validated by comparing with survey-based data on the most appreciated ecosystem types or other physical attributes (Gould et al., 2019).

5. Conclusion

There is an increasing urgency to address rapidly falling levels of human-nature interaction and policy-makers require relevant indicators to maximise the public health benefits generated by people's positive experiences of nature. CES indicators representing the contributions of nature to people's mental and physical health should therefore also account for this experiential aspect in their measures. In our study, we have demonstrated that social media, CV and NLP methods can identify some variation in the degree to which people gain a positive experience as a result of two CES: landscape enjoyment and the cultural appreciation of wildlife. We find that sentiment increases with aesthetic quality and that different wildlife observations are associated with varying levels of sentiment. These CES and sentiment measures were found to align with well-being measures reported in a national survey of nature recreation in England. At the same time, we find that physical settings and activities also influence people's positive experiences such as those related to social activities. Our findings illustrate the relevancy of applying these novel techniques for CES indicator development in recreational land use management.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data used in this study are open-source and publicly available. Scripts used in this study are available on request.

Acknowledgements

This research was supported by EU Horizon 2020 Grant 817527 (Mapping and Assessment for Integrated ecosystem Accounting) and the WIMEK Graduate Programme of Wageningen University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ecoser.2023.101581>.

References

- Abraham, A., Sommerhalder, K., Abel, T., 2010. Landscape and well-being: a scoping study on the health-promoting impact of outdoor environments. *Int. J. Public Health* 55, 59–69. <https://doi.org/10.1007/s00038-009-0069-z>.
- Abson, D.J., Fischer, J., Leventon, J., Newig, J., Schomerus, T., Vilsmaier, U., von Wehrden, H., Abernethy, P., Ives, C.D., Jager, N.W., Lang, D.J., 2017. Leverage points for sustainability transformation. *Ambio* 46, 30–39. <https://doi.org/10.1007/s13280-016-0800-y>.
- Aerts, R., Honnay, O., Van Nieuwenhuysse, A., 2018. Biodiversity and human health: mechanisms and evidence of the positive health effects of diversity in nature and green spaces. *Br. Med. Bull.* 127, 5–22. <https://doi.org/10.1093/bmb/ldy021>.
- Alshaabi, T., Arnold, M.V., Minot, J.R., Adams, J.L., Dewhurst, D.R., Reagan, A.J., Muhamad, R., Danforth, C.M., Dodds, P.S., 2021. How the world's collective attention is being paid to a pandemic: COVID-19 related n-gram time series for 24 languages on Twitter. *PLoS One* 16, e0244476.
- Alvarez-Codoceo, S., Cerda, C., Perez-Quezada, J.F., 2021. Mapping the provision of cultural ecosystem services in large cities: The case of The Andean piedmont in Santiago, Chile. *Urban For. Urban Green.* 66, 127390 <https://doi.org/10.1016/j.ufug.2021.127390>.

- Austen, G.E., Dallimer, M., Irvine, K.N., Maund, P.R., Fish, R.D., Davies, Z.G., 2021. Exploring shared public perspectives on biodiversity attributes. *People Nat.* 3, 901–913. <https://doi.org/10.1002/pan3.10237>.
- Baklien, B., Ytterhus, B., Bongaardt, R., 2016. When everyday life becomes a storm on the horizon: families' experiences of good mental health while hiking in nature. *Anthropol. Med.* 23, 42–53. <https://doi.org/10.1080/13648470.2015.1056056>.
- Barbieri, F., Camacho-Collados, J., Neves, L., Espinosa-Anke, L., 2020. TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification. <https://doi.org/10.48550/ARXIV.2010.12421>.
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>.
- Becken, S., Stantic, B., Chen, J., Alaei, A.R., Conolly, R.M., 2017. Monitoring the environment and human sentiment on the Great Barrier Reef: Assessing the potential of collective sensing. *J. Environ. Manage.* 203, 87–97. <https://doi.org/10.1016/j.jenvman.2017.07.007>.
- Bieling, C., Plieninger, T., Pirker, H., Vogl, C.R., 2014. Linkages between landscapes and human well-being: An empirical exploration with short interviews. *Ecol. Econ.* 105, 19–30. <https://doi.org/10.1016/j.ecolecon.2014.05.013>.
- Botzat, A., Fischer, L.K., Kowarik, I., 2016. Unexploited opportunities in understanding liveable and biodiverse cities. A review on urban biodiversity perception and valuation. *Glob. Environ. Chang.* 39, 220–233. <https://doi.org/10.1016/j.gloenvcha.2016.04.008>.
- Bratman, G.N., Hamilton, J.P., Daily, G.C., 2012. The impacts of nature experience on human cognitive function and mental health. *Ann. N. Y. Acad. Sci.* 1249, 118–136. <https://doi.org/10.1111/j.1749-6632.2011.06400.x>.
- Brindley, P., Cameron, R.W., Ersoy, E., Jorgensen, A., Maheswaran, R., 2019. Is more always better? Exploring field survey and social media indicators of quality of urban greenspace, in relation to health. *Urban For. Urban Green.* 39, 45–54. <https://doi.org/10.1016/j.ufug.2019.01.015>.
- Bush, L.E., 1973. Individual differences multidimensional scaling of adjectives denoting feelings. *J. Pers. Soc. Psychol.* 25 <https://doi.org/10.1037/h0034274>.
- Cao, H., Wang, M., Su, S., Kang, M., 2022. Explicit quantification of coastal cultural ecosystem services: A novel approach based on the content and sentiment analysis of social media. *Ecol. Indic.* 137, 108756 <https://doi.org/10.1016/j.ecolind.2022.108756>.
- Carrus, G., Scopelliti, M., Laforteza, R., Colangelo, G., Ferrini, F., Salbitano, F., Agrimi, M., Portoghesi, L., Semenzato, P., Sanesi, G., 2015. Go greener, feel better? The positive effects of biodiversity on the well-being of individuals visiting urban and peri-urban green areas. *Landsc. Urban Plan.* 134, 221–228. <https://doi.org/10.1016/j.landurbplan.2014.10.022>.
- Chan, K.M.A., Guerry, A.D., Balvanera, P., Klain, S., Satterfield, T., Basurto, X., Bostrom, A., Chuenpagdee, R., Gould, R., Halpern, B.S., Levine, J., Hannahs, N., Norton, B., Ruckelshaus, M., Russell, R., Tam, J., Woodside, U., 2012. Where are Cultural and Social in Ecosystem Services? A Framework for Constructive Engagement. *Bioscience* 62, 744–756. <https://doi.org/10.1525/bio.2012.62.8.7>.
- Chang, C., Cheng, G.J.Y., Nghiem, T.P.L., Song, X.P., Oh, R.R.Y., Richards, D.R., Carrasco, L.R., 2020. Social media, nature, and life satisfaction: global evidence of the biophilia hypothesis. *Sci. Rep.* 10, 4125. <https://doi.org/10.1038/s41598-020-60902-w>.
- Chesnokova, O., Taylor, J.E., Gregory, I.N., Purves, R.S., 2019. Hearing the silence: the middle ground in the spatial humanities? Extracting and comparing perceived silence and tranquillity in the English Lake District. *Int. J. Geogr. Inf. Sci.* 33, 2430–2454. <https://doi.org/10.1080/13658816.2018.1552789>.
- Clark, K., Khandelwal, U., Levy, O., Manning, C.D., 2019. What Does BERT Look at? An Analysis of BERT's Attention. in: Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP. Association for Computational Linguistics, Florence, Italy, pp. 276–286. <https://doi.org/10.18653/v1/W19-4828>.
- Cody, E.M., Reagan, A.J., Mitchell, L., Dodds, P.S., Danforth, C.M., 2015. Climate Change Sentiment on Twitter: An Unsolicted Public Opinion Poll. *PLoS One* 10, e0136092.
- Cooper, N., Brady, E., Steen, H., Bryce, R., 2016. Aesthetic and spiritual values of ecosystems: Recognising the ontological and axiological plurality of cultural ecosystem 'services'. *Ecosyst. Serv.* 21, 218–229. <https://doi.org/10.1016/j.ecoser.2016.07.014>.
- Cox, D.T.C., Gaston, K.J., 2015. Likeability of Garden Birds: Importance of Species Knowledge & Richness in Connecting People to Nature. *PLoS One* 10, e0141505.
- Cox, D.T.C., Shanahan, D.F., Hudson, H.L., Plummer, K.E., Siriwardena, G.M., Fuller, R.A., Anderson, K., Hancock, S., Gaston, K.J., 2017. Doses of Neighborhood Nature: The Benefits for Mental Health of Living with Nature. *Bioscience* 67, 147–155. <https://doi.org/10.1093/biosci/biw173>.
- Daily, G.C., Polasky, S., Goldstein, J., Kareiva, P.M., Mooney, H.A., Pejchar, L., Ricketts, T.H., Salzman, J., Shallenberger, R., 2009. Ecosystem services in decision making: Time to deliver. *Front. Ecol. Environ.* 7, 21–28. <https://doi.org/10.1890/080025>.
- Dallimer, M., Irvine, K.N., Skinner, A.M.J., Davies, Z.G., Rouquette, J.R., Maltby, L.L., Warren, P.H., Armsworth, P.R., Gaston, K.J., 2012. Biodiversity and the Feel-Good Factor: Understanding Associations between Self-Reported Human Well-being and Species Richness. *Bioscience* 62, 47–55. <https://doi.org/10.1525/bio.2012.62.1.9>.
- de Vries, S., van Dillen, S.M.E., Groenewegen, P.P., Spreeuwenberg, P., 2013. Streetscape greenery and health: Stress, social cohesion and physical activity as mediators. *Soc. Sci. Med.* 94, 26–33. <https://doi.org/10.1016/j.socscimed.2013.06.030>.
- Edens, B., Maes, J., Hein, L., Obst, C., Siikamaki, J., Schenau, S., Javorsek, M., Chow, J., Chan, J.Y., Steurer, A., Alfieri, A., 2022. Establishing the SEEA Ecosystem Accounting as a global standard. *Ecosyst. Serv.* 54, 101413 <https://doi.org/10.1016/j.ecoser.2022.101413>.
- Egarter Vigl, L., Marsoner, T., Giombini, V., Pecher, C., Simion, H., Stemle, E., Tasser, E., Depellegrin, D., 2021. Harnessing artificial intelligence technology and social media data to support Cultural Ecosystem Service assessments. *People Nat.* N/a. <https://doi.org/10.1002/pan3.10199>.
- Eisenstein, J., 2019. *Introduction to natural language processing*. MIT Press, Cambridge.
- Natural England, 2019. Monitor of Engagement with the Natural Environment 2009-19 [WWW Document]. URL <http://publications.naturalengland.org.uk/publication/2248731?category=47018>.
- Englund, O., Berndes, G., Cederberg, C., 2017. How to analyse ecosystem services in landscapes—A systematic review. *Ecol. Indic.* 73, 492–504. <https://doi.org/10.1016/j.ecolind.2016.10.009>.
- Fox, N., Graham, L.J., Eigenbrod, F., Bullock, J.M., Parks, K.E., 2021. Enriching social media data allows a more robust representation of cultural ecosystem services. *Ecosyst. Serv.* 50, 101328 <https://doi.org/10.1016/j.ecoser.2021.101328>.
- Fuller, R.A., Irvine, K.N., Devine-Wright, P., Warren, P.H., Gaston, K.J., 2007. Psychological benefits of greenspace increase with biodiversity. *Biol. Lett.* 3, 390–394. <https://doi.org/10.1098/rsbl.2007.0149>.
- García-Palomares, J.C., Gutiérrez, J., Mínguez, C., 2015. Identification of tourist hot spots based on social networks: A comparative analysis of European metropolises using photo-sharing services and GIS. *Appl. Geogr.* 63, 408–417. <https://doi.org/10.1016/j.apgeog.2015.08.002>.
- Ghermandi, A., Sinclair, M., 2019. Passive crowdsourcing of social media in environmental research: A systematic map. *Glob. Environ. Chang.* 55, 36–47. <https://doi.org/10.1016/j.gloenvcha.2019.02.003>.
- Ghermandi, A., Langemeyer, J., Van Berkel, D., Calcagni, F., Depietri, Y., Egarter Vigl, L., Fox, N., Havinga, I., Jäger, H., Kaiser, N., Karasov, O., McPhearson, T., Podschun, S., Ruiz-Frau, A., Sinclair, M., Venohr, M., Wood, S.A., 2023. Social media data for environmental sustainability: A critical review of opportunities, threats, and ethical use. *One Earth* 6, 236–250. <https://doi.org/10.1016/j.oneear.2023.02.008>.
- Go, A., Bhayani, R., Huang, L., 2009. Twitter sentiment classification using distant supervision. *Stanford, p. 1*. CS224N Proj. report.
- Gosal, A.S., Geizendorffer, I.R., Václavík, T., Poulin, B., Ziv, G., 2019. Using social media, machine learning and natural language processing to map multiple recreational beneficiaries. *Ecosyst. Serv.* 38, 100958 <https://doi.org/10.1016/j.ecoser.2019.100958>.
- Gould, R.K., Morse, J.W., Adams, A.B., 2019. Cultural ecosystem services and decision-making: How researchers describe the applications of their work. *People Nat.* 1, 457–475. <https://doi.org/10.1002/pan3.10044>.
- Gould, R.K., Bremer, L.L., Pascua, P., Meza-Prado, K., 2020. Frontiers in Cultural Ecosystem Services: Toward Greater Equity and Justice in Ecosystem Services Research and Practice. *Bioscience* 70, 1093–1107. <https://doi.org/10.1093/biosci/biaa112>.
- Graham, L.J., Eigenbrod, F., 2019. Scale dependency in drivers of outdoor recreation in England. *People Nat.* 1, 406–416. <https://doi.org/10.1002/pan3.10042>.
- Hartig, T., Mitchell, R., de Vries, S., Frumkin, H., 2014. Nature and Health. *Annu. Rev. Public Health* 35, 207–228. <https://doi.org/10.1146/annurev-publhealth-032013-182443>.
- Havinga, I., Bogaart, P.W., Hein, L., Tuia, D., 2020. Defining and spatially modelling cultural ecosystem services using crowdsourced data. *Ecosyst. Serv.* 43, 101091 <https://doi.org/10.1016/j.ecoser.2020.101091>.
- Havinga, I., Marcos, D., Bogaart, P.W., Hein, L., Tuia, D., 2021. Social media and deep learning capture the aesthetic quality of the landscape. *Sci. Rep.* 11, 20000. <https://doi.org/10.1038/s41598-021-99282-0>.
- Havinga, I., Marcos, D., Bogaart, P., Massimino, D., Hein, L., Tuia, D., 2023. Social media and deep learning reveal specific cultural preferences for biodiversity. *People Nat.* 5, 981–998. <https://doi.org/10.1002/pan3.10466>.
- Hoyle, H., Hitchmough, J., Jorgensen, A., 2017. All about the 'wow factor'? The relationships between aesthetics, restorative effect and perceived biodiversity in designed urban planting. *Landsc. Urban Plan.* 164, 109–123. <https://doi.org/10.1016/j.landurbplan.2017.03.011>.
- Hoyle, H., Norton, B., Dunnett, N., Richards, J.P., Russell, J.M., Warren, P., 2018. Plant species or flower colour diversity? Identifying the drivers of public and invertebrate response to designed annual meadows. *Landsc. Urban Plan.* 180, 103–113. <https://doi.org/10.1016/j.landurbplan.2018.08.017>.
- Huai, S., Van de Voorde, T., 2022. Which environmental features contribute to positive and negative perceptions of urban parks? A cross-cultural comparison using online reviews and Natural Language Processing methods. *Landsc. Urban Plan.* 218, 104307 <https://doi.org/10.1016/j.landurbplan.2021.104307>.
- Johnson, M.L., Campbell, L.K., Svendsen, E.S., McMillen, H.L., 2019. Mapping Urban Park Cultural Ecosystem Services: A Comparison of Twitter and Semi-Structured Interview Methods. *Sustain.* <https://doi.org/10.3390/su11216137>.
- Kashyap, R., Nahapetian, A., 2014. Tweet analysis for user health monitoring, in: 2014 4th International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH). pp. 348–351. <https://doi.org/10.1109/MOBIHEALTH.2014.7015983>.
- Keniger, L.E., Gaston, K.J., Irvine, K.N., Fuller, R.A., 2013. What are the Benefits of Interacting with Nature? *Int. J. Environ. Res. Public Heal.* <https://doi.org/10.3390/ijerph10030913>.
- Koblet, O., Purves, R.S., 2020. From online texts to Landscape Character Assessment: Collecting and analysing first-person landscape perception computationally. *Landsc. Urban Plan.* 197, 103757 <https://doi.org/10.1016/j.landurbplan.2020.103757>.
- Kokil, J., Salvatore, G., Andrew, S.H., L., K.M., H., U.L., C., E.J., 2020. Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-

- driven language methods. *Proc. Natl. Acad. Sci.* 117, 10165–10171. <https://doi.org/10.1073/pnas.1906364117>.
- Langemeyer, J., Calcagni, F., 2022. Virtual spill-over effects: What social media has to do with relational values and global environmental stewardship. *Ecosyst. Serv.* 53, 101400 <https://doi.org/10.1016/j.ecoser.2021.101400>.
- Lee, H., Seo, B., Cord, A.F., Volk, M., Lautenbach, S., 2022. Using crowdsourced images to study selected cultural ecosystem services and their relationships with species richness and carbon sequestration. *Ecosyst. Serv.* 54, 101411 <https://doi.org/10.1016/j.ecoser.2022.101411>.
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V., 2019. RoBERTa: A Robustly Optimized BERT Pretraining Approach. <https://doi.org/10.48550/ARXIV.1907.11692>.
- Loff, J., Reis, M., Martins, B., 2015. Predicting Well-Being with Geo-Referenced Data Collected from Social Media Platforms, in: Proceedings of the 30th Annual ACM Symposium on Applied Computing, SAC '15. Association for Computing Machinery, New York, NY, USA, pp. 1167–1173. <https://doi.org/10.1145/2695664.2695939>.
- Loureiro, D., Barbieri, F., Neves, L., Anke, L.E., Camacho-Collados, J., 2022. TimeLMs: Diachronic Language Models from Twitter. <https://doi.org/10.48550/ARXIV.2202.03829>.
- Mace, G.M., Norris, K., Fitter, A.H., 2012. Biodiversity and ecosystem services: a multilayered relationship. *Trends Ecol. Evol.* 27, 19–26. <https://doi.org/10.1016/j.tree.2011.08.006>.
- Methorst, J., Bonn, A., Marselle, M., Böhning-Gaese, K., Rehdanz, K., 2021. Species richness is positively related to mental health – A study for Germany. *Landscape Urban Plan.* 211, 104084 <https://doi.org/10.1016/j.landurbplan.2021.104084>.
- Miller, J.R., 2005. Biodiversity conservation and the extinction of experience. *Trends Ecol. Evol.* 20, 430–434. <https://doi.org/10.1016/j.tree.2005.05.013>.
- Mitchell, L., Frank, M.R., Harris, K.D., Dodds, P.S., Danforth, C.M., 2013. The Geography of Happiness: Connecting Twitter Sentiment and Expression, Demographics, and Objective Characteristics of Place. *PLoS One* 8, e64417.
- Moreno-Llorca, R., Méndez, F.P., Ros-Candeira, A., Alcaraz-Segura, D., Santamaría, L., Ramos-Ridao, Á.F., Revilla, E., Bonet-García, F.J., Vaz, A.S., 2020. Evaluating tourist profiles and nature-based experiences in Biosphere Reserves using Flickr: Matches and mismatches between online social surveys and photo content analysis. *Sci. Total Environ.* 737, 140067 <https://doi.org/10.1016/j.scitotenv.2020.140067>.
- Nahuelhual, L., Carmona, A., Latorra, P., Barrena, J., Aguayo, M., 2014. A mapping approach to assess intangible cultural ecosystem services: The case of agriculture heritage in Southern Chile. *Ecol. Indic.* 40, 90–101. <https://doi.org/10.1016/j.ecolind.2014.01.005>.
- Nordh, H., Evensen, K.H., Skår, M., 2017. A peaceful place in the city—A qualitative study of restorative components of the cemetery. *Landscape Urban Plan.* 167, 108–117. <https://doi.org/10.1016/j.landurbplan.2017.06.004>.
- Ostermann, F.O., 2021. Linking Geosocial Sensing with the Socio-Demographic Fabric of Smart Cities. *ISPRS Int. J. Geo-Inf.* <https://doi.org/10.3390/ijgi10020052>.
- Oteros-Rozas, E., Martín-López, B., Fagerholm, N., Bieling, C., Plieninger, T., 2018. Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. *Ecol. Indic.* 94, 74–86. <https://doi.org/10.1016/j.ecolind.2017.02.009>.
- Patterson, G., Xu, C., Su, H., Hays, J., 2014. The SUN attribute database: Beyond categories for deeper scene understanding. *Int. J. Comput. Vis.* 108, 59–81.
- Plieninger, T., Dijks, S., Oteros-Rozas, E., Bieling, C., 2013. Assessing, mapping, and quantifying cultural ecosystem services at community level. *Land Use Policy* 33, 118–129. <https://doi.org/10.1016/j.landusepol.2012.12.013>.
- Poria, S., Hazarika, D., Majumder, N., Mihalcea, R., 2020. Beneath the Tip of the Iceberg: Current Challenges and New Directions in Sentiment Analysis Research. *IEEE Trans. Affect. Comput.* 1 <https://doi.org/10.1109/TAFFC.2020.3038167>.
- Quinton, J.M., Duinker, P.N., 2018. Beyond burial: researching and managing cemeteries as urban green spaces, with examples from Canada. *Environ. Rev.* 27, 252–262. <https://doi.org/10.1139/er-2018-0060>.
- Raymond, C.M., Kenter, J.O., Plieninger, T., Turner, N.J., Alexander, K.A., 2014. Comparing instrumental and deliberative paradigms underpinning the assessment of social values for cultural ecosystem services. *Ecol. Econ.* 107, 145–156. <https://doi.org/10.1016/j.ecolecon.2014.07.033>.
- Richards, D.R., Tunçer, B., 2018. Using image recognition to automate assessment of cultural ecosystem services from social media photographs. *Ecosyst. Serv.* 31, 318–325. <https://doi.org/10.1016/j.ecoser.2017.09.004>.
- Russell, R., Guerry, A.D., Balvanera, P., Gould, R.K., Basurto, X., Chan, K.M.A., Klain, S., Levine, J., Tam, J., 2013. Humans and Nature: How Knowing and Experiencing Nature Affect Well-Being. *Annu. Rev. Environ. Resour.* 38, 473–502. <https://doi.org/10.1146/annurev-environ-012312-110838>.
- Samani, Z.R., Guntuku, S.C., Moghaddam, M.E., Preotjiuc-Pietro, D., Ungar, L.H., 2018. Cross-platform and cross-interaction study of user personality based on images on Twitter and Flickr. *PLoS One* 13, e0198660.
- Sandifer, P.A., Sutton-Grier, A.E., Ward, B.P., 2015. Exploring connections among nature, biodiversity, ecosystem services, and human health and well-being: Opportunities to enhance health and biodiversity conservation. *Ecosyst. Serv.* 12, 1–15. <https://doi.org/10.1016/j.ecoser.2014.12.007>.
- ScenicOrNot, 2015. ScenicOrNot Dataset [WWW Document]. accessed 5.3.20. <http://scenicornot.datascienceclab.co.uk>.
- Schwartz, A.J., Dodds, P.S., O'Neil-Dunne, J.P.M., Danforth, C.M., Ricketts, T.H., 2019. Visitors to urban greenspace have higher sentiment and lower negativity on Twitter. *People Nat.* <https://doi.org/10.1002/pan3.10045>.
- Seresinhe, C.L., Preis, T., MacKerron, G., Moat, H.S., 2019. Happiness is Greater in More Scenic Locations. *Sci. Rep.* 9, 4498. <https://doi.org/10.1038/s41598-019-40854-6>.
- Soga, M., Gaston, K.J., 2016. Extinction of experience: the loss of human–nature interactions. *Front. Ecol. Environ.* 14, 94–101. <https://doi.org/10.1002/fee.1225>.
- Stelzmüller, C., Tanzer, S., Schedl, M., 2021. Cross-City Analysis of Location-Based Sentiment in User-Generated Text, in: Companion Proceedings of the Web Conference 2021, WWW '21. Association for Computing Machinery, New York, NY, USA, pp. 339–346. <https://doi.org/10.1145/3442442.3451889>.
- Stevenson, N., Farrell, H., 2018. Taking a hike: exploring leisure walkers embodied experiences. *Soc. Cult. Geogr.* 19, 429–447. <https://doi.org/10.1080/14649365.2017.1280615>.
- van Zanten, B.T., Van Berkel, D.B., Meentemeyer, R.K., Smith, J.W., Tieskens, K.F., Verburg, P.H., 2016. Continental-scale quantification of landscape values using social media data. *Proc. Natl. Acad. Sci. U. S. A.* 113, 12974–12979. <https://doi.org/10.1073/pnas.1614158113>.
- Wartmann, F.M., Tieskens, K.F., van Zanten, B.T., Verburg, P.H., 2019. Exploring tranquillity experienced in landscapes based on social media. *Appl. Geogr.* 113, 102112 <https://doi.org/10.1016/j.apgeog.2019.102112>.
- Waterloo, S.F., Baumgartner, S.E., Peter, J., Valkenburg, P.M., 2017. Norms of online expressions of emotion: Comparing Facebook, Twitter, Instagram, and WhatsApp. *New Media Soc.* 20, 1813–1831. <https://doi.org/10.1177/1461444817707349>.
- White, M.P., Pahl, S., Ashbullby, K., Herbert, S., Depledge, M.H., 2013. Feelings of restoration from recent nature visits. *J. Environ. Psychol.* 35, 40–51. <https://doi.org/10.1016/j.jenvp.2013.04.002>.
- Wolf, I.D., Wohlfart, T., 2014. Walking, hiking and running in parks: A multidisciplinary assessment of health and well-being benefits. *Landscape Urban Plan.* 130, 89–103. <https://doi.org/10.1016/j.landurbplan.2014.06.006>.
- Wood, S.A., Guerry, A.D., Silver, J.M., Lacayo, M., 2013. Using social media to quantify nature-based tourism and recreation. *Sci. Rep.* 3, 1–7. <https://doi.org/10.1038/srep02976>.
- Workman, S., Souvenir, R., Jacobs, N., 2017. Understanding and Mapping Natural Beauty. *2017 IEEE Int. Conf. Comput. Vis.* 4, 5590–5599.
- Yan, T., Liu, F., 2021. Sentiment Analysis and Effect of COVID-19 Pandemic using College SubReddit Data. <https://doi.org/10.48550/ARXIV.2112.04351>.
- Zhang, F., Zhou, B., Ratti, C., Liu, Y., 2020. Discovering place-informative scenes and objects using social media photos. *R. Soc. Open Sci.* 6, 181375 <https://doi.org/10.1098/rsos.181375>.
- Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., Torralla, A., 2017. Places: A 10 million Image Database for Scene Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*
- Zvariková, M., Prokop, P., Zvarík, M., Ježová, Z., Medina-Jerez, W., Fedor, P., 2021. What Makes Spiders Frightening and Disgusting to People? *Front. Ecol. Evol.* 9 <https://doi.org/10.3389/fevo.2021.694569>.