Geo-Data for Mapping Scenic Beauty: Exploring the Potential of Remote Sensing and Social Media

2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS

Havinga, Ilan; Marcos, Diego; Bogaart, Patrick W.; Hein, Lars; Tuia, Devis

https://doi.org/10.1109/IGARSS47720.2021.9553417

GEO-DATA FOR MAPPING SCENIC BEAUTY: EXPLORING THE POTENTIAL OF REMOTE SENSING AND SOCIAL MEDIA

Ilan Havinga[1], Diego Marcos[2], Patrick W. Bogaart[3], Lars Hein[1], Devis Tuia[2,4]

[1]Environmental Systems Analysis Group, Wageningen University, the Netherlands

[2]Laboratory of Geo-Information Science and Remote Sensing, Wageningen University, the Netherlands

[3]National Accounts Department, Statistics Netherlands, the Netherlands

[4]Environmental Computational Science and Earth Observation Laboratory, Ecole Polytechnique Fédérale de Lausanne, Switzerland

ilan.havinga@wur.nl

ABSTRACT

Scenic beauty is an important contributing factor to peoples' well-being. Modelling scenic beauty has been made possible at large scales with the availability of open-source remote sensing products. At the same time, the metadata available through social media, including tags and descriptions, offer a novel modelling alternative with a personalised view from the ground. This is especially relevant to policy applications. Using a crowdsourced landscape aesthetics dataset called ScenicOrNot as ground truth, we develop and test models to predict scenic beauty based on remotely sensed indicators and image metadata from social media (Flickr). Initial results show that both model types generate strong predictions of scenic beauty and model accuracy is maximised when the two are combined. Our research shows that both a top-view measurement using remote sensing and a social media-based measurement from the ground can be used to model landscape aesthetics in support of sustainable policy goals.

***Index Terms*—** Landscape aesthetics, Social sensing, Social media, Big data, NLP

1. INTRODUCTION

In Europe, the beauty of a natural landscape, or *scenic beauty*, is a key source of cultural value and an important contributing factor to peoples' mental health [1]. This value is most commonly generated during peoples' outdoor recreation. Consequently, land use policy to manage the aesthetic quality of the landscape can play an important role in realising the Sustainable Development Goals (SDGs) in European countries. In particular, SDG 3, which specifically considers mental health.

Open-source remote sensing products have allowed large-scale measurement of sustainability indicators, including those relevant to peoples' quality of life [2, 3]. Quantitative studies of scenic beauty are based on spatially-explicit environmental indicators, many derived from remote sensing products [4]. In one national-level study, [5] modelled landscape aesthetics for the whole of Germany. Such indicator models have the advantage of being ubiquitous and readily obtained. However, these models are rigid and are fixed to land cover and landscape structure rather than the views of individuals. A lack of model validation with *in situ* data is another key challenge [4].

At the same time, the huge volumes of geo-referenced data being posted to social media sites such as Flickr have enabled new "social sensing" approaches for measuring sustainability-related indicators [6]. Such data can provide complementary information analogous to the purposes of remote sensing applications, but also richer in the sense that individuals' perspectives are taken into account. This is especially relevant in policy-focused contexts such as ecosystem service assessments [7]. Due to the volume of Flickr data, machine learning provides an appropriate method with which to model indicators at large scales, including scenic beauty.

Recent research has explored the potential of social media data in measuring the aesthetic contributions of the landscape. Drawing upon a user-generated image database of scenic beauty in Great Britain (GB), ScenicOrNot (SoN[1]), [8, 9] mapped landscape scenicness in ground images using deep learning. Subsequent work examined the factors that trigger the perception of scenic landscapes, e.g. by identifying groups of attributes determining scenicness [10], and connecting new attributes using ancillary text corpora [11]. Accurate predictions of scenicness have also been produced by models based on Geograph image tags [12], a web-based crowdsourcing project, and Flickr metadata, including occurrences of the word "scenic" and similar words [13]. In contrast to image-based models, model predictions based on the words users employ to describe their environments have the potential of accurately modelling landscape aesthetics while capturing the self-reported value of scenic beauty to users.

In our research, we explore this potential of Flickr image metadata in modelling individuals' perception of scenic beauty in comparison to (and in combination with) an environmental indicator approach. In doing so, we seek to advance landscape aesthetic modelling. We make the comparison in GB, using the SoN database as our ground-truth.

[1]http://scenicornot.datasciencelab.co.uk/

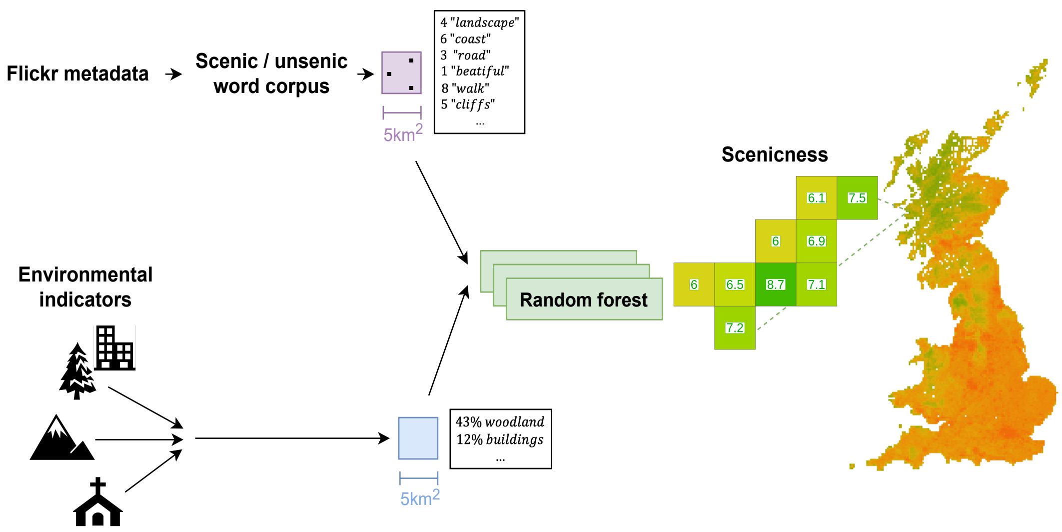

Fig. 1: The overall research design comparing model accuracy using environmental indicator variables and a variable set based on Flickr metadata at a $5km^2$ grid cell resolution.

2. METHODOLOGY

The research methodology is shown in Fig. 1. As variables, the models draw on 1) a reduced set of environmental indicators linked to visual concepts in the landscape aesthetics literature [14] and 2) Flickr metadata including relative term frequencies of the most and least scenic words.

2.1. Environmental indicator variables

Variable selection. We used four groups of indicator variables, each extracted using a $5km^2$ grid over GB: first, we considered *naturalness* as an important visual factor because of our innate biological need to affiliate with nature. Consequently, the percentage of ecosystem types was calculated, representing different levels of naturalness. This produced an initial set of 45 variables.

Visual scale is also a key driver of peoples' aesthetic experience because of our evolutionary history as both predator and prey. Thus, higher elevations and the possible refuge afforded by elevation differences generate greater aesthetic appeal. To reflect this, the difference in elevation in *m* was calculated.

In addition, *complexity* plays an important role because of our physiological need to explore. To capture this, the Patch Density Index (PDI), representing the number of continuous ecosystem type patches, and the Shannon Diversity Index (SDI) were calculated. SDI was calculated as:

$$SDI = \sum_{i=1}^{m} P_i \cdot \ln P_i, \qquad (1)$$

where m is the number of ecosystem types and P_i is the proportion of area covered by ecosystem type i.

The last indicator group captured *uniqueness*. A natural feature in the city can hold a much larger aesthetic value than in more natural contexts. Conversely, unique cultural elements in the landscape generate a similar effect. Thus, the number of historical Points of Interest (POI), a single variable, and the relative difference in percentage area of each ecosystem type within $10km^2$, an additional 45 variables, were also calculated.

Variable reduction: ecosystem variables that could be calculated for less than $100km^2$ (0.04% of the surface of GB) were first removed. Second, a check for variable collinearity was performed. The model accuracy effect of removing variables with a correlation $r \geq 0.7$ was measured and those with the smallest effect were removed. A final set of 41 variables were used to build a random forest (RF) regressor predicting scene beauty for each grid cell covering GB.

2.2. Flickr metadata variables

Scenic / unscenic corpus: we constructed a corpus of words from the text in Flickr images surrounding geo-located SoN images rated for scenicness. A balanced 10% sample across 10 rating categories (1-10) was taken of SoN images with the lowest variance in user ratings (each SoN image is rated by several people, see next section). Text from the titles, tags and descriptions of Flickr images within $500m$ of the SoN images were extracted. Stopwords, html, numbers, punctuation and words shorter than three characters were removed and made unique to each location. Local toponyms were removed using the OS Names API[2]. The total SoN image location frequency, mean and standard deviation in scenicness

[2] https://osdatahub.os.uk/docs/names/overview

was then calculated per word. The 100 most and least scenic words occurring in at least 1% of SoN locations were then extracted to compile the final corpus.

Variables: metadata was first reduced to one unique image per user per day per grid cell. This was done to achieve a more balanced set of user preferences. As a set of 200 variables for a Flickr-based RF regressor, the term frequency–inverse document frequency (tf-idf [15]) of words in the scenic / unscenic corpus was then calculated per grid cell using the unique words per intersecting Flickr image record.

3. EXPERIMENTAL SETTING

Three data sources were used to train the models: 1) a geo-tagged image dataset of landscape scenic beauty (or *scenicness*); 2) gridded and spatial feature datasets to calculate indicator variables, and 3) Flickr metadata records.

SoN images: SoN is a crowdsourcd dataset composed of more than 120k geo-located ground images of GB, each one labeled by volunteers with scores of scenic beauty. Per image, we took the average in user votes. To generate the ground truth for our experiments, the point coordinates of single SoN images were mapped and assigned a $5km^2$ grid cell. A mean scenicness rating was then calculated per cell. We randomly split the grid cells into a 70% training, 10% validation and 20% test dataset using a $50km^2$ grid overlay to reduce spatial autocorrelation. Image coordinates and scenicness scores were also used to extract the most scenic / unscenic words from the metadata of nearby Flickr images (see Section 2.2).

Indicator variables: The ecosystem-related variables were generated using the European Environment Agency (EEA) ecosystem type map which uses the EUNIS (European Nature Information System) habitat classification at $100m^2$ [16]. The EU Digital Elevation Model (DEM) at $25m^2$ was used to calculate elevation difference [17] and OpenStreetMap (OSM) provided the POI per grid cell.

Flickr metadata: We drew on Flickr image metadata composed of 9.8 million geo-located, outdoor images taken between 2004 and 2020 within the entire terrestrial area of GB. The metadata were retrieved through the Flickr API using a $1km^2$ search grid. Indoor images were removed by applying the pre-trained Places365-ResNet50 model[3] and taking it's binary indoor/outdoor predictions to filter out images related to indoor scenes. Flickr image text consisted of the image title, description and tags. The image content was not used.

4. RESULTS & DISCUSSION

The mean and standard deviation in scenicness for the words in the scenic / unscenic word corpus are shown in Figure 2.

[3]https://github.com/CSAILVision/places365

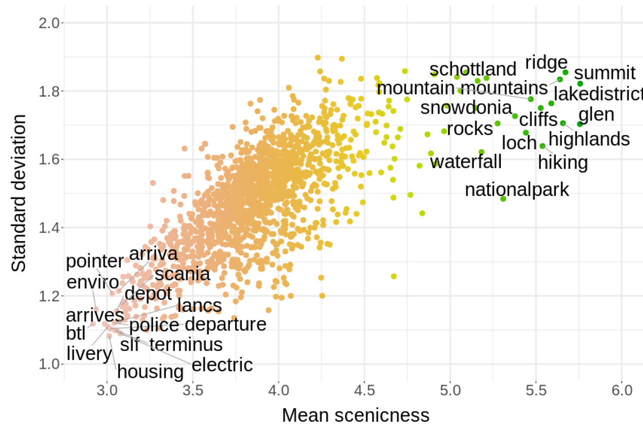

Fig. 2: Mean and standard deviation in scenicness for the words present in at least 1% of SoN locations in the Flickr text corpus.

Words with the highest scenicness are mostly nouns related to well-known aesthetic landscape features such as "mountains", "summit" and "waterfall" as well as words related to recreational activities such as "hiking". Conversely, the least scenic words relate to the urban environment such as "housing", "depot" and "police". The standard deviation shows that the most scenic words are employed much more widely, not just in very scenic environments.

Model accuracy results are reported in Table 1. Accuracy on the test set is reported using R^2, root mean squared error (RMSE) and Kendall's τ, a ranking correlation coefficient between -1 (*inverse correlation*) and 1 (*absolute correlation*). Figure 3 shows a spatial comparison between the SoN observed values and the predicted values of the indicator and Flickr models as well as a combination RF model, trained using both the indicator and Flickr-based variables. Taken individually, the Flickr-based model achieved a strong but less powerful level of model accuracy versus the environmental indicator model.

The Flickr model accuracy could be further improved if variables were generated using Flickr imagery as well as metadata, and we plan to explore this possibility in our future work. Still, the Flickr metadata-based variables enable a greater level of model accuracy when combined with the indicator model, providing the most accurate model of scenic beauty, which includes individuals' views of the landscape. This model reaches an R^2 of 0.829.

Table 1: Scenicness model accuracy results on the $5km^2$ grid test set derived from SoN labels.

Model	$\mathbf{R^2}$	**RMSE**	**Kendall's τ**
Flickr	0.679	0.626	0.631
Indicator	0.820	0.469	0.730
Combination	**0.829**	**0.459**	**0.739**

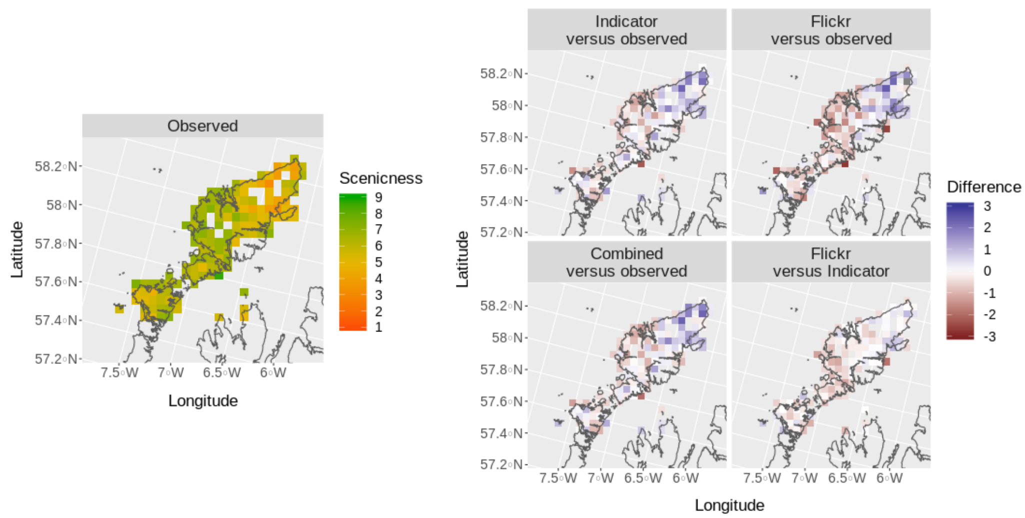

Fig. 3: Observed versus predicted values and inter-model comparisons on the Isle of Lewis, Scotland, using the indicator, Flickr and combination models.

5. CONCLUSIONS

We have shown that both environmental indicator and social media-based variables enable accurate models of landscape aesthetics at national level. In combination, model accuracy is further improved. Further research using Flickr should focus on using the images in combination with metadata to capture an even more accurate measure of landscape aesthetics.

6. REFERENCES

[1] A. Abraham, K. Sommerhalder, and T. Abel, "Landscape and well-being: a scoping study on the health-promoting impact of outdoor environments," *International Journal of Public Health*, vol. 55, no. 1, pp. 59–69, 2010.

[2] R. Avtar, A. A. Komolafe, A. Kouser, D. Singh, A. P. Yunus, J. Dou, P. Kumar, R. D. Gupta, B. A. Johnson, H. V. Thu Minh, A. K. Aggarwal, and T. A. Kurniawan, "Assessing sustainable development prospects through remote sensing: A review," *Remote Sensing Applications: Society and Environment*, vol. 20, pp. 100402, 2020.

[3] M. Sapena, M. Wurm, H. Taubenböck, D. Tuia, and L. Ruiz, "Estimating quality of life dimensions from urban spatial pattern metrics," *Computers Environment and Urban Systems*, vol. 85, pp. 101549, 2021.

[4] O. Karasov, M. Külvik, and I. Burdun, "Deconstructing landscape pattern: applications of remote sensing to physiognomic landscape mapping," *GeoJournal*, 2019.

[5] J. Hermes, C. Albert, and C. von Haaren, "Assessing the aesthetic quality of landscapes in Germany," *Ecosystem Services*, vol. 31, pp. 296–307, 2018.

[6] Y. Liu, X. Liu, S. Gao, L. Gong, C. Kang, Y. Zhi, G. Chi, and L. Shi, "Social Sensing: A New Approach to Understanding Our Socioeconomic Environments," *Annals of the Association of American Geographers*, vol. 105, no. 3, pp. 512–530, 2015.

[7] I. Havinga, P. W. Bogaart, L. Hein, and D. Tuia, "Defining and spatially modelling cultural ecosystem services using crowdsourced data," *Ecosystem Services*, vol. 43, pp. 101091, 2020.

[8] S. Workman, R. Souvenir, and N. Jacobs, "Understanding and Mapping Natural Beauty," *2017 IEEE International Conference on Computer Vision (ICCV)*, vol. 4, pp. 5590–5599, 2017.

[9] C. I. Seresinhe, P. Tobias, and H. S. Moat, "Using deep learning to quantify the beauty of outdoor places," *Royal Society Open Science*, vol. 4, no. 7, 2017.

[10] D. Marcos, R. Fong, S. Lobry, R. Flamary, N. Courty, and D. Tuia, "Contextual semantic interpretability," *Asian Conference on Computer Vision (ACCV)*, 2020.

[11] P. Arendsen, D. Marcos, and D. Tuia, "Concept Discovery for The Interpretation of Landscape Scenicness," *Machine Learning and Knowledge Extraction*, vol. 2, no. 4, 2020.

[12] O. Chesnokova, M. Nowak, and R. S. Purves, "A Crowdsourced Model of Landscape Preference," *13th International Conference on Spatial Information Theory (COSIT 2017)*, vol. 86, 2017.

[13] C. I. Seresinhe, H. S. Moat, and T. Preis, "Quantifying scenic areas using crowdsourced data," *Environment and Planning B: Urban Analytics and City Science*, vol. 45, no. 3, pp. 567–582, 2017.

[14] M. Tveit, Å. Ode, and G. Fry, "Key concepts in a framework for analysing visual landscape character," *Landscape Research*, vol. 31, no. 3, pp. 229–255, 2006.

[15] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*, Cambridge University Press, Cambridge, 2008.

[16] European Environment Agency, "Ecosystem types of Europe," 2019.

[17] European Environment Agency, "Copernicus Land Monitoring Service - EU-DEM," 2017.