CO-062

EXPLORING THE RELATIONSHIP BETWEEN LAND COVER AND SUBJECTIVE EVALUATION OF SCENIC BEAUTY THROUGH USER GENERATED CONTENT

STADLER B., PURVES R., TOMKO M.
University of Zurich, ZURICH, SWITZERLAND

INTRODUCTION

User generated content (UGC) offers the opportunity to explore classic geographic problems, such as defining the boundaries of vernacular regions (Montello et al., 2003; Hollenstein and Purves, 2010) or the nature of the terms used to describe geographic concepts (Edwardes and Purves, 2007) using the so-called "wisdom of the crowd". The basic notion used in such work is that, since very large volumes of data may describe either the location of a named object or the attributes of a particular location, it is possible to explore these properties empirically through the use of UGC. Essentially, UGC can simply be described as information uploaded onto the web by individuals, for example in the form of tagged Flickr images (www.flickr.com), reports on the state of repair of local streets (http://www.fixmystreet.com/), offerings of objects for sale in local classified adverts (http://www.gumtree.com/) and the generation of crowd sourced mapping (http://www.openstreetmap.org/). Where specifically geographic information is created, for example through adding coordinates to Flickr images, then UGC can also be termed Volunteered Geographic Information (VGI) defined by Goodchild (2007) as referring specifically to the "the widespread engagement of large numbers of private citizens, often with little in the way of formal qualifications, in the creation of geographic information".

UGC (and VGI) offers us an opportunity to explore a wide variety of opinions and processes in quasi-realtime, for example analysing crimes through the use of Twitter (White and Roth, 2010). However, in this paper we focus on subjective opinions of scenic beauty, as collected through UGC allowing individuals to rate the "scenicness" of georeferenced images in Great Britain displayed on the web. We hypothesise that such ratings provide us with a rich dataset capable of capturing a broadbrush view of where scenic locations are generally agreed to be, and use the resulting data to explore the relationship between scenic beauty and land cover, and important question for those who wish to predict scenic beauty using other, typically more readily available, geographic data.

Scenic or landscape beauty is important since it is often considered as an important public good, with a value which should be preserved and taken account of within decision making and planning processes. However, unlike properties such as the presence of a particular species of plant or animal, preferences for landscapes are subjective and hard to quantify. Indeed, the nature of landscape as a thing of worth is considered by Hull and Revell (1989) to vary according to cultural and social backgrounds. Nonetheless, many empirical studies rating landscapes (typically through the use of images) have been made, with for example Ribe (2002) showing a correlation between scenic beauty ratings for groups of evaluators with differing backgrounds. A variety of studies have explored how varying landscape elements and compositions influence their rating, with Grêt-Regamey et al. (2007) demonstrating that respondents valued landscapes without new developments such as ski pistes and buildings more, and Lothian (2008) showing the negative influence of the development of wind farms in areas rated as scenic. Palmer (2004) demonstrated that land cover classes composing a landscape had a measureable influence over the rating of landscape images, while Kaplan and Kaplan (1989) showed that landscapes were rated more highly where they were coherent – that is to say composed of relatively few homogeneous land cover classes. Despite considerable debate over the efficacy of images as a proxy for evaluating landscape beauty in situ (e.g. Hull et al. (1992); Meitner and Daniel (1997)), practical grounds mean that empirical studies of landscape beauty are primarily carried out using images, though care is required in defining the nominal location of the landscape evaluator and the resulting extent of the visible landscape.

In this paper we explore the relationship between scenicness, as evaluated for more than 175000 images, geography and land cover. In the following we first briefly introduce the data used in the study, before reviewing the methods applied to relate land cover to scenicness ratings. The results thus obtained are then presented, showing how the pattern of scenic beauty varies in Great Britain and the relationship between scenicness and land cover. Finally, we discuss the potential and limitations in the use of work such as this in exploring a broad range of geographic questions.

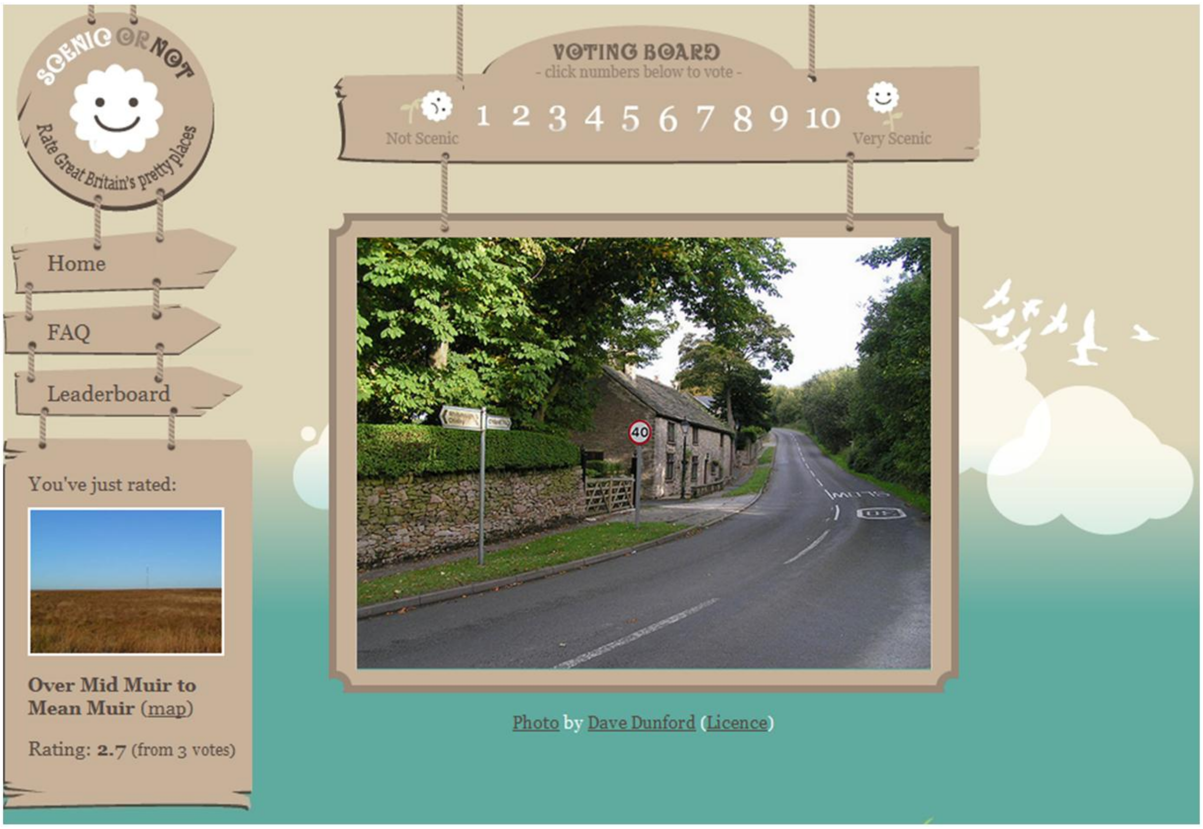

Figure 1. ScenicOrNot interface for rating images (scenic.mysociety.org)

DATA

In this paper we use the opportunity afforded to us by the collection of two UGC datasets, Geograph (http://www.geograph.org.uk/) and ScenicOrNot (http://scenic.mysociety.org/) to explore landscape preferences in Great Britain. ScenicOrNot is a web site which allows people to "Rate Great Britain's pretty places". It uses images collected from Geograph, a project with the aim to collect "geographically representative photographs and information for every square kilometre of the UK and the Republic of Ireland" to collect ratings of the scenic quality of images. After an image has received at least three ratings from different users it is made available along with each individual rating, a link to the original image and its location under a Creative Commons licence. Figure 1 shows the ScenicOrNot interface, where users are asked to rate images between 1 (Not scenic) and 10 (Very scenic). A total of 178717 images with three or more ratings were used for this study in October 2009. Geograph images are all located with a precision of at least 1km within grid squares, and more recently uploaded images also often have information on the direction in which the image was taken.

Landscape ratings were compared to land cover classes as represented by CORINE (http://www.eea.europa.eu/publications/COR0-landcover). CORINE is a European dataset, describing land cover at a nominal scale of 1:100000, with a resolution of 100m. CORINE has three levels of description, a top level with 5classes, an intermediate level with 15 classes and a detailed level with 44 classes. We also utilised a Digital Elevation Model from the Shuttle Radar Topography Mission (SRTM) with a nominal resolution of 90m to capture terrain properties for the calculation of viewsheds and surface roughness (http://www2.jpl.nasa.gov/srtm/).

LINKING LANDSCAPE RATINGS TO LAND COVER

The most important task in comparing the ratings collected from ScenicOrNot with land cover values obtained from CORINE was to develop a repeatable methodology for assigning land cover values to landscape ratings. The simplest method of relating a landscape rating to CORINE, is simply to use the land cover class found at the point associated with the image. However, a point is not representative of the view encompassed by an image and thus we experimented with relating some area (and associated land cover classes) with images. A number of approaches are possible, of which a subset are listed here, along with the justification for their application:

a) Only the land cover class at the location of the image is considered

b) 1km2 grid cell: Each Geograph image aims to be representative of a 1km grid square – thus the land cover classes found in the whole grid square are representative of the image

c) 1km2 buffer centred on image location: Method (b) ignores the position of the image with respect to the grid square and has variable maximum ranges to land cover classes in different cardinal directions.

d) Buffer with variable range, defined by land cover classes and terrain, centred on image location: Here the maximum visible distance is defined according to the land cover classes present (less for anthropogenic classes) and terrain roughness (less for rougher terrains)

e) Directed cone with variable range, defined by land cover classes and terrain (as d), centred on image location: A directed cone should better represent the actual visible area of a photograph than a circular buffer as used in c & d.

f) Directed viewshed, with maximum range of 2km: Viewsheds incorporate the actual visible area in terrain, and should thus give a good representation of the land cover classes visible in particular direction from a given location.

The scenicness rating of the image in question was then assigned to the land cover class identified, weighted by the area covered by that land cover class within the area assigned by methods b-f. Two variants were applied – in (i) only the polygon found at the image location was considered and in (ii) all land cover polygons found within the calculated area were assigned scenicness ratings. For (a) only the former was applied, since by definition a point can only be associated with a single land cover class. For (f), the viewshed, only the case with all land cover polygons was used, since a viewshed may have no visibility at the image location. Figure 2 shows the application of three of the methods described schematically for an exemplar image.

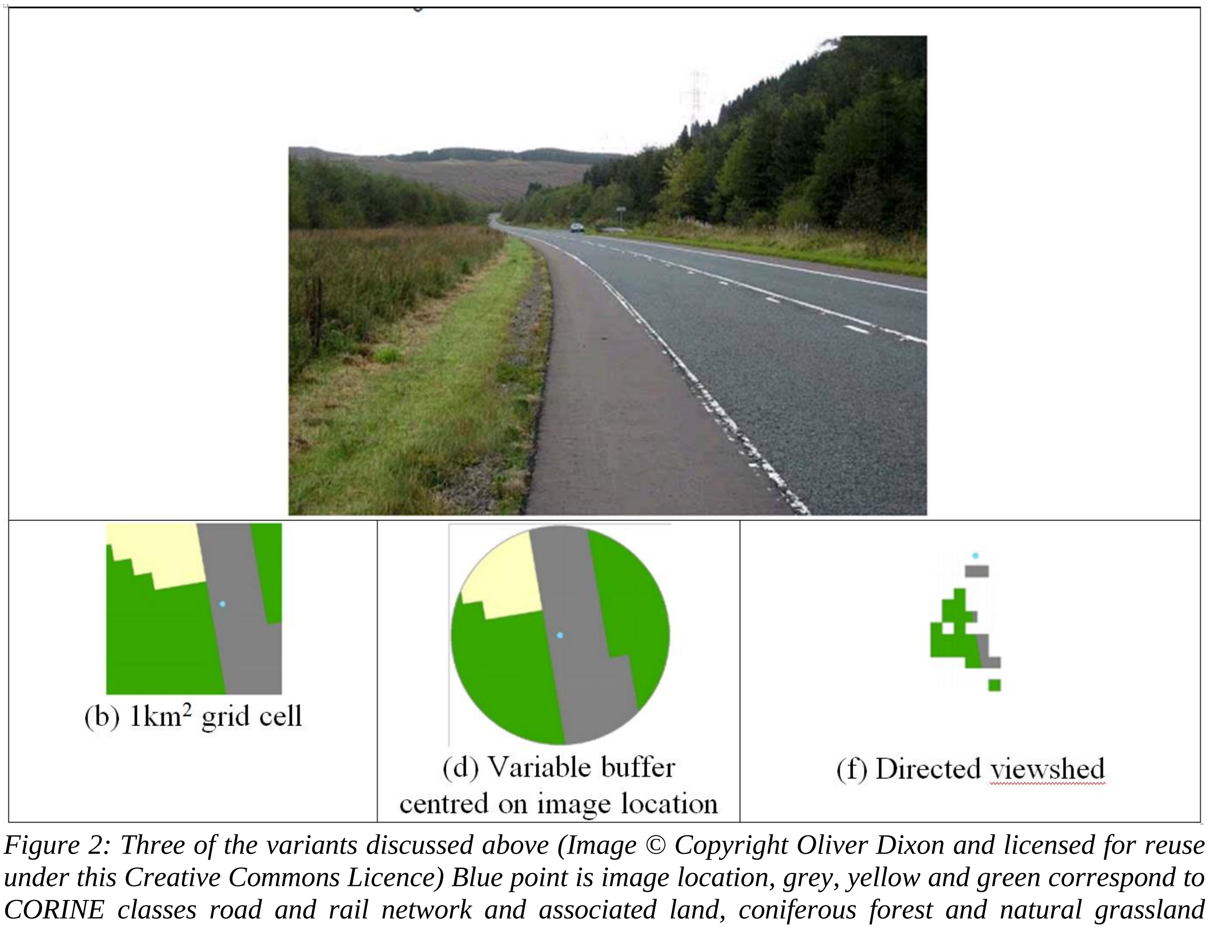

Figure 2: Three of the variants discussed above (Image © Copyright Oliver Dixon and licensed for reuse under this Creative Commons Licence) Blue point is image location, grey, yellow and green correspond to CORINE classes road and rail network and associated land, coniferous forest and natural grassland respectively.

RESULTS AND DISCUSSION

Before proceeding to explore the relationship between scenicness and land cover, it is important to explore how the area covered by CORINE classes relates to the number of images found within each class. If a CORINE class is significantly under or overrepresented then this may in turn have an effect on the scenicness rating for this class, and may also suggest that particular types of location are preferred for images in Geograph. Figure 3 shows a scatter plot of the relative area of CORINE classes and the number of images taken per CORINE class. Although a few classes are over or under represented (e.g. despite a total relative area of 12.3% only 8.6% of images are taken on moors and heathland), it is clear that in general images are proportionately distributed across all land cover types.

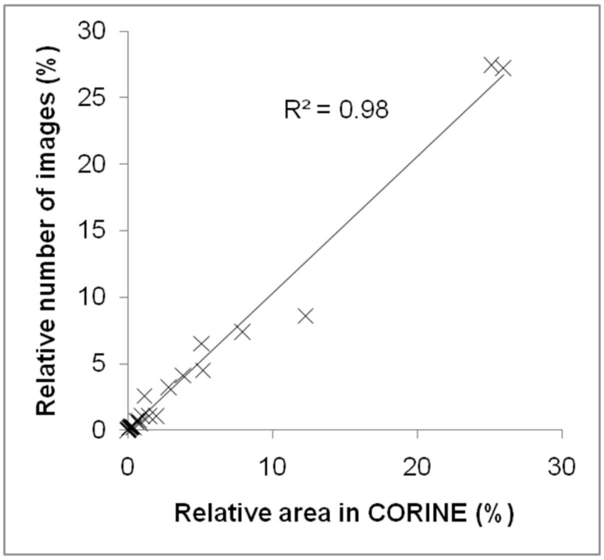

Figure 3: Scatter plot of relative area of 35 CORINE classes against number of images per CORINE class

Figure 4 shows the images rated as most and least scenic, with a clear tendency for images rated as scenic to be located in the north and west of Scotland, the Lake District and North Wales. By contrast, low rated images show a more general distribution, though with concentrations in more industrial and populated areas such as Scotland's central belt and the southeast of England.

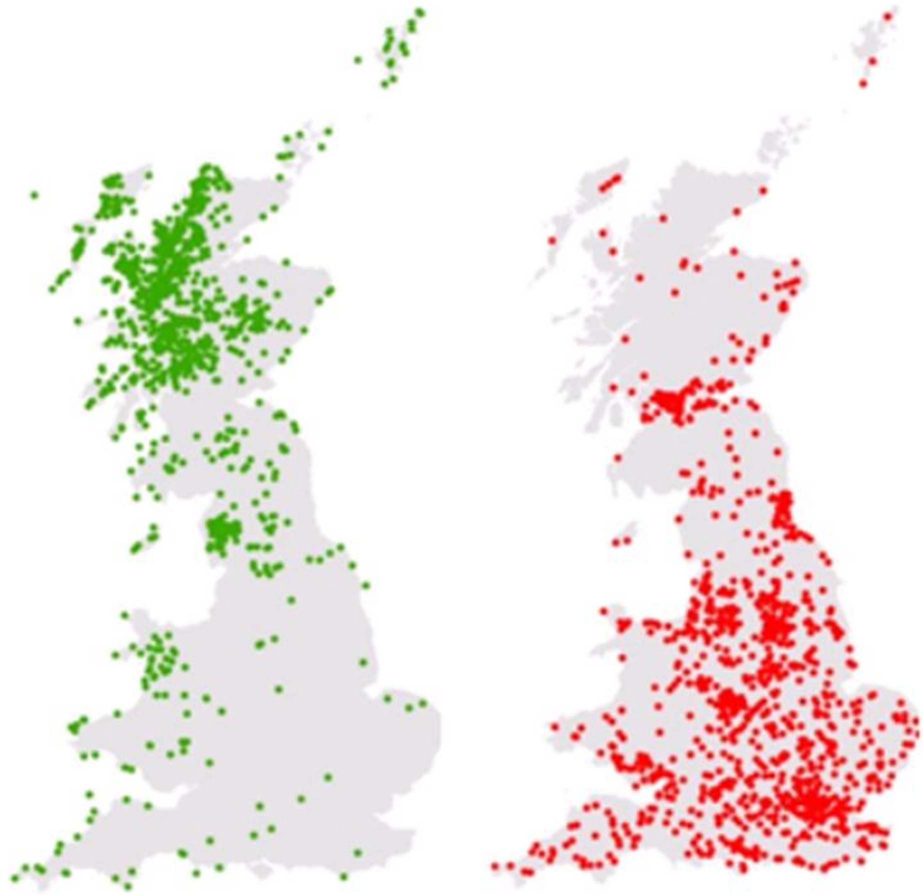

Figure 4: 1000 most images rated most (green) and 1045 images rated least (red) scenic (1045 images had the lowest possible scenicness rating)

Figure 5 illustrates the average and standard deviation of the landscape rating for all land cover classes, where the values were assigned using (d) (Variable buffer) for only the polygon found at the location of the image. Here, it is clear that, perhaps unsurprisingly, artificial surfaces have in general lower ratings than natural landscapes. Perhaps more surprising is the relatively small difference between agricultural areas and artificial surfaces.

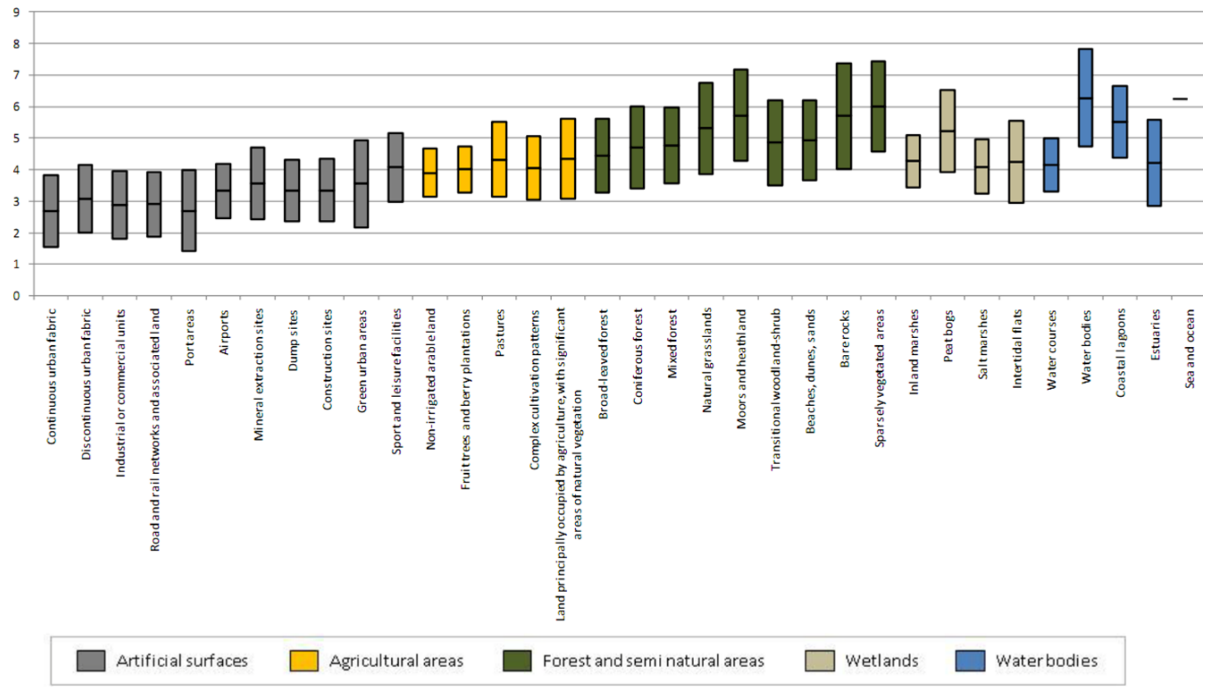

Figure 5: Mean and standard deviation in scenicness rating for CORINE classes

Table 1 shows the Spearman's correlation coefficients between land cover and landscape ratings for the six methods linking land cover to landscape rating described above, in all cases for the land cover class found at the image location, and in cases (b-f) for all land cover classes found within the area so defined.

Method (see Figure 2)	**r^2**
a (Point)	SP: 0.453
b (1km^2 grid cell)	SP: 0.459
	AP: 0.399
c (Centred 1km^2 buffer)	SP: 0.483
	AP: 0.446
d (Variable size buffer)	SP: 0.518
	AP: 0.280
e (Variable range cone)	SP: 0.451
	AP: 0.420
f (2km viewshed)	AP: 0.387

Table 1: Spearman's correlation between land cover class and average landscape rating for single polygon (SP) and all polygons (AP)

The strongest correlation, explaining a little more than 50% of the total variation in the landscape ratings, is found for the variable buffer where only the polygon found at the image location is taken account of. Interestingly, ratings considering all land cover polygons in an area show weaker correlations than those for the single polygon found at the image location. One implication may be that the land cover in the foreground of an image (so for example the road in Figure 2) has the most influence on its rating. Using more directed methods (a directed cone or a viewshed), at first glance surprisingly, does not appear to improve the correlation. This may be the result of one of two effects. The images taken in Geograph could well be taken with their general representativeness in mind, and thus participants are taking images which capture the general land cover (and not only that in a particular direction). Equally, it may be that uncertainties in the position of image taker and the precision of the CORINE data combine to make the more specific directional method more prone to errors. The result demonstrating that around 50% of the scenicness of a landscape image can be predicted using land cover data is similar to previous empirical work (e.g. Palmer, 2004) and suggests that UGC such as that produced through the ScenicOrNot and Geograph projects can be applied to explore important, and long standing, questions about landscape. In contrast to previous studies, the sample sizes reached in such work are very large, though important issues relating to bias, despite these large samples, cannot be ignored (Hollenstein and Purves, 2010). For example, it is not possible to evaluate who made the majority of ratings in ScenicOrNot, but previous works suggest a typical 80/20 rule (that is 80% of the content is provided by 20% of the users).

CONCLUSION AND FURTHER WORK

In this paper we set out to explore firstly, whether user generated content would allow us to explore the geographic distribution of scenicness in Great Britain, and secondly whether there was a relationship between ratings of landscape images and land cover data. It is clear that landscape ratings vary in space, and given the relatively even distribution of images across land cover classes, and Geograph's aim to collect geographically representative images (and thus its relatively homogenous distribution) that real geographic variations exist in landscape ratings.

These geographic variations can be related to land cover, with some 50% of the variation in landscape ratings being thus explained. Forest and semi-natural areas along with water bodies are rated as being most scenic, and ongoing work suggests that the images with larger vistas are generally considered to be more scenic. Somewhat counter intuitively, we find that linking land cover "more exactly" to image content, for example by the use of viewsheds, reduces the correlation between land cover and landscape rating, and that the strongest correlations are found for a buffer centred on the image location. This in turn suggests that, potentially, the instruction to take "geographically representative" images in Geograph has had a significant influence on the nature of the images taken.

In future work we will explore the local variability of scenicness in more depth, and attempt to compare data gathered from UGC with traditional empirical experiments where it is possible to gather more data on the reasoning behind landscape ratings.

ACKNOWLEDGEMENTS

This research would not have been possible without the efforts of those who developed Geograph and ScenicOrNot, and more importantly those who took images, rated them and make the content collected available under Creative Commons licenses. We would like to gratefully acknowledge contributors to Geograph British Isles, see http://www.geograph.org.uk/credits/, whose work is made available under the following Creative Commons Attribution-ShareAlike 2.0 Licence (http://creativecommons.org/licenses/by-sa/2.0/) and to ScenicOrNot whose work is made available under a Creative Commons Attribution-NonCommercial 3.0 Licence (http://creativecommons.org/licenses/by-nc/3.0/).

REFERENCES

Edwardes, A.J. & Purves, R.S. 2007. A theoretical grounding for semantic descriptions of place. In Proceedings of 7th Intl. Symposium on Web and Wireless Geographical Information Systems (W2GIS), 106-121.

Goodchild, M.F. 2007. "Citizens as sensors: the world of volunteered geography," GeoJournal 69, 4, 211-221.

Grêt-Regamey, A., Bishop, I. D & Bebi, P. (2007): Predicting the scenic beauty value of mapped landscape changes in a mountainous region through the use of GIS. Environment and Planning B, 50-67.

Hollenstein, L. & Purves, R.S.. 2010. Exploring place through user-generated content: Using Flickr tags to describe city cores. Journal of Spatial Information Science, 1, 21-48.

Hull, R. B & Stewart, W. P. 1992. Validity of Photo-Based Scenic Beauty Judgments. Journal of Environmental Psychology, 101-11

Hull, R. B. IV; Reveli, G. R. B. 1989. Cross-Cultural Comparison of Landscape Scenic Beauty Evaluations: A Case Study in Bali. Journal of Environmental Psychology, 177-191.

Kaplan, R. & Kaplan, S. 1989. The experience of nature: A psychological perspective. Cambridge University Press, New York. 360 pp.

Lothian, A. 2008. Scenic Perceptions of the Visual Effects of Wind Farms on South Australian Landscapes. Geographical Research, 196-207.

Meitner, M. J. & Daniel, T. C. 1997. The effects of animation and interactivity on the validity of human responses to forest data visualizations. In: Proceedings of Data Visualization '97.

Montello, D., Goodchild, M., Gottsegen, J. & Fohl, P. 2003. Where's Downtown?: Behavioral Methods for Determining Referents of Vague Spatial Queries. Spatial Cognition and Computation 3, 185–204.

Palmer, J. F. 2004. Using spatial metrics to predict scenic perception in a changing landscape: Dennis, Massachusetts. Landscape and Urban Planning, 201-218.

Ribe, R. G. 2002. Is Scenic Beauty a Proxy for Acceptable Management? The Influence of Environmental Attitudes on Landscape Perceptions. Environment and Behavior, 757-780.

Ribe, R. G. 2002. Is Scenic Beauty a Proxy for Acceptable Management? The Influence of Environmental Attitudes on Landscape Perceptions. Environment and Behavior, 757-780.

White, J.J.D & Roth, R.E. 2010 TwitterHitter: Geovisual Analytics for Harvesting Insight from Volunteered Geographic Information. In Extended Abstract Proceedings GIScience 2010.