Landscape and Urban Planning 203 (2020) 103904

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Using geographically weighted models to explore how crowdsourced landscape perceptions relate to landscape physical characteristics

Yi-Min Chang Chien*, Steve Carver, Alexis Comber

School of Geography, University of Leeds, UK

A R T I C L E I N F O

Keywords:
Landscape character assessments
Crowdsourcing
Wildness
Scenic-Or-Not

A B S T R A C T

This study explores how formal measures of landscape wildness (i.e. absence of human artefacts, perceived naturalness of land cover, remoteness from mechanised access, and ruggedness of the terrain) correlate with crowdsourced measures of landscape aesthetic quality as captured in Scenic-Or-Not data for Great Britain. It evaluates multiple linear regression (MLR) and two spatially varying coefficients models: geographically weighted regression (GWR) and multiscale geographically weighted regression (MGWR). The MLR provided a baseline model in an analysis of national data, exhibiting the presence of spatially autocorrelated residuals and suggesting that geographically weighted models may be appropriate. A standard GWR was found to exacerbate local collinearity between covariates, both overfitting and underfitting the model with highly varied and localised results. This was due to its single one-size-fits-all bandwidth and the assumption that all relationships between the target and predictor variables operate over the same spatial scale. MGWR relaxes this assumption by determining parameter-specific bandwidths, mitigating the local collinearity issues found in a standard GWR and resulting in more spatially stable and consistent coefficient estimates. The findings also indicated that the relationship between some covariates (such as remoteness) and perceived landscape quality varied little spatially, while clear gradients were found for other covariates. For example, naturalness was stronger in the north and west, ruggedness was stronger in the south and east, and the absence of human artefacts was weaker in Scotland and the north than in England and the south. Overall, the study showed that MGWR is more sensitive than GWR to the spatial heterogeneity in the statistical relationships between landscape factors and public perceptions. These findings provide nuanced understandings of how these relationships vary spatially, underscoring the value of such approaches in landscape scale analyses to support policy and planning. The discussion section of this paper considers the MGWR as the default geographically weighted model, assessing the potential for the use of crowdsourced data in landscape studies. In so doing, it illustrates how such approaches could be used to explore both subjective and objective landscape evaluations.

1. Introduction

The aesthetic quality of landscapes has a clear positive correlation with human health and well-being, and aesthetics have been recognised as a key benefit of landscapes in ecosystem service modelling (Zoderer, Tasser, Carver, & Tappeiner, 2019). However, aesthetic preferences vary widely across social and cultural contexts (Dramstad, Tveit, Fjellstad, & Fry, 2006; Zube & Pitt, 1981), making objective evaluations difficult. As a result, there is a long-standing tension between objectivist and subjectivist paradigms in landscape assessment (Daniel, 2001). At the heart of this ideological rift lies the question of whether a landscape's quality is determined by inherent physical landscape properties, or by how it is perceived (Lothian, 1999). The objectivist paradigm is based on landscape's visual properties and biophysical features, often as defined by specialists such as landscape architects. This is the most prevalent approach in formal landscape assessment practices. The subjectivist model focuses on human perceptions, opinions and preferences. However, there is a general consensus is that landscape quality is derived from the interaction between biophysical and perceived components (Daniel, 2001). Integrated approaches linking both subjectivist and objectivist considerations provide a basis for enhancing landscape planning and decision making, and an analytical framework is needed to link the two paradigms and handle discrepancies between them. However, effective landscape assessments involving both expert and non-expert perspectives also pose a challenge, as demonstrated by the landscape character assessments (LCA) (Swanwick, 2002) in the

* Corresponding author at: School of Geography, University of Leeds, LS2 9JT, United Kingdom.
E-mail addresses: gyymcc@leeds.ac.uk (Y.-M. Chang Chien), S.J.Carver@leeds.ac.uk (S. Carver), A.Comber@leeds.ac.uk (A. Comber).

https://doi.org/10.1016/j.landurbplan.2020.103904
Received 3 April 2020; Received in revised form 14 July 2020; Accepted 23 July 2020
Available online 06 August 2020

United Kingdom. This approach uses a classification system to assess and value landscapes in a two-phase process: characterisation and evaluation. Characterisation sub-divides the landscape into distinct areas based on the visual continuity of physical characteristics (such as geology, landform, and land cover), applied through the lens of spatial hierarchical mapping. Evaluation occurs through *in situ* site visits, during which landscape character descriptions are formulated qualitatively. The practice of LCA often fails in its stated aim of centring public perceptions, as both phases are typically undertaken by professionals and therefore do not capture collective or public landscape perceptions (Butler & Berglund, 2014; Conrad et al., 2011). The disconnect between public and professional perceptions in this field illustrates the need for integrated assessment frameworks, accommodating both subjectivist-based landscape evaluations (i.e. non-expert opinions) and objectivist-based ones (i.e. expert opinions).

In recent decades, the increased availability of crowdsourced geo-information offers the potential for new avenues of research to further understand links between perceptions and objective landscape measures. Such data has already been applied in numerous research areas, including place preferences (Gliozzo, Pettorelli, & Haklay, 2016), values (van Zanten et al., 2016) and perceptions (Dunkel, 2015). The Scenic-Or-Not campaign in the UK (http://scenicornot.datasciencelab.co.uk) captures public evaluations and perceptions of landscapes using photographs. Scenic-Or-Not data have been used to investigate the impact of scenic environments on human well-being (Seresinhe, Preis, & Moat, 2015) and happiness (Seresinhe, Preis, MacKerron, & Moat, 2019), enabling a clearer understanding of public perceptions regarding landscape composition and scenic beauty (Seresinhe, Preis, & Moat, 2017). The dataset is geo-referenced with national coverage, enabling spatial analyses of how public preferences and aesthetic perceptions are related to objective indicators of landscape quality.

Wilderness-related research has developed several formal methods for measuring landscape character by wilderness and wildness, and many people intuitively associate the concept of wilderness with certain aesthetic values (Carlson, 2019). The term ‘wilderness’ can be understood in multiple ways: it is partially a human construct based on romantic notions about nature and landscape, and partly an ecological reality of intact ecosystems devoid of human influence (Nash, 1982). Although there is little wilderness (in the term’s truest sense) left within Great Britain, the concept of a wilderness continuum – an idea which models anthropogenic environmental modification using inherent underlying landscape characteristics (Fritz, Carver, & See, 2000) – is still a useful tool for mapping the spectrum of relative wildness. So-called ‘wild land areas’ refer to large natural areas that are relatively undisturbed by human activity (Carver, Comber, McMorran, & Nutter, 2012). Aesthetic values, meanwhile, are more closely related to perceptions of scenic beauty. Many studies use multi-criteria approaches to capture and link the various spatial characteristics of wilderness areas. These assess wilderness quality based on four principal characteristics: absence of modern human artefacts, perceived naturalness of land cover, remoteness from mechanised access, and rugged and physically challenging nature of the terrain (Carver et al., 2012; Carver, Evans, & Fritz, 2002; Comber et al., 2010; Fritz et al., 2000). These four indicators can be used to identify landscapes that are highly valued and thought to merit conservation due to their wilderness qualities. It is unclear whether these formal wildness measures could contribute to landscape aesthetic assessments, and to what extent these indicators are associated with the public’s landscape preferences. Nonetheless, such approaches have been adopted by the United States National Park Service to model, map and monitor variations in wilderness character (Carver, Tricker, & Landres, 2013).

Previous studies that have examined the relationships between measures of landscape values or qualities and features (topography, land cover, etc.) have typically applied global statistical models. In these models, the relationships between input variables are assumed to be spatially invariable (Makeschin, Koschke, Frank, Fürst, & Witt, 2013; Schirpke, Tasser, & Tappeiner, 2013; van Zanten et al., 2016). However, the relationships between landscape-related predictor and response variables may vary in different locations (i.e. exhibit process spatial heterogeneity). Spatially-varying coefficient models such as geographically weighted regression (GWR) can be used to identify and explore these relationships, supporting an enhanced understanding of geographical processes (Brunsdon, Fotheringham, & Charlton, 1996). GWR uses a moving kernel to generate subsets of the data from which local regression models are determined. It has been applied in several landscape studies to understand local processes (Hong & Jeon, 2017; Luo & Wei, 2009; Su, Foody, & Cheng, 2012; Sun, Xie, & Chen, 2018). The critical consideration in any GWR analysis is the specification of the kernel size, or bandwidth. This determines the number of observations that are included in each local subset, thus establishing the degree of spatial smoothing in the model’s outputs. GWR bandwidths can be implemented at a fixed or an adaptive distance (where adaptive includes the same number of observations in each subset). They are optimally determined using some measure of model fit such as Akaike Information Criterion (Akaike, 1973) or leave-one-out cross-validation (Bowman, 1984; Brunsdon et al., 1996; Cleveland, 1979). Although a standard GWR can capture process and relationship heterogeneity, its single kernel size assumes that each response-to-predictor relationship operates over the same spatial scale. Multiscale geographically weighted regression (MGWR) relaxes this assumption and identifies the individual scale at which each response-to-predictor relationship operates (Fotheringham, Yang, & Kang, 2017; Yang, 2014), thus elucidating geographic processes.

This study explores how measures of wildness (Carver et al., 2012) correlate with crowdsourced perceptions of landscape aesthetics from Scenic-or-Not using both non-spatial and spatial statistical models. The aim is to better understand the relationship between objective and subjective measures of landscape quality – with particular attention to variations across space and spatial scale – to develop a more holistic model for landscape character assessments. To this end, bivariate correlations were initially evaluated, and the global relationships were examined through multiple linear regression (MLR). A GWR was then applied to examine spatial non-stationarity in the relationships. The analysis was refined by applying an MGWR to examine the differing scales of the relationships.

2. Data and methods

2.1. Scenic-Or-Not data (response variable)

The Scenic-Or-Not data are freely available. At the time of writing, the dataset includes 212,212 images covering nearly 80% of the Ordnance Survey (OS) 1 km^2 grid squares of Great Britain. Each grid square contains at least three ratings. The dataset uses Geograph geo-referenced photographs taken and uploaded by members of the public. Scenic-Or-Not participants are presented with randomly selected photographs and are invited to rate each one on a scale of 1–10, wherein 1 is the least scenic and 10 is the most scenic. The mean scenic rating, which captured an average measure of public perceptions of landscape scenic beauty, was used as the response variable in the scenic quality regression models of this study. However, these methodologies feature some limitations: in most cases, landscape visual aesthetic quality or preference values were given for a single photograph, which was assumed to capture the local landscape characteristics present in a 1 km^2 region. The mechanism of representative image selection for each grid cell in Scenic-Or-Not is unclear, and visual inspection of some photographs reveals potential sources of bias in subject choice and framing. For example, a focus on a barn in the composition of a rural landscape photograph for aesthetic effect may misrepresent the local landscape. Such biases illustrate the problem of the uncertain reliability and quality of crowdsourced datasets (Comber, Mooney, Purves, Rocchini, & Walz, 2016; Oteros-Rozas, Martín-López, Fagerholm, Bieling, &

Plieninger, 2018). Additionally, the image locations reported in the Scenic-Or-Not dataset may vary by 100 m from those reported in Geograph, and some Scenic-Or-Not images may have been removed from the Geograph repository altogether. Thus, the measures captured via Scenic-Or-Not may be representative of the landscape scenic quality of a broader area with better accuracy.

2.2. Wildness components (predictor variables)

Formal measures of wildness quality, as described in full by (Carver et al., 2012) in the context of Scotland and later extended across the United Kingdom, were used as explanatory variables of landscape aesthetic quality. Overall, wildness quality can be defined by four attributes: absence of modern human artefacts, perceived naturalness of land cover, remoteness from mechanised access, and rugged and challenging terrain. These were calculated over a 25 m grid and summarised below:

- Absence of modern human artefacts (absence):

This indicator measures the visual absence of man-made structures in a 360-degree arc at a given location. Structures were extracted from OS MasterMap data and included linear features (e.g. railways and roads), non-natural vegetation (e.g. hard-edged plantation forestry), built features (e.g. buildings and structures), engineering structures (e.g. pylons and hydro-electric/reservoir drawdown lines), and novel industrial features (e.g. wind turbines). The absence measure at each location was derived from the proportions of these structures within the 360-degree field of view (FOV) in a GIS-viewshed. The cumulative percentage of the view that was obstructed by man-made features based on the horizontally and vertically visible proportions of the features was calculated over a digital surface model (DSM). This voxel viewshed approach accounts for the effects of visual distance decay and relative size (Carver & Washtell, 2012).

- Perceived naturalness of land cover (naturalness):

The evaluation of naturalness was based on a reclassification of the Land Cover Map 2007 (LCM2007) (Morton et al., 2014), using ancillary forest data from the National Forest Inventory (https://www.gov.uk/guidance/access-forestry-commission-datasets). Each LCM2007 class was allocated a naturalness score of 0–5 based on its level of human intervention (see Table 1). These allocations were visually checked against aerial photography and local knowledge to identify any inconsistencies. The area weighted mean naturalness score was calculated within a 250-metre radius for each grid cell.

- Remoteness from mechanised access (remoteness):

Remoteness refers to the time needed to walk to a destination from the nearest road access. This measurement accounts for the effects of distance, relative gradient, ground cover, and barrier features such as open water and steep terrain. It is essentially an adaption of Naismith's rule (Naismith, 1892) which allocates 15 min of walking time for 1 km on horizontal surfaces, plus 10 min for every 100 m of ascent. The rule includes an assumed speed of 5 km per hour over flat terrain (i.e. slopes between 0° and 5°) and corrections for the slope and angle at which the terrain is crossed. For example, it features penalties of 30 min for every 300 m of ascent and 10 min for every 300 m of descent on slopes greater than 12°. Table 2 details the derivation of the factors that were used to generate the cumulative cost surface.

- Rugged and physically challenging nature of the terrain (ruggedness):

This indicator was devised to capture physical variations in terrain morphology, as well as weather conditions caused by the nature of the terrain (in cases where the challenging weather at high altitudes can influence human perceptions). The OS landform profile 10-metre digital elevation model (DEM) was used to initially derive indices of terrain complexity that account for gradient, aspect and relative relief. Ruggedness was calculated from 2 standard deviations of terrain curvature within a 250-metre radius of the target cell, combined by linear summation with altitude from the DEM, to reflect the weather conditions at higher locations with lower temperatures and greater wind speeds.

Hereafter, the response and the explanatory covariates are referred to simply as 'scenicness', absence, naturalness, remoteness and ruggedness.

Table 1
Land cover naturalness scores, adapted with permission from (Carver et al., 2012).

LCM2007 class	Naturalness score
Broad-leaved woodland: semi-natural	5
Broad-leaved woodland: mixed	4
Broad-leaved woodland: planted	3
Coniferous woodland: semi-natural	5
Coniferous woodland: mixed	4
Coniferous woodland: planted	3
Arable and horticultural	2
Improved grass	2
Neutral grass	3
Calcareous grass	3
Acid grass	4
Bracken	4
Dwarf shrub heath	4
Bog	5
Inland water: natural	5
Inland water: raised	4
Inland water: impounded	3
Montane habitats	5
Inland rock	5
Built up areas	0
Supra littoral rock	5
Supra littoral sediment	5
Littoral rock	5
Littoral sediment	5
Saltmarsh	4
Sea/Estuary	5

2.3. Sampling scheme

To overcome potential sampling bias, the Scenic-Or-Not data were aggregated over 5 km hexagonal grid cells. Hexagonal grids enable the exploration of more subtle spatial patterns than square grids due to their more consistent connectivity (Wang, Ai, Shen, & Li, 2020).

The median values of both response and explanatory variables within the cells were determined for each of the 11,786 grid cells. Fig. 1 shows the spatial pattern of the aggregated data for the scenicness response and the standardised covariates.

2.4. Data analysis

A multiple linear regression (MLR) model was constructed to model the relationships between the predictor and target variables as follows:

$$y_i = \beta_0 + \sum_{j=1}^{m} \beta_j x_{ij} + \varepsilon_i \tag{1}$$

where for observations indexed by $i = 1, \cdots, n$, y_i is the target variable, x_{ij} is the value of the j^{th} predictor variable, m is the number of predictor variables, β_0 is the intercept term, β_j is the regression coefficient for the j^{th} predictor variable and ϵ_i is the random error term. The coefficients β_j are commonly estimated by the ordinary least squares

Table 2
The calculations of walking time for the remoteness indicator.

	Data source	Specific type	Speed (km/h)	Cost (second)	Criteria
Ground cover influence	LCM2007	Heather and forest	3	$T = 1.2 * \Delta S$	self-defined
	OS MasterMap	Bog	2	$T = 1.8 * \Delta S$	
		Other types	5	$T = 0.72 * \Delta S$	
		Crossable rivers	0.03	$T = 120 * \Delta S$	
		Roads and tracks	15	$T = 0.24 * \Delta S$	
Gradient influence	DEM	Uphill (slope $> 0^{\circ}$)	+ 10mins/100 m of ascent	$T = a * \Delta S + 6 * \Delta H$	Naismith's rule
		Slight downhill ($-5^{\circ} <$ slope $< 0^{\circ}$)	5	$T = a * \Delta S$	Langmuir's correction
		Moderate downhill ($-12^{\circ} <$ slope $< -5^{\circ}$)	− 10min/300 m of descent	$T = a * \Delta S + 2 * \Delta H$	
		Steep downhill (slope $< -12^{\circ}$)	+ 10min/300 m of descent	$T = a * \Delta S - 2 * \Delta H$	
Barrier influence	OS MasterMap	Unfordable rivers (i.e. polygons)			self-defined

where T is time in second.
ΔS is the horizontal cell distance/resolution in metres.
ΔH is the vertical elevation difference between cells in metres.
a is the horizontal cost factor according to different land cover types.

(OLS) method. A MLR model frequently suffers from two commonly observed effects in spatial data: spatial autocorrelation of observation and process spatial heterogeneity (Anselin, 2010). To overcome these effects, a GWR can be applied (Brunsdon et al., 1996). A GWR is similar to a linear regression, except that it calculates a series of local linear regressions rather than a global one. It uses data falling within a moving window or kernel at a series of discrete locations, such as grid cells. In this process, it gathers data from nearby locations and thereby generates local and spatially varying coefficient estimates. A GWR model has locations associated with the coefficient terms and can be expressed as:

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^{m} \beta_j(u_i, v_i)x_{ij} + \varepsilon_i \tag{2}$$

where (u_i, v_i) is the spatial location of the i^{th}observation and $\beta_j(u_i, v_i)$ is a realization of the continuous function $\beta_j(u, v)$ at point$i$. As with the linear regression model, the set of ε_i obeys an independent normal distribution with a zero mean and common variance σ^2.

Critical to any GWR is the specification of the kernel, which selects and weights data to be used in each local model. This geographical weighting process produces data nearer to the kernel's centre, making a greater contribution to the estimation of local regression coefficients at each local regression calibration point. The bandwidth can either be specified as a constant (fixed) distance value or as an adaptive one, in which the number of nearest neighbours is fixed. In this study, a Gaussian kernel was used to determine the optimal fixed bandwidth.

However, a uniform bandwidth specified in a standard GWR may be inappropriate in situations in which different predictor variables operate over different spatial scales and, therefore, have unique spatial relationships with the target variable (Fotheringham et al., 2017; Yang, 2014). A standard GWR, as previously outlined, ignores these differences and identifies a best-on-average scale of relationship non-stationarity for a single kernel bandwidth. This approach may be limited because it implicitly assumes the same spatial scale for each predictor, and these scales may be incorrect. To rectify this problem, a mixed (or semiparametric) GWR (MX-GWR) can be applied (Brunsdon, Fotheringham, & Charlton, 1999; Mei, Xu, & Wang, 2016), in which some relationships are assumed to be stationary (i.e. globally fixed as in a standard OLS), whereas others are assumed to be non-stationary (i.e. locally varied as in a standard GWR). However, a mixed GWR only partially addresses the problem, as locally-varying relationships are assumed to operate at one of two spatial scales. Consequently, a multiscale GWR was proposed by (Fotheringham et al., 2017; Yang, 2014). In a MGWR model, an individual bandwidth is determined for each predictor variable. This allows the scale of relationship non-stationarity

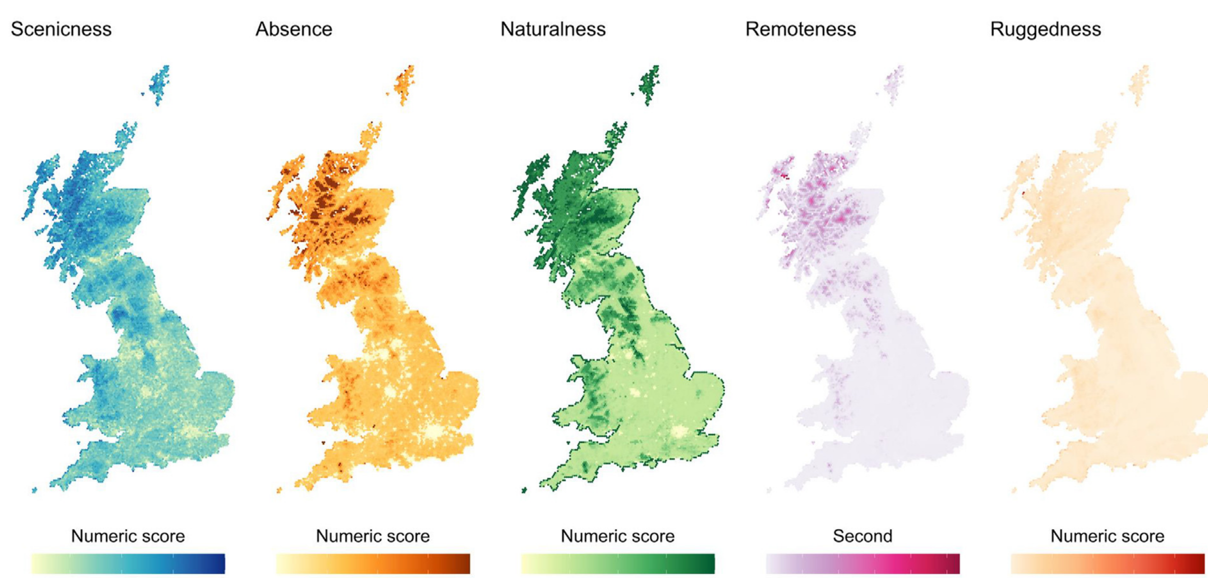

Fig. 1. The standardised Scenic-Or-Not ratings (scenicness) and the four wilderness components (i.e. absence, naturalness, remoteness and ruggedness) for Great Britain aggregated over a hexagonal grid with a cell width of 5 km.

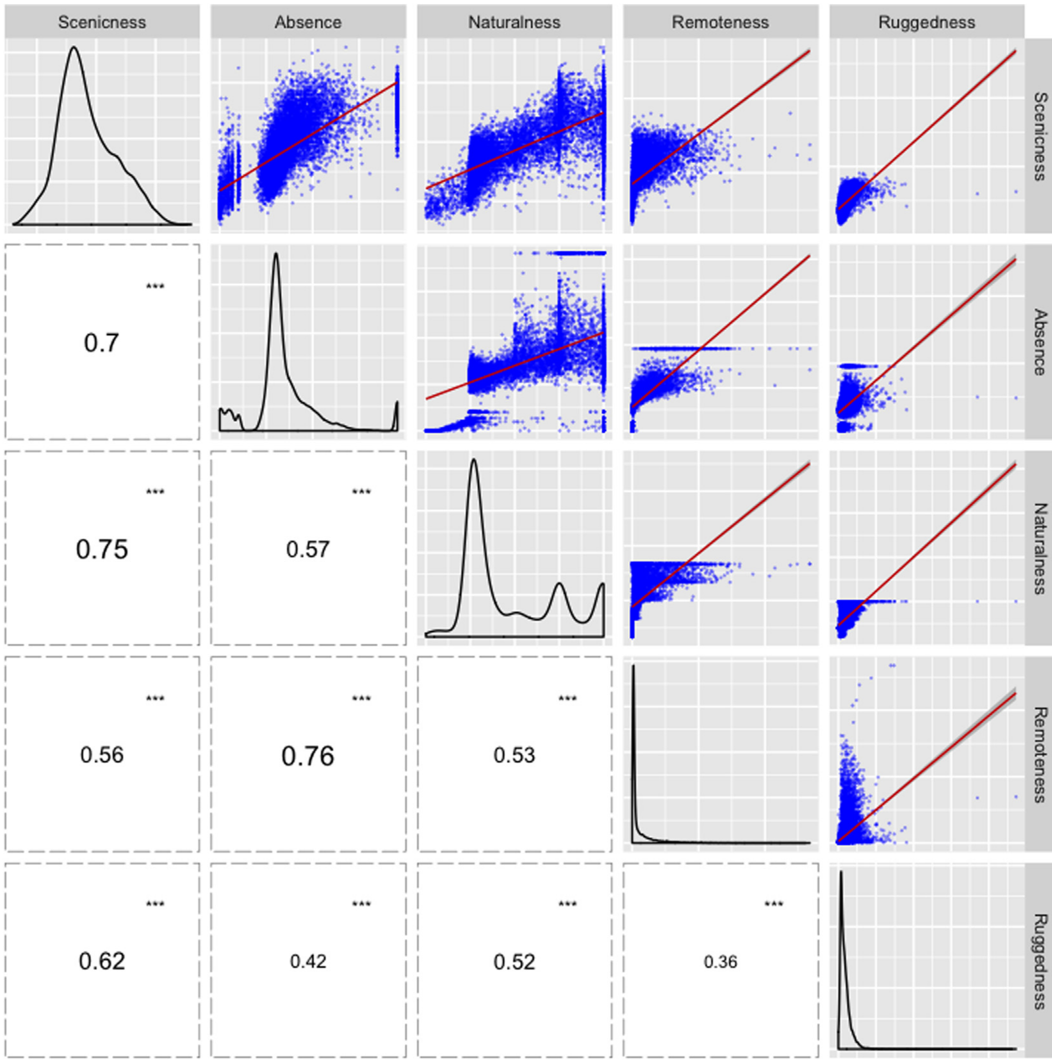

Fig. 2. Pearson pairwise correlation, scatterplots and distributions of the input data (significance indicated by *** < 0.001, ** < 0.01, * < 0.05).

to vary for each target-to-predictor variable relationship, as described in Equation (3):

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^{m} \beta_{bwj}(u_i, v_i)x_{ij} + \varepsilon_i \tag{3}$$

where bwj in β_{bwj} indicates the bandwidth used to calibrate the j^{th} conditional relationship. The MGWR model calibration uses an iterative back-fitting procedure; thus, the computational overheads are high when handling a large number of observations (Oshan, Li, Kang, Wolf, & Fotheringham, 2019).

3. Results

3.1. Exploratory analysis

The pairwise Pearson correlation analysis is shown in Fig. 2. It reveals significant positive associations between each wildness component and scenicness. Naturalness has the highest correlation ($\gamma = 0.75$, $p < 0.001$), and the scatter plot shows that the association approximates to a linear relationship. Similar values were found for absence ($\gamma = 0.7$, $p < 0.001$), ruggedness ($\gamma = 0.62$, $p < 0.001$), and remoteness ($\gamma = 0.56$, $p < 0.001$). There is little evidence of bivariate correlation among explanatory variables except for that between absence and remoteness ($\gamma = 0.76$, $p < 0.001$). This correlation is plausible; a lack of intervening man-made features is likely to be confounded by inaccessibility. Hence, two multiple regression analyses were used to deduce whether remoteness acted as a confounder, coupled with the diagnostics of collinearity. Variable collinearity may have adverse effects on the estimation of MLR coefficients (O'Brien, 2007). Local collinearity may be found in local data subsets in a GWR, even when not observed globally (Wheeler & Tiefelsdorf, 2005). However, more recent research has suggested that collinearity is unproblematic where the correlation is < 0.8 or > -0.8 (Comber & Harris, 2018). The robustness of GWR to the effects of multicollinearity has been also demonstrated, particularly with a large sample size (A Stewart Fotheringham & Oshan, 2016; Páez, Farber, & Wheeler, 2011).

3.2. Multiple linear regression

Two MLR models of scenicness were fitted, one with remoteness and one without. The inclusion of remoteness mildly influenced the coefficient estimates of the other predictors (Table 3). The sign of the coefficient estimate for remoteness was negative, contradicting the positive correlation reported in the previous section but indicating interaction amongst predictors. The variance inflation factor (VIF) diagnostics for each predictor confirmed the lack of collinearity in both models with all VIFs values below 10 (Belsley, Kuh, & Welsch, 1980). A marginally improved model fit with all covariates was found, as indicated by the adjusted R-squared and corrected Akaike information criterion (AICc)

Table 3
The coefficient estimates and associated p-values of the MLRs with and without remoteness.

Variable	MLR without Remoteness					MLR with Remoteness				
	Coefficient Estimate	Standard Error	t-value	p-value	VIF	Coefficient Estimate	Standard Error	t-value	p-value	VIF
Intercept	4.606	0.006	778.220	0.000	–	4.606	0.006	779.099	0.000	–
Absence	0.421	0.007	57.440	0.000	1.533	0.454	0.010	47.222	0.000	2.641
Naturalness	0.489	0.008	62.630	0.000	1.741	0.496	0.008	62.732	0.000	1.787
Remoteness	–	–	–	–	–	−0.048	0.009	−5.261	0.000	2.415
Ruggedness	0.303	0.007	42.890	0.000	1.423	0.303	0.007	42.942	0.000	1.423
	Adjusted R^2 = 0.709, AICc = 23,027					Adjusted R^2 = 0.710, AICc = 23,001				

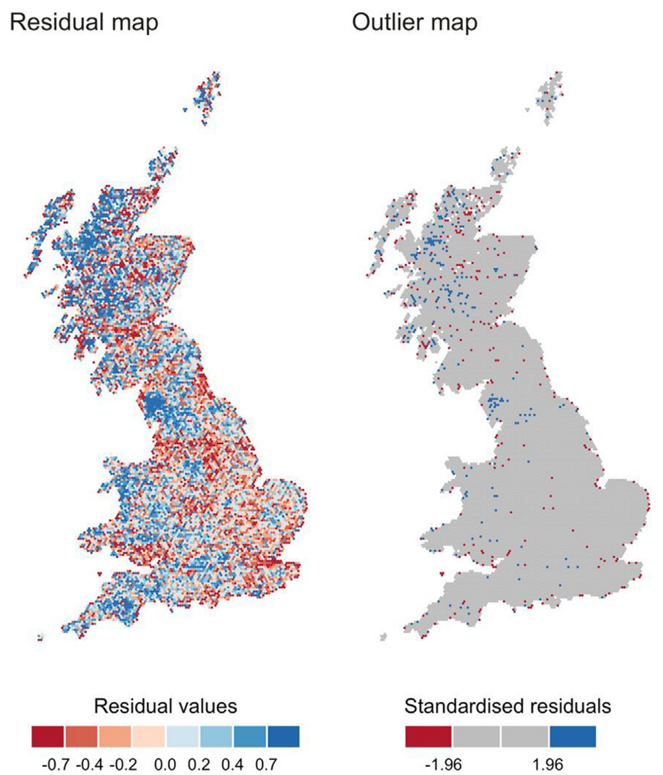

Fig. 3. The quantile-classified residual map (left) and the outlier map (right) highlights areas where the global model overestimated (red) and underestimated (blue) landscape scenic beauty.

values (see Table 3). The model had an adjusted R-squared of 0.71, suggesting that the 71% variation in public scenic ratings can be explained by them. The coefficient estimates in Table 3 indicate that all covariates are significantly associated with scenicness. Absence, naturalness, and ruggedness exhibited significantly positive relationships with scenicness, while remoteness exhibited a negative one. However, the MLR coefficient estimates should be interpreted with caution as the model residuals were found to be spatially autocorrelated (Moran's I = 0.267, $p < 0.001$; Jarque-Bera statistic = 15074, p-value < 0.001). The map of residuals (Fig. 3) highlights areas where the global model overestimated (red) and underestimated (blue) landscape scenic beauty, showing some evidence of clustering (and, therefore, spatial autocorrelation). The overpredictions tended to occur in urbanised regions, including major cities in England, Wales and Scotland, whilst the underpredictions emerged predominantly in rural regions. The map of outliers (i.e. where t-values are greater than +1.96 or less than −1.96) (Fig. 3) indicates that negative outliers were largely found along the coastline. Positive ones were clustered around the Lake District and the Northwest Highlands, both of which are scenic mountainous landscapes with high cultural value. A plausible explanation could be that cultural and topographical characteristics not captured by the covariates (e.g. agro-pastoral scenery and terrain openness) may positively influence perceptions of aesthetic value in these areas. The Koenker's studentised Breusch-Pagan statistic was used to further determine if there was a non-constant variance in the residuals. It was found to be statistically significant (BP = 2337.7, df = 4, p-value < 0.001), indicating that the relationships between some or all of the predictors and the response were non-stationary. This finding emphasizes the need for methodologies such as the GWR and MGWR, which can explore spatial heterogeneity in data relationships and account for the spatial autocorrelation of the input variables. The following analyses and comparisons were undertaken using all four covariates.

3.3. Standard GWR and multiscale GWR

As collinearity may be present in local subsets under the GW framework (Wheeler & Tiefelsdorf, 2005) despite a global absence, the GWR and MGWR analyses were coupled with the local collinearity diagnostic tests using the *mgwr* Python package (Oshan et al., 2019). Fig. 4 shows the variability of the local condition numbers (CN) for both the GWR and the MGWR models. In the GWR model, some areas (predominantly in Southern England) were highly affected by collinearity, with many areas having a CN greater than 30. These numbers

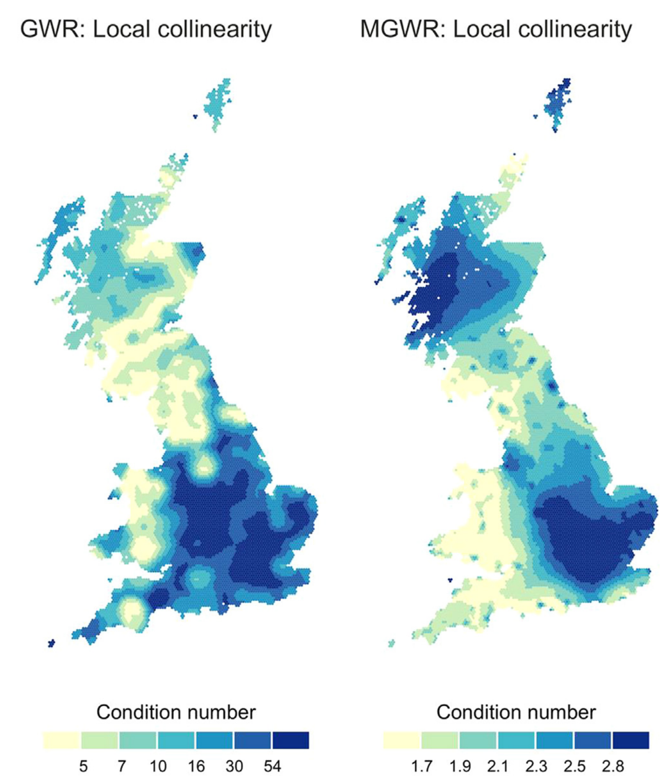

Fig. 4. The diagnostic tests of the local collinearity for the GWR (left) and the MGWR (right) models using quantile breaks.

Table 4
The coefficient estimates arising from the GWR and MGWR models (1Q = 1st quartile, Med = median, 3Q = 3rd quartile, IQR = interquartile range).

Parameter	GWR 1Q	Bandwidth: 15.2 km Med	3Q	IQR	MGWR Bandwidth (km)	1Q	Med	3Q	IQR
Intercept	4.636	4.821	5.103	0.467	5.7	4.440	4.648	4.956	0.516
Absence	0.148	0.387	0.547	0.399	32.9	0.151	0.326	0.486	0.335
Naturalness	0.217	0.353	0.504	0.287	118.6	0.308	0.336	0.355	0.047
Remoteness	−0.090	0.086	0.546	0.636	1944.2	0.035	0.035	0.035	0.000
Ruggedness	0.264	0.429	0.628	0.364	48.7	0.217	0.325	0.444	0.227
	GWR: adjusted R^2 = 0.818, AICc = 18,430;					MGWR: adjusted R^2 = 0.831, AICc = 18,313			

are indicative of significant collinearity amongst the predictor variables (Belsley et al., 1980; Gollini, Lu, Charlton, Brunsdon, & Harris, 2015). This collinearity may be caused by the single GWR bandwidth, which can increase collinearity between variables (Oshan & Fotheringham, 2018). All of the local MGWR models were found to have CNs of less than 3.

Bandwidth selections for both the GWR and MGWR models were optimised using a cross-validation approach under a Gaussian weighting kernel. Table 4 summarises the spatial distribution and variation of the coefficient estimates from the two analyses, along with the MGWR bandwidths. The GWR and MGWR improve the fit as expected (GWR: adjusted R^2 = 0.818; MGWR: adjusted R^2 = 0.831) over the MLR (adjusted R^2 = 0.710). However, it would be unwise to compare the three models by their adjusted R^2 only. Cross-model fits can be compared more effectively using specific information criteria such as the AICc, which accounts for both model parsimony and prediction accuracy. Large improvements (decreases) in the AICc fit were found using GWR and MGWR models (AICc = 18,430 and 18,313 respectively) than that found using a MLR model (AICc = 23,001). Overall, the GWR coefficient estimates show a higher variation than the MGWR ones – as indicated by the interquartile range (IQR) – except for the intercept. The low variation of the intercept could be caused by the single average bandwidth of the GWR model, which is narrower than the bespoke bandwidth for the individual predictor but wider than the bandwidth for the intercept from the MGWR.

Figs. 5 and 6 show the mapped GWR and MGWR coefficient estimates for the intercept and each covariate along with their statistical significance (i.e. t-values over 1.96 or below −1.96), as indicated by the grid outlines, creating darker areas on the maps. Comparisons of coefficient surfaces can deepen understandings of spatial and scale variations. Some marked differences between the standard GWR and MGWR models are present. First and foremost, all of the covariate coefficient estimates in the GWR model inflect from negative (red) to positive (blue), indicating both negative and positive associations with scenicness. Nearly all of the coefficient estimates in the MGWR model are positive, with some highly localised negative values for absence (highlighting the limitations of a standard GWR with a 15.2 km bandwidth, which may misrepresent parameter-specific relationship scales). This is confirmed by the MGWR bandwidths of 32.9 km for absence, 118.6 km for naturalness, 1944.2 km for remoteness, and 48.7 km for ruggedness. Similarly, the GWR model has the largest variation in coefficient estimates for remoteness (IQR = 0.704), with its effects changing in sign for England in particular but with little significance. The MGWR output for remoteness shows limited variation, indicating a largely stationary process. This stationary quality is reflected by its wide bandwidth; it has a weak relationship with scenicness compared to the other covariates. This weak correlation is plausible given that remoteness is mainly concerned with landscape accessibility. While accessibility is essential for stimulating people's perceptions of a landscape, it does not necessarily contribute to an area's scenic attractiveness.

The MGWR bandwidths for the intercept and the other covariates indicate their degree of localness in their relationships with perceived landscape scenic beauty. The intercept operates at a highly localised scale of 5.7 km, with a similar spatial pattern to that observed in the map of MLR residuals (Fig. 3). This suggests that much of the residual autocorrelation may have been captured by the locally varying intercepts which could help guide further data acquisition and analysis. The MGWR coefficient estimates for absence are similar to the GWR estimates because the MGWR bandwidth of 32.9 km is broadly similar to the GWR bandwidth of 15.2 km. The difference between the GWR and MGWR is in the significance of those relationships; however, a greater

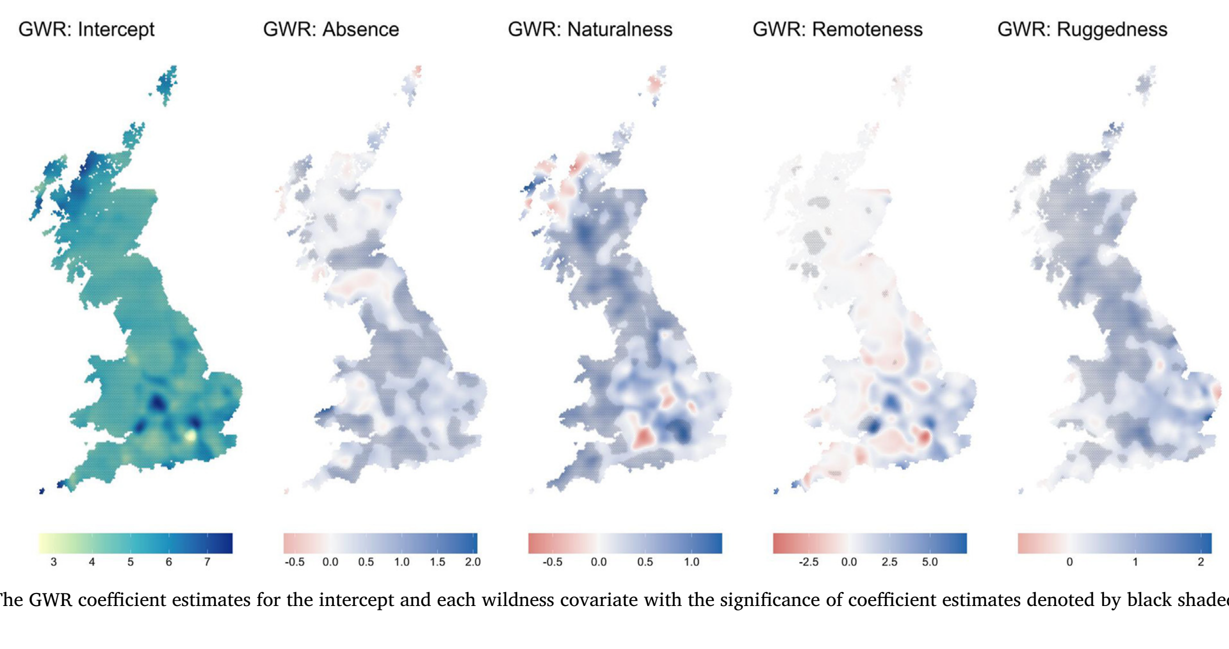

Fig. 5. The GWR coefficient estimates for the intercept and each wildness covariate with the significance of coefficient estimates denoted by black shaded outlines.

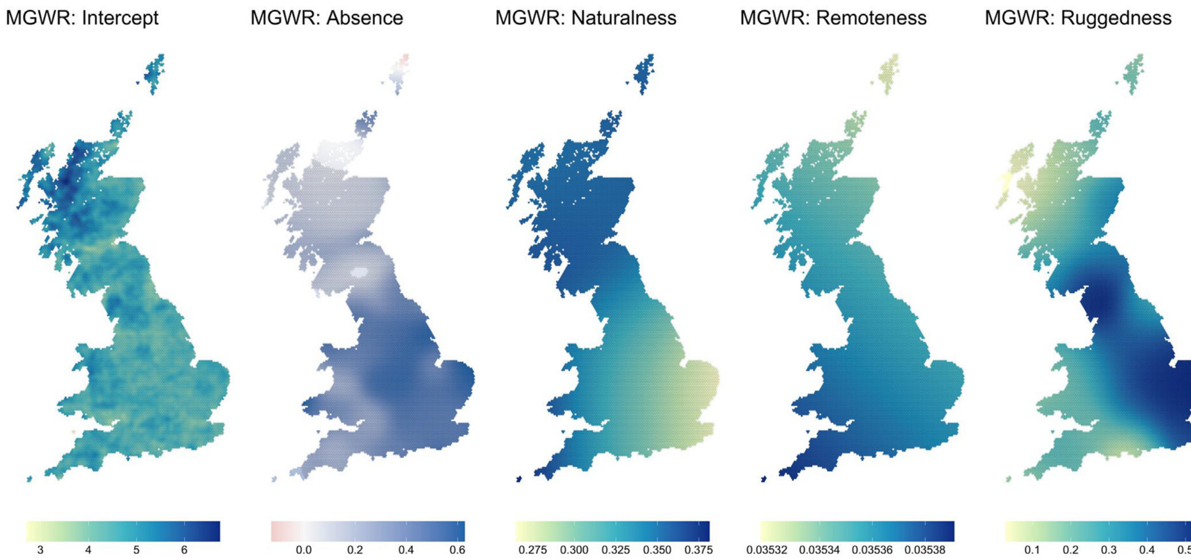

Fig. 6. The MGWR coefficient estimates for the intercept and each wildness covariate with the significance of coefficient estimates denoted by black shaded outlines.

number of locations have significant coefficient estimates obtained from the MGWR calibration.

The MGWR results shown in Table 4 demonstrate that absence has a relatively strong relationship with scenicness (a median coefficient estimate of 0.326). However, this relationship was somewhat localised; it occurred with a MGWR bandwidth of 32.9 km and considerable local variation, as shown by the IQR of the local coefficient estimates (0.335). Naturalness has a similar median coefficient value (0.336) and a wider bandwidth (118.6 km). However, it also has a low IQR (0.047), indicating weak spatial variation and overall tendencies towards a global trend. The coefficient estimates for ruggedness has a median value (0.325), a moderate IQR (0.227), and a localised bandwidth (48.7 km), indicating that the relationship between this variable and scenicness varies locally within the study area. The maps in Fig. 6 illustrate the spatial variation of the coefficient estimates derived from the MGWR calibration. The MGWR coefficient estimates for naturalness show a clear pattern, with a strongly positive effect in Scotland, suggesting that naturalness may be of particular importance in areas that are widely renowned for their natural beauty. Comparatively, a decline in East of England suggests that public perceptions of scenic beauty in England may be context-dependent – what is perceived as naturalness in an urban setting might not be seen as such in a more natural context. Likewise, there are clear differences from west to east in Wales. The MGWR ruggedness coefficient estimates highlight two areas with high values: the Lake District, which comprises many areas with rugged characteristics, and East of England, which does not. In some of the most rugged landscapes, such as the Northwest Highlands, the association was weakly positive. This also suggests that the effects of ruggedness on landscape scenic beauty are relative and context-dependent.

4. Discussion

4.1. Model estimation

In this study, a MLR was fitted as a baseline model after confirming that the variable collinearity was not an issue globally. The MLR model did not account for spatial context and its residuals exhibited autocorrelation, emphasizing the applicability of spatially varying coefficient models such as GWR. Consequently, a standard GWR was used to explore the local variations of the relationships between the response and predictor variables under a single kernel bandwidth, which resulted in significant levels of local variable collinearity (T. Oshan et al., 2019). The MGWR analysis, incorporating variable-specific bandwidths, was found to eliminate local collinearity with a greater number of locations at which the covariates were found to be significant. MGWR has thus been advanced as the default geographically weighted model (Comber et al., 2020; Fotheringham et al., 2017; Lu, Brunsdon, Charlton, & Harris, 2017; Murakami et al., 2018; Wolf, Oshan, & Fotheringham, 2018) as it makes fewer assumptions about the spatial scales of processes related to individual covariates, reducing susceptibility to collinearity.

Of the MGWR estimates, absence has a weaker relationship with scenicness in Scotland than in England, whereas naturalness showed strong to weak gradients running north to south and west to east. Absence has a stronger relationship with scenicness in parts of the Midlands, East of England and Southwest Wales, with the remainder of Great Britain either weakly positive or largely non-existent, particularly Scotland (the landscape with the fewest human modifications). Yet, there are clear exceptions to this pattern. One such exception was Scotland's Central Lowlands – where the country's largest cities (i.e. Edinburgh and Glasgow) are located – and the Orkney Islands. Remoteness was found to have a weak relationship with scenicness and varied little, and ruggedness was a stronger predictor of scenic beauty to the south and east – almost the inverse of naturalness. These results suggest that, aside from remoteness, the factors associated with crowdsourced measures of landscape aesthetic quality vary by location and the local landscape contexts. In areas with high urban density, ruggedness and the absence of human of artefacts have a greater impact on public landscape preferences. Perceived naturalness, by contrast, was more strongly associated with scenic beauty in areas with a sparser population and fewer urban centres. While recognizing that the wildness covariates may not fully capture landscape aesthetic values (for example, by failing to capture the cultural aspects of landscapes) (Tieskens, Van Zanten, Schulp, & Verburg, 2018) these findings highlight strategies for future landscape enhancement and conservation throughout the United Kingdom.

4.2. Limitations and future research

This analysis used data aggregated to 5-km hexagonal grid cells. All analyses of spatial data are subject to the modifiable areal unit problem (MAUP) (Openshaw, 1984a, 1984b). In brief, the MAUP posits that statistical distributions, relationships and trends exhibit widely different properties when the same data are aggregated or combined over various reporting units at different spatial scales. It describes the process of distortion in calculations and differences in outcomes due to aggregation (the scale effect), as well as the configuration of the zoning system (the zoning effect) (Fotheringham & Wong, 1991). Future work will examine the effects on the findings of different scales of aggregation and zonings, particularly in the context of determining optimal MGWR bandwidth and the process scales they suggest.

A further limitation relates to the opinions captured in the Scenic-Or-Not dataset. Each image in the Scenic-Or-Not database has at least three ratings, but nothing is known about the demography of the contributors. It is well known, however, that different groups interpret landscapes in different ways (Comber et al., 2016) and that these interpretations may or may not be representative of general public opinion (Oteros-Rozas et al., 2018). The Scenic-Or-Not data may represent a biased sample of landscape aesthetics preferences. Additionally, the motivations of contributors for their scores were unknown. Finally, the use of photographs as a proxy for the in-person experience of a landscape may cause bias associated with aesthetic considerations or framing. Perceptions of an online photograph do not always relate to *in situ* direct observations and perceptions (Palmer & Hoffman, 2001).

This work showed how spatially explicit approaches such as MGWR support enhanced understandings of the relationships between landscape covariates and public landscape preferences. Such methods (including the use of crowd-sourced data, such as the dataset provided by Scenic-Or-Not), can be effective exploratory tools for spatially unpacking socio-environmental relationships. These methods offer a bridge between subjectivist and objectivist paradigms in support of local planning. Landscape planners and practitioners might benefit from using this technique to facilitate targeted management, thus conserving valuable landscape characteristics and features. The identification of spatially varying relationships can also be used to guide further data acquisition and analysis, augmenting the development of more informed landscape policies. This supports integrated mapping approaches for incorporating data from perception-based surveys. By supplementing inputs into current LCA evaluations and complementing current conceptual frameworks for CES (Kerebel, Gélinas, Déry, Voigt, & Munson, 2019), such efforts sensitively connect both the human and the natural components of landscapes.

5. Conclusions

This study explored the relationships between crowdsourced measures of perceived landscape scenic beauty as captured in the Scenic-Or-Not dataset (scenicness), alongside components of formal landscape wildness (i.e. absence of human artefacts, perceived naturalness of land cover, remoteness from mechanised access and rugged and challenging terrain). It used both non-spatial (standard regression) and spatial regression (GWR and MGWR) models. The results of this analysis illustrate the limitations of a standard GWR, which is liable to overfit some variables while underfitting others. The variable-specific, or bespoke, bandwidths in the MGWR resulted in a more spatially nuanced model with the potential to facilitate deeper understandings of landscape processes and relationships.

Under a standard regression model, the model residuals (errors) were found to be spatially autocorrelated. A standard GWR was undertaken but was found to both overfit and underfit the model due to the use of a single bandwidth for all variables. This resulted in highly localised patterns of variation in the coefficient estimates, demonstrating both positive and negative associations with perceived landscape beauty in different locations. To address this limitation, a MGWR was undertaken to allow the parameter-specific scale of the relationship between the target variable and each landscape factor to vary, enabling local (spatially non-stationary) and global (stationary) relationships between them. The MGWR results indicate that the relationship between remoteness and scenicness operates on a global scale, whereas the relationships for absence, naturalness and ruggedness operate over several degrees of localness. These findings support the use of MGWR as an exploratory tool, reinforcing the notion that it should function as the default geographically weighted model. It holds great potential for bridging objectivist and subjectivist paradigms and supporting integrated landscape assessments. A standard GWR should only be undertaken if there is evidence that the covariates have the same scale of relationship with the target variable. Unfortunately, most existing applications of a GWR in landscape literature and practice do not do this.

CRediT authorship contribution statement

Yi-Min Chang Chien: Software, Validation, Formal analysis, Investigation, Writing - original draft, Visualization. **Steve Carver:** Conceptualization, Investigation, Resources, Writing - review & editing, Supervision. **Alexis Comber:** Methodology, Investigation, Software, Visualization, Writing - review & editing, Supervision.

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.

Acknowledgements

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Chang Chien is grateful for the sponsorship offered by Taiwan's Ministry of National Defence. Comber is funded by the UK Natural Environment Research Council grant NE/S009124/1. We gratefully acknowledge the editor and the anonymous reviewers for their valuable and constructive comments, resulting in an improved paper. We also thank all the developers, maintainers and contributors of Scenic-Or-Not and Geograph.

Appendix A. Supplementary data

https://doi.org/10.1016/j.landurbplan.2020.103904

References

Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. Á In: Petrov, BN and Csaki, F. In Second International Symposium on Information Theory. Akademiai Kiado, Budapest, pp. 276Á281.

Anselin, L. (2010). Thirty years of spatial econometrics. *Papers in Regional Science, 89*(1), 3–25. https://doi.org/10.1111/j.1435-5957.2010.00279.x.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). Regression Diagnostics. Regression Diagnostics. Hoboken, NJ, USA: John Wiley & Sons, Inc. doi: 10.1002/0471725153.

Bowman, A. W. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika, 71*(2), 353–360. https://doi.org/10.1093/biomet/71.2.353.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1999). Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Science, 39*(3), 497–524. https://doi.org/10.1111/0022-4146.00146.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis, 28*(4), 281–298. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x.

Butler, A., & Berglund, U. (2014). Landscape Character Assessment as an Approach to Understanding Public Interests within the European Landscape Convention. *Landscape Research, 39*(3), 219–236. https://doi.org/10.1080/01426397.2012.716404.

Carlson, A. (2019). Environmental aesthetics. Retrieved from https://plato.stanford.edu/archives/sum2019/entries/environmental-aesthetics/.

Carver, S., Comber, A., McMorran, R., & Nutter, S. (2012). A GIS model for mapping spatial patterns and distribution of wild land in Scotland. *Landscape and Urban Planning, 104*(3–4), 395–409. https://doi.org/10.1016/j.landurbplan.2011.11.016.

Carver, S., Evans, A. J., & Fritz, S. (2002). Wilderness attribute mapping in the United Kingdom. *International Journal of Wilderness, 8*(1), 24–29.

Carver, S., Tricker, J., & Landres, P. (2013). Keeping it wild: Mapping wilderness character in the United States. *Journal of Environmental Management, 131*, 239–255. https://doi.org/10.1016/j.jenvman.2013.08.046.

Carver, S., & Washtell, J. (2012). Real-time visibility analysis and rapid viewshed calculation using a voxel-based modelling approach. *In GISRUK 2012 Conference* (pp. 11–13). .

Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association, 74*(368), 829–836. https://doi.org/10.1080/01621459.1979.10481038.

Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., ... Harris, P. (2020). The GWR route map: A guide to the informed application of Geographically Weighted Regression, 1–34. *Retrieved from*. http://arxiv.org/abs/2004.06070.

Comber, A., Carver, S., Fritz, S., McMorran, R., Washtell, J., & Fisher, P. (2010). Different methods, different wilds: Evaluating alternative mappings of wildness using fuzzy MCE and Dempster-Shafer MCE. *Computers, Environment and Urban Systems, 34*(2), 142–152. https://doi.org/10.1016/j.compenvurbsys.2009.10.006.

Comber, A., & Harris, P. (2018). Geographically weighted elastic net logistic regression. *Journal of Geographical Systems, 20*(4), 317–341. https://doi.org/10.1007/s10109-018-0280-7.

Comber, A., Mooney, P., Purves, R. S., Rocchini, D., & Walz, A. (2016). Crowdsourcing: It matters who the crowd are. The impacts of between group variations in recording land cover. *PLOS ONE, 11*(7), Article e0158329. https://doi.org/10.1371/journal.pone.0158329.

Conrad, E. F., Cassar, L., Jones, M., Eiter, S., Izaovičová, Z., Barankova, Z., & Fazey, I. (2011). Rhetoric and Reporting of Public Participation in Landscape Policy. *Journal of Environmental Policy & Planning, 13*(1), 23–47. https://doi.org/10.1080/1523908X.2011.560449.

Daniel, T. C. (2001). Whither scenic beauty? Visual landscape quality assessment in the 21st century. *Landscape and Urban Planning, 54*(1–4), 267–281. https://doi.org/10.1016/S0169-2046(01)00141-4.

Dramstad, W. E., Tveit, M. S., Fjellstad, W. J., & Fry, G. L. A. (2006). Relationships between visual landscape preferences and map-based indicators of landscape structure. *Landscape and Urban Planning, 78*(4), 465–474. https://doi.org/10.1016/j.landurbplan.2005.12.006.

Dunkel, A. (2015). Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning, 142*, 173–186. https://doi.org/10.1016/j.landurbplan.2015.02.022.

Fotheringham, A. S., & Wong, D. W. S. (1991). The Modifiable Areal Unit Problem in Multivariate Statistical Analysis. *Environment and Planning A: Economy and Space, 23*(7), 1025–1044. https://doi.org/10.1068/a231025.

Fotheringham, A Stewart, & Oshan, T. M. (2016). Geographically weighted regression and multicollinearity: Dispelling the myth. *Journal of Geographical Systems, 18*(4), 303–329. https://doi.org/10.1007/s10109-016-0239-5.

Fotheringham, A Stewart, Yang, W., & Kang, W. (2017). Multiscale geographically weighted regression (MGWR). *Annals of the American Association of Geographers, 107*(6), 1247–1265. https://doi.org/10.1080/24694452.2017.1352480.

Fritz, S., Carver, S., & See, L. (2000). New GIS approaches to wild land mapping in Europe. In S. F. McCool, D. N. Cole, W. T. Borrie, & J. O'Loughlin (Eds.). *USDA Forest Service Proceedings RMRS-P-15-VOL-2* (pp. 120–127). Montana: Missoula.

Gliozzo, G., Pettorelli, N., Haklay, M., & (Muki) (2016). Using crowdsourced imagery to detect cultural ecosystem services: A case study in South Wales, UK. *Ecology and Society, 21*(3), art6. https://doi.org/10.5751/ES-08436-210306.

Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2015). GWmodel: An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. *Journal of Statistical Software, 1*(17), 2015. https://doi.org/10.18637/jss.v063.i17.

Hong, J. Y., & Jeon, J. Y. (2017). Exploring spatial relationships among soundscape variables in urban areas: A spatial statistical modelling approach. *Landscape and Urban Planning, 157*, 352–364. https://doi.org/10.1016/J.LANDURBPLAN.2016.08.006.

Kerebel, A., Gélinas, N., Déry, S., Voigt, B., & Munson, A. (2019). Landscape aesthetic modelling using Bayesian networks: Conceptual framework and participatory indicator weighting. *Landscape and Urban Planning, 185*, 258–271. https://doi.org/10.1016/j.landurbplan.2019.02.001.

Lothian, A. (1999). Landscape and the philosophy of aesthetics: Is landscape quality inherent in the landscape or in the eye of the beholder? *Landscape and Urban Planning, 44*(4), 177–198. https://doi.org/10.1016/S0169-2046(99)00019-5.

Lu, B., Brunsdon, C., Charlton, M., & Harris, P. (2017). Geographically weighted regression with parameter-specific distance metrics. *International Journal of Geographical Information Science, 31*(5), 982–998. https://doi.org/10.1080/13658816.2016.1263731.

Luo, J., & Wei, Y. H. D. (2009). Modeling spatial variations of urban growth patterns in Chinese cities: The case of Nanjing. *Landscape and Urban Planning, 91*(2), 51–64. https://doi.org/10.1016/j.landurbplan.2008.11.010.

Makeschin, F., Koschke, L., Frank, S., Fürst, C., & Witt, A. (2013). Assessment of landscape aesthetics—Validation of a landscape metrics-based assessment by visual estimation of the scenic beauty. *Ecological Indicators, 32*, 222–231. https://doi.org/10.1016/j.ecolind.2013.03.026.

Mei, C.-L., Xu, M., & Wang, N. (2016). A bootstrap test for constant coefficients in geographically weighted regression models. *International Journal of Geographical Information Science, 30*(8), 1622–1643. https://doi.org/10.1080/13658816.2016.1149181.

Morton, R. D., Rowland, C. S., Wood, C. M., Meek, L., Marston, C. G., & Smith, G. M. (2014). Land Cover Map 2007 (25m raster, GB) v1.2. NERC Environmental Information Data Centre. Doi: 10.5285/a1f88807-4826-44bc-994d-a902da5119c2.

Murakami, D., Lu, B., Harris, P., Brunsdon, C., Charlton, M., Nakaya, T., & Griffith, D. A. (2018). The importance of scale in spatially varying coefficient modeling. *Annals of the American Association of Geographers, 1–21*. https://doi.org/10.1080/24694452.2018.1462691.

Naismith, W. W. (1892). Excursions: Cruach Ardran, Stobinian, and Ben More. *Scottish Mountaineering Club Journal, 2*(3), 136.

Nash, R. (1982). *Wilderness and the American mind.* New Haven, Connecticut: Yale University Press.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity, 41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6.

Openshaw, S. (1984a). Ecological fallacies and the analysis of areal census data. *Environment & Planning A, 16*(1), 17–31. https://doi.org/10.1068/a160017.

Openshaw, S. (1984b). *The modifiable areal unit problem: Concepts and Techniques in Modern Geography.* Norwich: Geo Books.

Oshan, T., Li, Z., Kang, W., Wolf, L., & Fotheringham, A. (2019). mgwr: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating Process Spatial Heterogeneity and Scale. *ISPRS International Journal of Geo-Information, 8*(6), 269. https://doi.org/10.3390/ijgi8060269.

Oshan, T. M., & Fotheringham, A. S. (2018). A Comparison of Spatially Varying Regression Coefficient Estimates Using Geographically Weighted and Spatial-Filter-Based Techniques. *Geographical Analysis, 50*(1), 53–75. https://doi.org/10.1111/gean.12133.

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. *Ecological Indicators, 94*, 74–86. https://doi.org/10.1016/j.ecolind.2017.02.009.

Páez, A., Farber, S., & Wheeler, D. (2011). A Simulation-Based Study of Geographically Weighted Regression as a Method for Investigating Spatially Varying Relationships. *Environment and Planning A: Economy and Space, 43*(12), 2992–3010. https://doi.org/10.1068/a44111.

Palmer, J. F., & Hoffman, R. E. (2001). Rating reliability and representation validity in scenic landscape assessments. *Landscape and Urban Planning, 54*(1–4), 149–161. https://doi.org/10.1016/S0169-2046(01)00133-5.

Schirpke, U., Tasser, E., & Tappeiner, U. (2013). Predicting scenic beauty of mountain regions. *Landscape and Urban Planning, 111*(1), 1–12. https://doi.org/10.1016/j.landurbplan.2012.11.010.

Seresinhe, C. I., Preis, T., MacKerron, G., & Moat, H. S. (2019). Happiness is Greater in More Scenic Locations. *Scientific Reports, 9*(1), 4498. https://doi.org/10.1038/s41598-019-40854-6.

Seresinhe, C. I., Preis, T., & Moat, H. S. (2015). Quantifying the Impact of Scenic Environments on Health. *Scientific Reports, 5*(1), 16899. https://doi.org/10.1038/srep16899.

Seresinhe, C. I., Preis, T., & Moat, H. S. (2017). Using deep learning to quantify the beauty of outdoor places. *Royal Society Open Science, 4*(7), Article 170170. https://doi.org/10.1098/rsos.170170.

Su, Y.-F., Foody, G. M., & Cheng, K.-S. (2012). Spatial non-stationarity in the relationships between land cover and surface temperature in an urban heat island and its impacts on thermally sensitive populations. *Landscape and Urban Planning, 107*(2), 172–180. https://doi.org/10.1016/J.LANDURBPLAN.2012.05.016.

Sun, R., Xie, W., & Chen, L. (2018). A landscape connectivity model to quantify contributions of heat sources and sinks in urban regions. *Landscape and Urban Planning, 178*, 43–50. https://doi.org/10.1016/J.LANDURBPLAN.2018.05.015.

Swanwick, C. (2002). Landscape character assessment: Guidance for England and Scotland. Countryside Agency and Scottish Natural Heritage.

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

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. *Proceedings of the National Academy of Sciences, 113*(46), 12974–12979. https://doi.org/10.1073/pnas.1614158113.

Wang, L., Ai, T., Shen, Y., & Li, J. (2020). The isotropic organization of DEM structure and extraction of valley lines using hexagonal grid. *Transactions in GIS, 24*(2), 483–507. https://doi.org/10.1111/tgis.12611.

Wheeler, D., & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems, 7*(2), 161–187. https://doi.org/10.1007/s10109-005-0155-6.

Wolf, L. J., Oshan, T. M., & Fotheringham, A. S. (2018). Single and multiscale models of process spatial heterogeneity. *Geographical Analysis, 50*(3), 223–246. https://doi.org/10.1111/gean.12147.

Yang, W. (2014). *An extension of geographically weighted regression with flexible bandwidths.* University of St Andrews.

Zoderer, B. M., Tasser, E., Carver, S., & Tappeiner, U. (2019). An integrated method for the mapping of landscape preferences at the regional scale. *Ecological Indicators, 106*, Article 105430. https://doi.org/10.1016/j.ecolind.2019.05.061.

Zube, E. H., & Pitt, D. G. (1981). Cross-cultural perceptions of scenic and heritage landscapes. *Landscape Planning, 8*(1), 69–87. https://doi.org/10.1016/0304-3924(81)90041-1.