Integrating expert-based objectivist and nonexpert-based subjectivist paradigms in landscape assessment

DISSERTATION

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Geography

By
Yi-Min Chang Chien

January 2024

Declaration

The present thesis is based on the following three papers where the contributions of the PhD candidate will be referred to by their Roman numerals.

I. Chang Chien, Y. M., Carver, S., & Comber, A. 2020. Using geographically weighted models to explore how crowdsourced landscape perceptions relate to landscape physical characteristics. Landscape and Urban Planning, 203(August), 103904. (10.1016/j.landurbplan.2020.103904)

II. Chang Chien, Y.-M., Carver, S., & Comber, A. 2021. An Exploratory Analysis of Expert and Nonexpert-Based Land-scape Aesthetics Evaluations: A Case Study from Wales. Land, 10(2), 192. (10.3390/land10020192)

III. Chang Chien, Y.-M., Carver, S., & Comber, A. Using crowdsourced data and wildness measures to model landscape aesthetic quality: an integrative approach using supervised machine learning. ready for submission to Landscape and Urban Planning.

Publication I: The PhD candidate contributed to co-developing research idea, data processing and analysis, mapping the results, writing the draft manuscript and incorporating co-author's feedback.

Publication II: The PhD candidate contributed to developing research idea, data processing and analysis, visualising the results, writing the draft manuscript and incorporating co-author's feedback.

Paper III: The PhD candidate contributed to developing research idea based on the findings of the previous two publications, data processing and analysis, data visualisation including creating the map results, writing the draft manuscript and incorporating co-author's feedback.

Dedication

To my loving parents
for their love, support and understanding.

Acknowledgements

I wish to express my deepest gratitude to the people who have supported and assisted me throughout the duration of my studies. First and foremost, I would like to extend my most profound thanks to my principal and second supervisors, Prof. Alexis Comber and Dr. Steve Carver respectively, for encouraging and spurring me on to do my best. They have guided and helped me to overcome many challenges, and I have been immensely inspired by their easygoingness and sense of humour. In particular, I want to thank my principal supervisor for introducing me to the field of geospatial analysis. Prof. Comber always lent a helping hand whenever I reached a dead-end. With his mentoring and guidance, I have transformed from a novice in academia into a researcher who can work independently. I also extend my thanks to Prof. Guy Ziv and Prof. David Miller for their insightful comments and suggestions that significantly improved the quality of this thesis. Furthermore, I would like to express my gratitude to Jacqui and Nicolas, both of whom helped with my extension requests in response to the Covid-19 pandemic.

No word can express my gratitude to my family, who provided countless support throughout this journey. I am immeasurably grateful for their rock-steady support and companionship during this solitary journey; to my parents for their never-ending patience, encouragement, and care; to my cousin Jill for providing English consultation needed at the beginning and the end of this journey; and to many good friends I have met during my PhD: Leon, Marie, Ye Zi, Wanyun, Sungje, Yuanxuang, Rakeshi, Joey, Nikee, Antonio—thank you for your continued friendship, comfort food, hugs, board games, pints along the way.

Special thanks also go to my high school mentor, Mrs. Jen-Chiu Lee and her husband, Admiral Hsi-Min Lee (Ret.). Without their invaluable encouragement and assistances, I would never step on this path. I truly appreciate their belief in my abilities.

I could not have done any of this without you. Thank you all.

Taiwan, January 2024

Summary

There is an increasing realisation that public perceptions of landscape should be taken into account in national landscape governance and planning. The international treaty of European Landscape Convention (ELC) has placed the public central to any understanding of landscape and the signatories of the Convention are obligated to work towards that goal. Despite the call for of public involvement, most national landscape assessments in practice are still highly expert-centred and in a top-down manner which hardly ever matches the participatory stance. Limited resources and methodological challenges often restrict public engagement and consultation on a large scale. The lack of more inclusive and citizen-centric approaches to collect different views of the public on landscapes are identified. Furthermore, more pragmatic and inclusive approaches that facilitate a systematic integration of multiple perspectives of landscapes are required for countries that have ratified the Convention, such as the United Kingdom, to narrow the gap between the participatory rhetoric and practice.

With the development of Web 2.0 technologies and the proliferation of smartphones and other location-aware devices, people are enabled to contribute contents (e.g., photographs and texts) with georeferenced information, often in relation to how they individually perceive the environment. In this respect, prodigious amounts of such crowdsourced geographic information associated with subjective opinions on perceived environment have become ubiquitously available. Despite a rapid increase of efforts to incorporate these bottom-up data collection and analysis in the field of landscape assessment, relatively little progress has been made in better understanding the association of expert-based metrics of landscape quality with these public inputs. The expert-based objectivist paradigms of landscape assessment might be prone to fall into inconsistency with the nonexpert-based subjectivist one. This could, in turn, lead to less informed decision and practices in the context of public participation in landscape conservation, planning, and management.

The aim here sought to integrate the objective and subjective measures of landscape aesthetic quality, regarded as complementary and supplementary to each other. This thesis concentrated on crowdsourced geo-information, specifically regarding public perceived values of landscape aesthetics or environmental aesthetics in a broader sense, as captured in the Scenic-Or-Not initiative for Great Britain. GIS-based Wildness spatial layers—consisting of four components: absence of modern human artefacts (absence), perceived naturalness of land cover (naturalness), remoteness from mechanised access (remoteness) and rugged and physically challenging nature of the terrain (ruggedness)—as well as the LANDMAP dataset regarding visual and sensory landscape classification

for Wales were also utilised herein, both of which served as expert-based objectivist paradigms of landscape assessment.

To approach this aim, several studies were conducted: first, an investigation into the degree of correlations between the four formal measures of wildness components and the crowdsourced measures of landscape aesthetic quality, and to further explore the spatial variation and the geographic scale (or influence scope) in their relationships across Great Britain using the multiscale geographically weighted regression (MGWR) framework which is a recent extension to the geographically weighted regression (GWR) model (Publication I). Second, a comparison of landscape aesthetic perceptions held by experts and citizens was drawn based on the results of five generalised linear models (GLMs). These models were constructed over the visual and sensory areas from the LANDMAP classification (a national baseline about landscape in Wales). The similarities and discrepancies between these two perspectives of scenic quality concerning the ranking orders of the landscape typologies were pinpointed (Publication II). Finally, an integrative approach that conflates the aforementioned three datasets – LANDMAP, Scenic-Or-Not, Wildness – was proposed to address the lack of any systematic and reproducible evaluations of landscape quality and aesthetics for England and Scotland. As the LANDMAP data of landscape aesthetics was undertaken by two separate expert groups, two extreme gradient boost (XGBoost) models were constructed over the LANDMAP areas for each group and then applied to predict the spatial patterns and distribution of landscape aesthetics across England and Scotland (as well as Wales) (Paper III).

The first study evidenced that the relationships between objective measures of landscape wildness quality and subjective measures of landscape aesthetics are statistically significant; each wildness component display different degrees of correlations. Overall, MGWR is more sensitive than GWR to the analysis of spatial heterogeneity in the statistical relationships between landscape factors and public perceptions. Notably, the MGWR output for remoteness exhibited very limited variation and a wide bandwidth, indicating a stationary or globally fixed process. Remoteness is mainly concerned with the opposite side of the ease of reaching or travelling to a specific landscape, which can therefore be interpreted as a contextual factor for aesthetic perceptions of landscapes. In contrast, the relationships for absence, naturalness and ruggedness showed different degrees of localness which can be used to facilitate targeted landscape management. Based on these findings, both wildness and scenicness measures were feasible to be incorporated into formal landscape aesthetic assessment. The second study further provided a few observations that were summarized as follows: first, water-related landforms generally received higher appraisal from citizens, followed by upland, lowland, and development while upland typologies were found to be highly variable, followed by lowland, water, and development. Second, a higher level of agreement between experts and non-experts was observed in assessing landscapes at both extremes of the scenic spectrum – that is, those universally perceived as highly aesthetically pleasing or the contrary. Conversely, there was less consensus on landscapes falling in the middle of the spectrum, where aesthetic judgements tend to be more subjective and varied. This consultation with on-the-ground views of landscape scenic quality has made it possible to identify the contesting landform typologies and subsequently suggestions were made in relation to the deficiency in the principally vertical, bird's-eye assessment led by professionals.

These results motivated the third study to construct models of LANDMAP classes evaluated by two groups of experts using data from the Scenic-Or-Not initiatives and Wildness spatial layers. Two extreme gradient boost (XGBoost) models were constructed over the LANDMAP areas for each group and then applied to predict the spatial patterns and distribution of landscape aesthetics in the manner of LANDMAP across Great Britain. The two predictive models reach Overall classification accuracies of 67.3% and 74.5%, and the Kappa statistics of 0.50 and 0.64, and are comparable to the previous studies based on traditional statistical models that use spatial metrics or landscape features (e.g., terrain roughness and proximity to water) to predict scenic beauty. If different expert perceptions are explored (as in Figure 5.4), then a consensus of important regions of outstanding scenery, can be determined, across different evaluations (as in Figure 5.5). The resulting maps can potentially be used to complement current Landscape Character Assessment (LCA) practices in support of relevant landscape policy decisions.

From a data point of view, one should be aware of the limitations and biases inherent in those data generated from the crowdsourcing initiative in relation to their data quality such as spatial coverage and sampling as well as representative uncertainties. From a methodological point of view, the aggregation strategy was required to alleviate the biases in crowdsourced data and make the computation manageable and operational. Caution should also be exercised in aggregating the point-based Scenicness scores and grid-based wildness measures to a hexagonal grid as well as the LANDMAP areas. The classic geospatial analytic issue, associated with the modifiable areal unit problem (MAUP), which refers to a statistical bias occurred when point-based observations aggregated over different areal units and scales, could produce misleading prediction or inference results, has been recognised and discussed.

This thesis is structured as follows. Chapter 1 describes the motivation and relevance, while Chapter 2 puts the questions analysed in this project into a broader context of background. The 3 papers (2 published and 1 for submission as below) are presented in Chapter 3, Chapter 4, and Chapter 5. Finally, Chapter 6 devotes to the discussion of each presented paper as well as the overall limitations with regard to data and methods along with the conclusion.

Publication I:
Chang Chien, Y. M., Carver, S., & Comber, A. 2020. Using geographically weighted models to explore how crowdsourced landscape perceptions relate to landscape physical characteristics. *Landscape and Urban Planning*, 203(August), 103904. (10.1016/j.landurbplan.2020.103904)

Publication II:
Chang Chien, Y.-M., Carver, S., & Comber, A. 2021. An exploratory analysis of expert and nonexpert-based landscape aesthetics evaluations: a case study from Wales. *Land*, 10(2), 192. (10.3390/land10020192)

Paper III:
Chang Chien, Y.-M., Carver, S., & Comber, A. Using crowdsourced data and wildness measures to model landscape aesthetic quality: an integrative approach using supervised machine learning. ready for submission to *Landscape and Urban Planning*.

List of Contents

List of Tables

List of Figures

List of Abbreviations

CES – Cultural Ecosystem Services
CGI – Crowdsourced Geographic Information
COSP – Change Of Support Problem
EDA – Exploratory Data Analysis
EDSA – Exploratory Spatial Data Analysis
ELC – European Landscape Convention
GB – Great Britain
GIS – Geographical Information System
ICT – information and communications technology
LANDMAP – Landscape Assessment Decision Making Process
LCM – Land Cover Map
LP&P – Landscape Perception and Preference
LQO – Landscape Quality Objective
MAUP – Modifiable Areal Unit Problem
MCE – Multiple Criteria Evaluation
XGBoost – eXtreme Gradient Boosting
SHAP – Shapley Additive exPlanations
SVC – Spatially Varying Coefficient

Chapter 1 Introduction

1.1 Motivation and relevance

Landscape and land are related constructs, but landscape involves human perception while land does not. Landscapes hold different meanings for various individuals in diverse contexts and play a crucial role in people's identity and well-being, evoking strong emotions and providing numerous benefits, such as stress relief. Over the past two decades, landscape issues have gained prominence within European policy agenda, significantly influenced by the European Landscape Convention (Council of Europe, 2000). The Convention defines landscape as 'an area, as perceived by people, whose character is the result of the action and interaction of natural and/or human factors.' Article 6 of the ELC mandates public consultation for defining landscape quality objectives (LQO), placing the public at the centre of any understanding of landscapes (see Section 2.6). The Convention promotes participatory and democratic approaches (Prieur, 2006), prioritising public involvement in landscape policy discussions.

However, there is currently no consensus on best practices for public consultation and public input has not been adequately incorporated into landscape assessment and planning, leading to limited influence on national policy issues (Conrad, F. Cassar, et al., 2011). The first challenge of implementing the ELC involve the elicitation of public landscape perception. Traditional methods, such as household surveys and in-situ interviews, have primarily focused on visual perception, but are limited to small-scale case studies with restricted respondent groups due to associated costs (Daniel, 2001). The lack of large-scale or national-level perception data hinders fulfilling the ELC's participatory stance (Conrad, F. Cassar, et al., 2011). Thus, there is a need for innovative approaches to gather varying perceptions of landscapes from broader populations and operationalize public participation on a larger scale.

Advances in information and communications technology (ICT) have paved the way for research on landscape perception and preference (LP&P) and cultural ecosystem services (CES). The emergence of Web 2.0 technology enables individuals to create and share large amounts and varieties of content with each other online (O'Reilly, 2005). This "bottom-up" content generation, combined with top-down organisational goals is referred to as crowdsourcing (Brabham, 2013). For instance, the advent of photo-sharing platforms with location-based services (e.g., Geograph and Flickr) has facilitated the contribution of crowdsourced, geotagged photographs and information that is relevant to landscapes (Bubalo et al., 2019). These online initiatives enhance understanding of how people perceive their environments and facilitate the engagement of diverse public groups, which is often unattainable with traditional data collection methods, such as field questionnaires (Dunkel, 2015). From a practical perspective, crowdsourcing strategies can help in understanding people's perceived landscapes and meeting the requirements of the ELC, which forms an integral part of this thesis.

Landscape studies attract interest from numerous disciplines, each with various, disparate focuses, and distinct methods and tools. There are different methodologies for assessing visual landscape quality (Daniel, 2001). Among these, an important conceptual distinction can be made between objective (physical) and subjective (psychological) methodologies, often associated with quantitative and qualitative evaluation approaches. The objectivist interprets visual quality as inherent to the landscape, whereas the subjectivist views it as a construct of the observer (Lothian, 1999). Previous studies have established relationships between expert-based and perception-based assessments (Dramstad et al., 2001; Dramstad et al., 2006; Schirpke et al., 2019). It has been suggested that the integration of objective and subjective methodologies could be the ideal tool for any understanding of landscapes (Lothian, 1999; Daniel, 2001). In a more recent review, a research trend towards transdisciplinary and integrative landscape assessment practices that evaluate objective and subject factors in combination has also been observed (Medeiros et al., 2021).

Over the past few decades, landscape character assessment (LCA) methods have been widely employed to analyse, classify, and map landscape types, using a combination of a map-based objective characterisation with an expert-based subjective assessment (Van Eetvelde and Antrop, 2009). Initially developed in the United Kingdom (UK), LCA has become an essential property of landscape in the ELC's definition (Van Eetvelde and Antrop, 2009). As a key communicative framework, it serves as a basis for landscape management and planning (Swanwick, 2002). However, little effort has been made to incorporate public inputs into LCA, allowing for direct comparisons between expert and public opinions. By identifying similarities and differences, this could facilitate the reconciliation of varying approaches. Moreover, landscape characterization relies on the visual interpretation and integration of mapped data sources, which can be labour-intensive and time-consuming. In response to rapid landscape changes, land managers and policymakers require more efficient methods to evaluate landscape scenic quality.

1.2 Research aim

This thesis not only seeks to ascertain the extent to which crowdsourced geographic information can supplement existing landscape assessment practices, but also aims to demonstrate the application of cutting-edge modelling techniques that support integration. Consequently, the research objectives are outlined and discussed in a separate chapter as follows:

- Objective 1: Explore the link between the objective and subjective assessments of landscapes in terms of spatial patterns (Chapter 3).
- Objective 2: Conduct a comparative study of landscape scenic quality evaluations made by experts versus non-experts (Chapter 4).
- Objective 3: Develop an integrative approach that efficiently combines expert and public perspectives relevant to landscape aesthetic quality (Chapter 5).

1.3 Study area

This thesis comprises three studies, all conducted within Great Britain (GB), which consists of three countries: England, Scotland, and Wales (Figure 1.1). Each country follows a similar expert-based methodology for assessing landscapes. However, Wales adopts a unique approach to establishing a landscape baseline that includes perceptual information, used as expert-based landscape evaluations. The difference in practices among these countries will be discussed in greater detail in Section 2.5. Chapter 3 focuses on the geographic region of GB while Chapter 4 is centred solely on Wales. Chapter 5 uses data from Wales for modelling and the predictive models are subsequently applied to map scenic quality across GB.

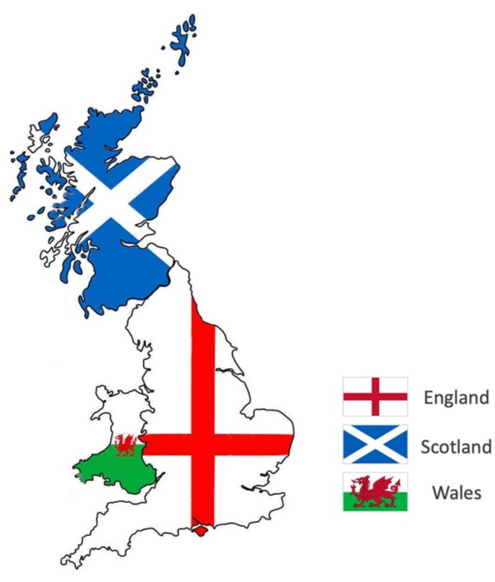


Figure 1.1 The study area (GB) comprising three countries: England, Scotland, and Wales.

1.4 Chapter outline

The remainder of this thesis, including the corresponding publications, is structured as follows:

- Chapter 2: Literature Review - This chapter reviews different approaches and relevant concepts related to studying visual landscape quality, as well as methodological approaches for modelling.
- Chapter 3: Publication I - Published in Landscape and Urban Planning (Chang Chien et al., 2020), this paper quantitatively explores the relationships between formal measures of landscape wildness and crowdsourced measures of perceived landscape scenic quality, illustrating the potential of a more spatially nuanced model.
- Chapter 4: Publication II - Published in Land (Chang Chien et al., 2021), this paper compares expert and public evaluations of landscape scenic quality and discusses potentially contested landscape typologies.
- Chapter 5: Publication III with the title: "Using crowdsourced scenic ratings and wildness measures to model landscape aesthetic quality: an integrative approach using supervised machine learning" – Submitted to Landscape and Urban Planning, this paper demonstrates how information and data from diverse sources can be integrated with formal evaluations of landscape scenic quality to provide extensive and consistent landscape quality assessments.
- Chapter 6: Discussion and conclusion - This chapter synthesizes the research findings, discusses potential broader implications for current assessment practices, and examines overall limitations. Several future research directions are suggested to make current landscape assessment practices more consistent, tractable, democratic, and accountable.

Chapter 2 Literature review

2.1 Objectivist and subjectivist landscape assessments

Approaches for assessing landscape quality can be grouped into two general categories: objectivist and subjectivist (Lothian, 1999). These contrasting views stem from the long-standing question of whether beauty lies 'in the object' or 'in the eyes of the beholder' (Meinig, 1976). Objectivist approaches, which primarily rely on expert knowledge to measure physical landscape characteristics, have long dominated visual landscape quality assessments. In contrast, subjectivist approaches, which mainly focus on understanding observers' perceptual reactions to landscapes in varying contexts, have been widely adopted in much landscape research. A range of different approaches for visual landscape quality assessment can be roughly situated between these two positions. However, these approaches differ in the relative importance they confer on the landscape and the observer.

Lothian (1999) proposed that landscape research and practices should shift from objectivist paradigms to subjectivist ones (Lothian, 1999). Several researchers also believed that visual landscape quality should be assessed and measured through the appreciation of the observer (de la Fuente de Val et al., 2006; G Fry et al., 2009). Yet, the subjectivist approaches have been found difficult to apply in practice due to large individual differences in individual perceptions and preferences. Thus far, it has become generally accepted that landscape quality is derived from the interaction between biophysical and perceived components of landscapes. In this respect, combining both objective and subjective assessments is deemed as a more inclusive and appropriate practice (Daniel, 2001).

Zube, Sell, and Taylor (1982) proposed a classification scheme which grouped the different approaches concerning visual landscape quality into four paradigms: expert, psychophysical, cognitive, and experiential (see Table 2.1) (Zube et al., 1982). Later, (Daniel and Vining, 1983) suggested a rather similar classification criteria where these approaches can be split into the ecological, formal-aesthetic, psychophysical, psychological, and phenomenological models. The difference in these two reviews is that (Daniel and Vining, 1983) further distinguished the former's expert paradigm into the ecological and formal-aesthetic models and described the former's cognitive and the experiential paradigms as the psychological and phenomenological models. The expert, ecological and formal-aesthetic sets are adherents of the objectivist, whereas the psychophysical, cognitive/psychological, and experiential/phenomenological sets hold the subjectivist position.

The ecological model assumes landscape quality as directly related to naturalness, or ecosystem integrity which can be entirely determined by ecological and biological features in the landscape and is, consequently, independent of the observer's judgement. Implicit in such model is the assumption that areas left undisturbed by human activity have the highest scenic quality. The formal-aesthetic model is primarily concerned with the visual characteristics of landscapes. Visual elements such as shapes, lines, colours, and textures are applied to represent the innate visual qualities of a landscape. The relationships between these elements are then examined to classify each area in terms of aesthetic qualities such as unity, symmetry, harmony, and contrast. The ecological and formal-aesthetic models are usually carried out by qualified and trained experts in the field of ecology, art and design or resource management. Both models can therefore be cross-referenced to the expert paradigm noted by (Zube et al., 1982). However, such expert-based assessment approach that stresses the evaluation of "scene" has been criticised as deficient because people's "sensory" responses are ignored.

The psychophysical model aims to investigate the statistical relationship between the physical characteristics of a landscape and the perceptual judgements of human observers, operating under the assumption of a stimuli-response relationship. Remote sensing and geospatial technologies are usually employed to measure landscape features (e.g., perceived naturalness and landscape type). Various representations of landscapes, such as photographs and street view images, have been used to elicit people's perceptions (Dubey et al., 2016; Zoderer et al., 2019). The perceptual response is generally limited to a single dimension, focusing on scenic and visual qualities of landscapes, and measured using an interval rating scale. Consequently, these perceptual measures of landscape quality can be effectively integrated with expert-based measures, including spatial indicators and metrics, that quantify specific aspects of the landscape. Psychophysical assessments, which capture a consensus of opinions between experts and laypeople, have demonstrated reliability in various management contexts (Daniel and Vining, 1983). The conceptual framework presented in Section 2.9 is inspired by this psychophysical model.

The cognitive/psychological model also takes a position in between the objectivist and subjectivist approach which is regarded as methodologically rigorous as well (Taylor et al., 1990). However, the cognitive/psychological approaches tend to be not directly applied to landscape quality assessments as with an in-depth focus on the human aspect of human-landscape interaction in terms of human cognitive, emotional, and behavioural responses to landscapes. Over decades, enormous efforts have also been put into understanding the impact of a variety of factors (e.g., gender, age, occupation, hobbies, education, profession, familiarity, nationality, and religion) on landscape perception and experience (Aoki, 1999). These studies provide the basis for explaining human aesthetics to landscape through evolutionary processes and functional demands and form the theoretical foundations (see Section 2.2).

Finally, the experiential/phenomenological model places greater emphasis on individual's personal experiences to comprehend the underlying meanings of human–environment interactions (Ohta, 2001). Such approaches treat landscapes as "more-than-visual and more-than-symbolic" product from the interactions. Both are mutually shaping and being shaped by each other. As such, landscape is imbued with emotions, meanings, identity, and hidden narratives ascribed by people. To describe this, a number of inventories such as surveys, verbal questionnaires, open-ended interview questions have been developed for inspiring individual expression and creativity. The information is recorded descriptively and qualitatively that is often without numerical support.

Table 2.1 summarises these two classification schemes that range on a scale from objectivist to subjectivist. The objectivist paradigm comprises expert, ecological and formal aesthetic typologies, which are based on objective measurements and assessments. In contrast, the subjectivist paradigm encompasses psychophysical, psychological/cognitive, experiential/phenomenological aspects, which are more subjective and based on human perception and experience. These two seminal works provide an overview of the diverse contemporary paradigms for landscape assessment which remains salient to the work of landscape planners. It has been hypothesised that neither of these models are capable of achieving all the goals of aesthetic landscape assessment, but that landscape assessment practices that incorporate objective and subjective methodologies are considered to be balanced, valid, and trustworthy, enabling better-informed policymaking (Gobster et al., 2019).

Table 2.1. Comparison of paradigms of landscape quality assessment (Lothian, 1999).

Lothian, 1999	Objectivist/Physical paradigm		Subjectivist/Psychological paradigm		
Zube et al., 1982	Expert		Psychophysical	Cognitive	Experiential
Daniel and Vining, 1983	Ecological Aesthetic	Formal Aesthetic	Psychophysical	Psychological	Phenomenological

2.2 Landscape aesthetic theories

Accordingly, theories explaining landscape aesthetic preferences can be broadly categorised into two distinct premises, evolutionary and cultural preference theories, relating to innate and learned human behaviours respectively (Tveit et al., 2018) which will be explained in more detail in the following sections.

2.2.1 Evolutionary theories

Evolutionary theories posit that human aesthetic responses to landscape are partially innate and fixed in the course of evolution. The basic tent of evolutionary explanations is that our primitive ancestors preferred landscapes that enhanced their survival and well-being and the genetic bases for such landscape preference would still be inherent in modern humans today. This may explain why certain landscapes are generally preferred across population groups, such as the desire for a landscape with water. Several theories such as Appleton's prospect-refuge theory, Orians' habitat theory, and Kaplan's information-processing theory are salient examples. The prospect-refuge theory focuses on the role of our primitive ancestors as both predator and prey, thus needing to be able to survey the landscape for food and to hide from large predators (Appleton, 1975). The presence of prospect and refuge in a landscape that allow to see without being seen is, thus, favourable to survival in early human communities. Following Appleton, Gordon Orians postulated the habitat theory, explaining environmental preferences as the results of the search for a suitable habitat (Orians, 1980). It is stated that humans tend to favour savannah-like environments because these are suitable habitats for our distant ancestors (Orians, 1980). The Kaplans' information processing theory added to the evolutionary explanation by stating that people have intrinsic needs for understanding and exploring environments (Kaplan and Kaplan, 1989). Landscapes that aid in rapid understanding and incite further exploration of the environment would hence be favoured over those that fail to satisfy or even hinder these needs. The four "informational variables"—coherence, legibility, complexity, and mystery—were identified by crossing the two needs, i.e., understanding and exploration, with the two levels of interpretation, i.e., immediacy and inference, that is required in extracting the information in two-dimensional and three-dimensional space. Table 2.2 depicts the preference matrix. Coherence and legibility help one understand the environment at different levels of immediacy or degree of inference while complexity and mystery encourage further exploration in a similar vein. In this regard, the more coherent, complex, legible, and mysterious the scene, the more desirable it is. Further, a close link between Kaplan's and Appleton's theories (e.g., legibility and prospect, mystery and refuge) can be seen.

Table 2.2. Relationship between factors predicting environmental preference.

	Informational needs	
Level of interpretation	Understanding	Exploration
Immediacy (two-dimensional)	Coherence	Complexity
Inference (three-dimensional)	Legibility	Mystery

2.2.2 Cultural theories

Cultural theories, in contrast to evolutionary ones, explain landscape preference as learnt and shaped by social, cultural, and personal characteristics, focusing on the influences of individual attributes such as age, gender, occupation, ethnicity, education, and familiarity, rather than the immediate and affective preference responses (Bell, 2012). Much quoted cultural theories include topophilia and the ecological aesthetics. First, topophilia, coined by (Tuan, 1990), means "the affective bond between people and place or setting", implying that familiarity and experience plays an important role in the forming of landscape preference (Tuan, 1990). Thus, this theory focusses on the historical aspect of landscape perception and advocates that a person tends to have a strong attachment to familiar places. Congruent with topophilic ideas, (Adevi and Grahn, 2012) discovered strong general support for an influence of the childhood landscape on adult preferences. Several variations of topophilia, such as sense or spirit of place, place-identity, and place-attachment, were developed to express the idea that 'place' goes beyond landscape and involves characteristics such as identity, history, and memory (Bell, 2012). Second, the ecological aesthetic recognises knowledge is an important driver of landscape preference, arguing that scientific knowledge about the ecological functions of a landscape will lead to appreciate its beauty (Gobster, 1999). This new type of beauty has largely been advocated to address issues concerning the preservation of ecologically significant landscapes which is seen as an attack on traditional notions of scenic beauty (Gobster, 2008). There have been many attempts to expand this scope of landscape aesthetics in order to better understand and bridge the gap between aesthetics and ecology (Carlson, 2004). However, this concept is omnipresent in debates about its overemphasis on the role of ecological knowledge and the neglect of richly aesthetic experience of landscape that is affected by different contextual environments (Brady and Prior, 2020). Other cultural theories include landscape heritage approaches highlighting the importance of cultural heritage, such as stone walls, archaeological ruins, and grave sites (Fairclough et al., 1999) and aesthetics of care highlighting the role of human agency in creating and maintaining landscapes (Nassauer, 2011).

2.2.3 Integrative theoretical framework

Thus far, a widely accepted theory of landscape, providing an all-encompassing framework with which to understand and to predict landscape preferences does not exist. However, there is a growing recognition of the interplay between biological and cultural components in shaping human landscape preferences (Zube et al., 1982; Bourassa, 1990; Dramstad et al., 2006) that promotes the development of integrative theoretical frameworks. These frameworks rely on a combination of evolutionary and cultural preference theories where the assumption is made that human aesthetic preferences for landscapes are rooted in the same evolutionary history, and subsequently modified by cultural, social and personal factors, leading to the divergence in aesthetic preferences observed across different cultures and individuals (Tveit, 2009; Bell, 2012). Some landscape elements, such as water, appear to be rather universally valued whereas the evaluations of other attributes, such as openness, vary depending on observer characteristics (M. Tveit et al., 2006; Sevenant and Antrop, 2010).

Tveit et al. (2006) developed a seminal framework for analysing visual landscape characteristics, identifying nine key concepts: stewardship, coherence, disturbance, historicity, visual scale, imageability, complexity, naturalness, ephemera (see Table 2.3). These visual landscape characteristics are also presented in the visual guidelines of LCA (see Section 2.5.1). The interpretation of these visual aspects is, nevertheless, context-dependent where the surrounding environment come into play as well as the observer's interest and values (G. Fry et al., 2009). In order to position the following section, it is necessary to mention that besides complexity (Kaplan and Kaplan, 1989), naturalness is a widely used concept as a key aspect of visual landscape quality (Ode et al., 2009). A dimension closely related to naturalness is the idea of wilderness that is regarded as the most extreme manifestation of naturalness in a biological and ecological sense. Studies on mapping wilderness perception are a more specified sub-research area of the broader field of landscape perception whereas many approaches and methods are similar.

These visual concepts can be further categorised into four levels of above-mentioned aesthetic cognition (Nohl, 2001). Perceptual level includes visual scale, coherence, and complexity; expressive level includes ephemera; interpretative level includes naturalness, disturbance, and stewardship; and symbolic level includes historicity and imageability (Lee and Son, 2017). These levels of cognitive processes can be referred to Section 2.6.2.

Table 2.3. Theory-based concepts of visual landscapes character with their definition (complied from Tveit et al., 2006; Ode et al., 2008)

Concept	Definition	Theory
Stewardship	Degree of human care for the landscape through active and careful management, contributing to a perceived accordance to an 'ideal' situation	Aesthetic of care
Coherence	Unity of a scene, repeating patterns of colour and texture, correspondence between land use and natural conditions	Information processing Theory
Disturbance	Lack of contextual fit and coherence, constructions and interventions	Information processing theory, Biophilia hypothesis
Historicity	Historical continuity and historical richness, different time layers, amount and diversity of cultural elements	Topophilia
Visual scale	Landscape rooms or perceptual units: their size, shape and diversity, degree of openness	Prospect-refuge theory Habitat theory
Imageability	Qualities of a landscape present in totality or through elements; landmarks and special features, both natural and cultural, making the landscape create a strong visual image in the observer, and making landscapes distinguishable and memorable	Spirit of place, Vividness, Topophilia
Complexity	Diversity, richness of landscape elements and features, interspersion of pattern	Information processing theory, Biophilia hypothesis
Naturalness	Closeness to a preconceived natural state	Biophilia hypothesis
Ephemera	Changes with season, weather or other temporal effects	Restorative environments

2.3 Conceptual links between wilderness and aesthetics

2.3.1 Wildness of wilderness

The concept of wilderness is complex and debated, as it encompasses both tangible natural environments as well as abstract, subjective feelings and preconceptions onto it that are constantly evolving and changing (Bell, 2012). As Roderick Nash observes, "wilderness is so heavily freighted with meaning of a personal, symbolic, and changing kind as to resist easy definition." Studies on wilderness experiences demonstrate its multidimensional nature, encompassing naturalness, primitiveness, remoteness, solitude, freedom, spirituality, and aesthetics, all of which are important in cognitive and affective senses. It is now widely accepted in academic circles (though not uncontroversial) that the idea of nature and wilderness is culturally constructed (Warren, 2009).

Despite of the controversy, there is an increasingly accepted idea that any land can be characterised by two fundamental and independent qualities: naturalness and freedom (Aplet et al., 2000). Naturalness refers to a state of a land free from human effects while freedom is determined by the degree to which it is free from human intent. Wilderness is that portion of the land that possess the most pristine and self-willed while built environments are at the most artificial and controlled ends of the spectra. As there is very little pristine wilderness left in modern time due to the prolonged anthropocentric activities, the concept of 'wild land' has been created, which refers to wilderness in cultural landscape context. The term refers to lands that have been, or are, uninhabited or less influenced by human activities in terms of their characters and qualities (Fairclough, 2006). Figure 2.1 illustrates the wilderness continuum concept and gives a few broad types of land (i.e., non-wilderness wildland, semi-wildland, and ex-urban non-wildland) along the spectrum of naturalness and freedom.

On these grounds, there is a shift in language and concept from wilderness to wildness (Warren, 2009). As Howard Zahniser, the author of the 1964 Wilderness Act, stated, "We must remember always that the essential quality of the wilderness is its wildness" (Zahniser, 1992), wilderness is a place dedicated to the wildness of nature and wildness refers to a perceptive quality that is a function of both naturalness and freedom from human control. This suggest that wildness may be a more appropriate term to describe natural landscapes, rather than the controversial notion of wilderness, which often means different things to different people.

Wildness refers to a perceptive quality that is a function of both naturalness and freedom from human control (Aplet et al., 2000). The continuum concept takes respect to the vague nature of wildness and represents it as relative rather than absolute quantification which allows to distinguish areas with superior qualities for conservation efforts. GIS modelling allows to characterise landscapes on a wilderness continuum, and Section 2.4 will describe a specific method for mapping wildness quality in the study area.

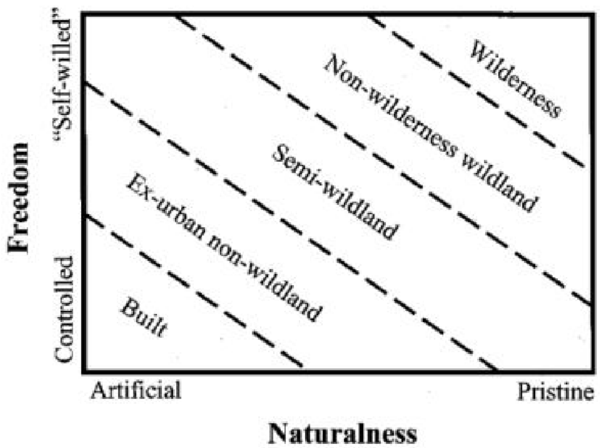


Figure 2.1 The wilderness continuum concept (source: Aplet et al., 2000).

2.3.2 Landscape naturalness and aesthetics

Nature beauty is an integral part of the traditional realm of aesthetics (Bell, 2012), and the concept of naturalness has theoretical support in the Biophilia hypothesis (Kellert and Wilson, 1993), as well as theories about restorative landscapes (Ulrich, 1986; Kaplan and Kaplan, 1989). Empirical studies have highlighted the importance of naturalness in shaping landscape aesthetic preferences (Purcell and Lamb, 1998; Ode et al., 2009). However, the definition of naturalness remains vague and elusive. For example, human-modified environments, such as well-managed parks, manicured gardens, and cultivated fields, might also be perceived as naturalness, and evoke aesthetic experiences. Although these environments adhere to cultural expectations for orderly landscapes, they often tend to be less ecologically functioning (Nassauer, 1995). Conversely, many ecologically significant wildlands, such as tidal wetland or prairies, are consistently attributed to low aesthetic values, as their beauty mainly relies on the hidden riches of health and sustainability (Gobster et al., 2007).

The intricate conceptual relationship between naturalness and aesthetics can be better understood from two perspectives. From an ecological standpoint, naturalness is linked to the ecological integrity of an ecosystem, and an ecologically sound landscape is preferred when considering ethical dimensions (Machado, 2004). Previous studies have also suggested that the ecological knowledge associated with an individual's background is significantly relevant to proper aesthetic appreciation of landscapes at a higher cognitive level (A. Carlson, 2001). However, the scientific basis about ecosystem health is likely to change with scientific advancements (Matthews, 2002).

Perceptually, naturalness is associated with the perceived closeness to a natural state, primarily based on sensory experience rather than a clear biophysical assessment of ecological processes, which can be influenced by cultural norms. Therefore, perceived naturalness may not always align with ecological naturalness (M Tveit et al., 2006; Ode et al., 2008). For example, an intensively managed park with neatly trimmed lawns and flowerbeds may be perceived as highly natural and favoured by people despite containing few or no native species and lacking ecological processes found in natural ecosystems. As outlined above, ecological naturalness is at a higher cognitive level while perceived naturalness is simply of the sensuous surface of nature.

While the concepts of wildness, naturalness, and aesthetics are somehow intertwined, the relationship between landscape wildness and aesthetics, specifically in terms of their spatial patterns, remains empirically unexplored. As previously discussed, some wildlands may be positively perceived as scenic, while others might not evoke the same response. Since these qualities are related to human subjective perception and do not have a specific form, their connections are highly context dependent. There are myriad factors influencing aesthetic preferences for landscapes, such as local landscape features, which can vary significantly across different regions and cultures. As such, it would be worthwhile to further examine the spatial variability in their relationships using an approach that is sensitive to local contexts. This thesis aims to contribute to the existing body of knowledge on landscape aesthetics by exploring the relationships between landscape wildness and landscape aesthetics qualities, and the spatial variability in these relationships.

2.4 Mapping of wildness in the UK

Wilderness has been analysed and mapped across scales, from global to local level (Aplet et al., 2000). Prior qualitative studies into the characterisation of wild spaces, such as Australia's National Wilderness Inventory, have focussed focused on a set of generally accepted attributes of wildness, pertaining to perceived degree of naturalness and remoteness (Lesslie et al., 1993). In the wilderness mapping literature, the methodologies have considered man as a central role in characterising the relativity of the wilderness concept based on measurable objective criteria and using GIS modelling techniques. Researchers typically use available spatial data, such as human infrastructures, land cover, and topography, as proxies for wilderness quality. Expert knowledge is also incorporated to determine the degree of naturalness and ecosystem modification.

Compared to England and Wales, Scotland has relatively unspoilt natural environments in the study area. However, there are no true wilderness areas left in Scotland, only wild lands which are defined as "uninhabited and often relatively inaccessible countryside, where the influence of human activity on the character and quality of the environment has been minimal" (Scottish Natural Heritage, 2002). Despite this, these wild lands still possess certain attributes of wildness. Carver et al. (2012) developed a GIS model for Scottish Natural Heritage (SNH) to map relative wildness across Scotland, using four basic physical attributes: human impact, naturalness, remoteness, and ruggedness, describe in more detail in the following subsections and shown in Figure 2.2 (Carver et al., 2012). These four attributes were identified and further used to define the associated components. These components are weighted according to a uni-modal perception survey and combined using multi-criteria evaluation (MCE) to estimate overall wildness qualities (see Figure 2.3). This GIS-based approach was applied to GB in this research.

2.4.1 Absence of modern human artefacts

The visual perception of wildness was adversely affected by the presence of built features (e.g., buildings and structures), energy infrastructures (e.g., pylons, dams, and wind turbines), and recreational amenities (e.g., off-road tracks, hiking routes, and ski lifts). These man-made elements can create a sense of disturbance that spoil the perception of wildness. A high degree of perceptual human disturbance is likely to result in a low degree of natural conditions in an area. Further, the disturbance allows the unity of the scene to be gauged (G. Fry et al., 2009), which can be closely linked to the visual concept of coherence, associated with one of the Kaplan's information factors in landscape aesthetics and Kaplan, 1989). The lack of coherence in a visual scene can lower its aesthetic preference ratings. Carver and Washtell (2012) have developed a voxel viewshed approach that quantifies the horizontal and vertical proportion of the view obstructed by these human artefacts relative to terrain and distance. This approach provides an informed characterisation of how a person might perceive the relative levels of disturbance within a landscape setting. This can also be used as a visual indicator for landscape aesthetics and applied in landscape assessment and planning.

2.4.2 Perceived naturalness of land cover

From the visual point of view, the perceived naturalness in a landscape is broadly associated with its vegetation and land cover patterns. These patterns are shaped by natural processes and influenced by the level of land management. Land management activities, such as fencing, improved pasture, and stocking rates, and the presence of natural or near-natural vegetation patterns reflects human care for the landscape, which play a role in shaping perceptions of wilderness. The land cover data was categorised into 1 to 5 naturalness classes by utilising ancillary information and expert knowledge (refer to Table 3.1). To ensure accuracy, the reclassification was cross-checked with aerial photography and local knowledge to detect any inconsistencies. The area-weighted mean naturalness score was then computed within a 250-metre radius neighbourhood for each target cell to account for the influence of land cover patterns upon perceived naturalness. This unitless value reflects the overall perceived naturalness in the immediate vicinity of the observer at a given location. As discussed in Section 2.3, this unitless indicator indicates the perceived closeness to a natural state that aligns with the visual concept of naturalness and stewardship, and it may be correspond to the visually aesthetic qualities of landscapes.

2.4.3 Remoteness from mechanised access

Remoteness from mechanised access, that is mainly connected to landscape accessibility, can coincide with the feeling of solitude and tranquillity intrinsic to the wilderness experience. Landscape accessibility can be measured through the walking time from a location to roads to the nearest road access, as expressed in second. Carver and Fritz (1999) have developed anisotropic measures of remoteness using an adaptation of Naismith's rule which assumes differentiated relative traveling time depending on terrain, land cover, and river networks (see Table 3.2). Patterns of remoteness can be mapped as cost surfaces using least-cost (or cost-distance) modelling in GIS that that incorporates both the travelled distance and the traversed costs across landscapes, including the influences of ground cover, relative gradient, and barrier features. Remoteness, in contrast to the other wildness components, does not indicate the visual characteristics of the environment which may not have a significant impact on landscape aesthetics. Its primary importance lies in assessing human access and evoking a sense of areas relatively free from human influence.

2.4.4 Rugged and physically challenging nature of the terrain

According to SNH's definition, wild lands are generally related to rugged terrain and dramatic landscapes formed by natural processes. Thus, the indicator of terrain ruggedness was devised to capture both the topographical variation and the likelihood of encountering harsh weather conditions at higher altitudes (Carver et al., 2012). The ruggedness of terrain, along with the greater wind speeds and lower temperature caused by the increase at altitudes, can contribute to the sense of wildness in a landscape. A 10-metre digital elevation model (DEM) was used to obtain measures of terrain curvature that account for gradient, aspect, and relative relief. The visual ruggedness index is defined as the 2 standard deviation (SD) of terrain curvature within a 250-metre radius of the observer, combined with the altitude data. As with perceived naturalness, the choice of a 250-metre radius corresponds to the immediate landscape that an individual may experience. Ruggedness is related to the visual concept of complexity and may contribute to perceived naturalness, which can evoke aesthetically pleasing perceptions. Hereafter, the four wildness components are simply referred to as absence, naturalness, remoteness, and ruggedness.

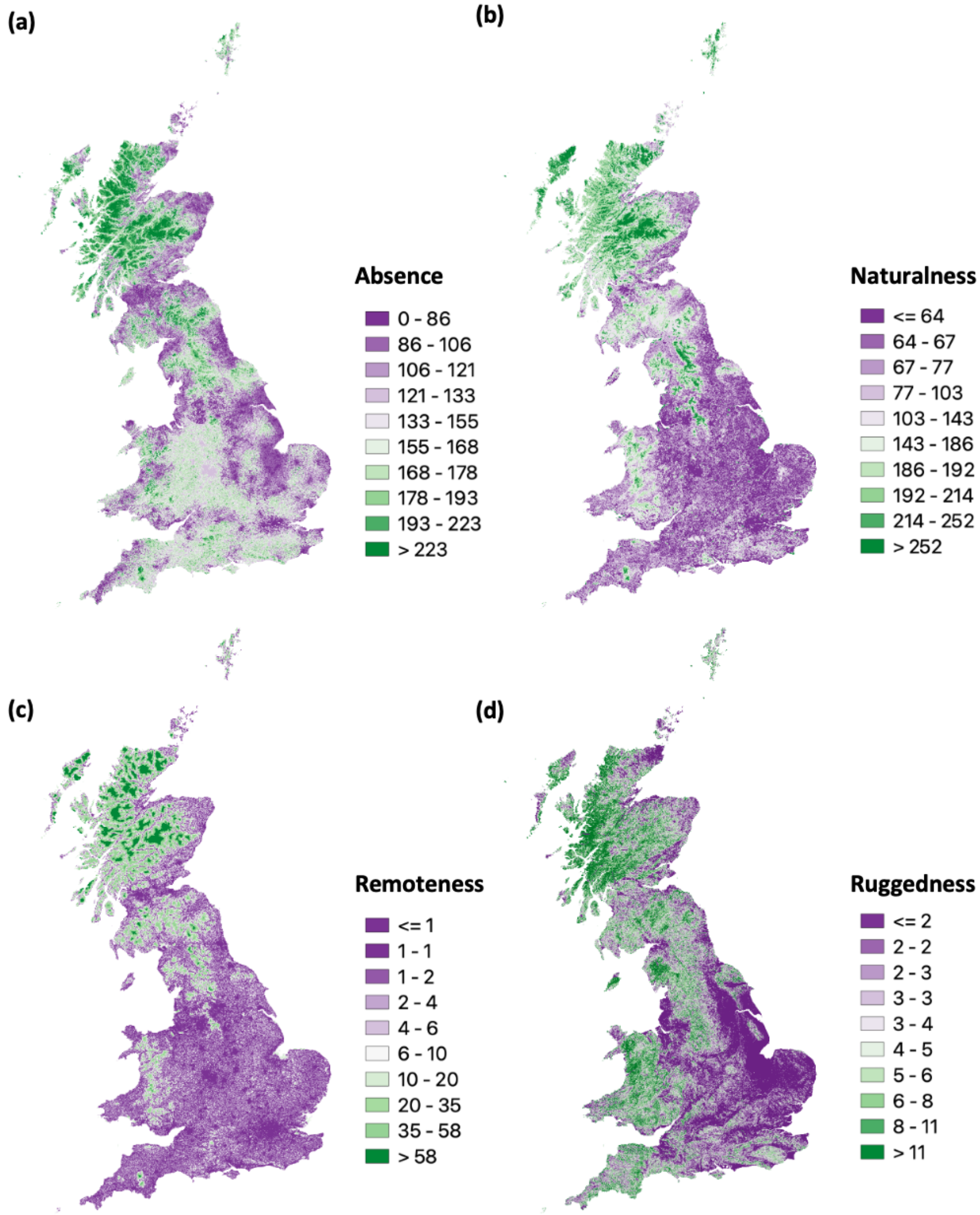


Figure 2.2 Maps of the four wildness components. (a) Absence of modern human artefacts (Absence), (b) Perceived naturalness of land cover (Naturalness), (c) Remoteness from mechanized access (Remoteness), and (d) Rugged and physically challenging nature of the terrain (Ruggedness). These components are rescaled using a 256-level scale, which are displayed using quantile breaks. Reproduced with permission of Dr. Steve Carver.

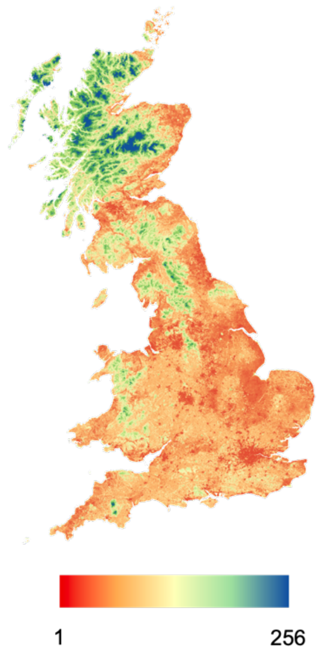


Figure 2.3 Map of the wildness index, derived from an equally weighted combination of the four wildness components, using a multi-criteria evaluation (MCE) approach. Reproduced with permission of Dr. Steve Carver.

Based on the conceptual common ground, the four attributes of wildness are expected to be related to the aesthetic characteristics of landscapes in terms of their spatial patterns. However, there is limited empirical research on the relationships between these cognitive characteristics to date. The emergence of crowdsourced datasets on public scenic perceptions of landscapes provides an opportunity to examine the spatial correlation between wildness and scenic beauty. Therefore, it is worthwhile to explore spatial variations in model calibration results in order to better understand the relationship between wildness characteristics and scenic qualities of landscapes (as described in Chapter 3).

2.5 Landscape character assessments (LCA) in the UK

The ELC recognises landscape 'character' as an essential property of landscape (Van Eetvelde and Antrop, 2009). Landscape character refers to a unique, recognisable and consistent pattern of elements in a landscape that distinguishes it from another/others (rather than better or worse) (Swanwick, 2002). It is not to be confused with landscape quality, which is mainly dependent on the assigned functions including aesthetic, recreational, economic and ecological aspects.

2.5.1 LCA procedures

Character-based approaches, such as Landscape Character Assessment (LCA), originated in the UK and have been adopted in various modified forms in many other countries. LCA is a process of identifying and describing the distinctive characteristics of all landscapes through the development of a hierarchical classification system (Swanwick, 2002). It has become an increasingly formalized tool for understanding and articulating the character of the landscape to facilitate communication in landscape policymaking and management at the national and local government levels.

The LCA is an iterative process, consisting of two main stages: desk study and field survey, which are not separable and refined iteratively in practice. During the desk study, natural and socio-cultural information, as listed in Table 2.4, along with remotely sensed images, are collated and analysed to identify landscape character types and delineate boundaries of landscape character areas using GIS. The characterisation can be conducted at nested and multiple scales depending on the goal of the assessment and the availability of spatial information with the corresponding level of detail. For each landscape character area, additional dimensions of landscape character are examined via field study, including aesthetic and perceptual or experiential factors such as memories, associations, preferences, sensory experiences through sight, sound, smell, touch/feel. Visual characteristics such as colour, texture, pattern, form, scale, enclosure, balance, movement) are also used in assessing the overall aesthetic quality of a landscape and informing management decisions to maintain or enhance its visual appeal.

Table 2.4. Factors likely to be considered at the desk study stage (Tudor, 2014)

	Landscape	**Desk study**
Natural factors	Geology	Geology (solid and drift)
	Landform	Landform/topography
		Geomorphology
	Hydrology	Rivers and drainage
		Water quality and water flows
	Air and climate	Climate
		Microclimate
		Patterns of weather
	Soils	Soils
		Agricultural Land Classification (ALC)
	Land cover/flora and fauna	Habitats/biodiversity

		Land cover
		Vegetation cover
		Tree cover – forest/woodland etc
Socio-cultural factors	Land use (and management)	Land cover
		Agricultural land use
	Settlement	Settlement patterns
		Building types, styles, and materials
		Built structures
	Enclosure	Pattern and type of field enclosure (rural)
		Urban morphology
	Land ownership	Land ownership and tenure
	Time depth	Archaeology and the historic dimension
Cultural associations	Art, literature, descriptive writings, music, myth/legend/folklore, people, events, and associations	Obtained through desk review
Perceptual and aesthetic factors (largely ascertained via field study)	Memories	Obtained via stakeholder engagement
	Associations	
	Perceptions	Some aesthetic factors might be identified as part of the desk study e.g., sense of wildness, remoteness and tranquillity
	Touch/feel	Identified largely via field survey
	Smells/sounds	
	Sight	

2.5.2 National/Regional level LCAs

In the study area, each country has developed its own methodologies for undertaking LCA. Focusing on nationwide coverage and public availability, Table 2.5 summarises the assessments conducted by the three countries at both national/regional and local authority scale. Key differences in these existing LCAs are briefly outlined below.

England and Wales both have compatible assessments at the national/regional scale (1:250,000), specifically National Character Areas (NCAs) in England and National

Landscape Character Areas (NLCAs) in Wales. These assessments capture broad patterns of landscape character variation, providing a broader context for more detailed LCAs. England's NCAs consist of 159 distinct natural areas while Welsh NLCAs comprise 48 individual character areas. These areas are characterised by a unique combination of landscape, biodiversity, geodiversity, history, and cultural and economic activity. In contrast, Scotland lacks a comparable characterisation and assessment approach at this strategic level. Instead, it features a national suite of generic landscape character types (LCTs) that evolved from local authority initiatives and did not follow a prescriptive classification system. Although the existing landscape assessments provide nationwide coverage, their focus on character identification and documentation rather than quality evaluation limits their contribution to national policy on designating high-quality landscapes. Nonetheless, these spatial units are developed into landscape typologies that form a reference basis for future landscape evaluations.

2.5.3 Landscape assessment decision making process (LANDMAP) in Wales

In Wales, all LCAs begin with reference to the Landscape Assessment Decision Making Process, known as LANDMAP, the most comprehensive landscape evidence baseline created through 'top-down' expert assessments and evaluations of various landscape aspects using predetermined typologies. LANDMAP offers a nationally consistent framework for developing regional and local LCAs, ensuring 'top-down' and consistent assessments across the country. Landscape information is compiled at county level (generally by consultants) as well as organised and validated at national level. This approach achieves greater consistency and provides defensible information for both national- and local-scale planning and policy-making. It both classifies and evaluates landscape resources in a hierarchical way based on predetermined typologies, consisting of five themes: geological, ecological, visual and sensory, historic, and cultural aspects of landscape. Each evaluated aspect is delineated as a spatial layer, managed through GIS. For each aspect area identified, specialists develop a field survey record that describes and documents the specific landscape character, features, and qualities with a set of criteria. These aspect layers can be used individually or in combination to interrogate the data to support analysis and decision-making. This thesis concentrates on the visual and sensory aspect, particularly its qualitative evaluation of scenic quality, which is an important consideration for designation. A key feature of this evaluative output is its comparative and overall larger scale (1:2,5000) framework, allowing for the analysis of professional perspectives on landscape visual and perceptual qualities with other compatible datasets. Conversely, England and Scotland lack comparable baseline information that can be consistently applied across both countries. The absence of a national 'top-down' perspective may result in variations in classifications and qualities among different study areas. Therefore, their regional and local LCAs, regarded as 'bottom-up' assessments, lack the same legal status and standardised quality assurance procedures.

Table 2.5. Summary comparison of the national LCA systems in the study area, adapted from (Julie Martin Associates and Swanwick, 2003).

Country	System	Scale	Responsibility	Approach	Principal Units
England	NCAs	National/ regional level, mapped at 1:250,000	Natural England	'Top down' assessment, coordinated by a single consultant	159 Countryside Character Areas
Scotland	LCTs	Detailed, local authority level, mapped at 1:50,000	NatureScot in collaboration with other government agencies and local authorities	'Bottom up' assessments by consultants, later amalgamated on GIS and database without using predetermined typology	394 Landscape Character Types
Wales	NLCA	National/ regional level, mapped at 1:250,000	Natural Resources Wales	'Top down' assessment,	48 individual character areas
	LANDMAP	Hierarchical with a focus on detailed, local authority level, mapped at various scales		'Top down' expert assessment and evaluation of different landscape aspects using predetermined typologies, coordinated and quality-assured by NRW	Level 1, 2, 3 and 4 aspect classifications (1,991 aspect areas for the visual and sensory classification)

2.6 Public participation and crowdsourcing

2.6.1 Public participation according to the ELC

Table 2.6. summarises the ELC's provisions related to public participation. As discussed in Section1.1, one of the core challenges in implementing the ELC is to define landscape quality objectives (LQO) by establishing effective public participation procedures. This is particularly important because experts often formulate LQO without any direct input from the general public. In order to truly reflect the opinions and perceptions of the public, these participation processes should incorporate the public's landscape preferences and values into the definition of LQO.

Table 2.6. Public participation according to the ELC (Jones, 2007).

Article 1 – Definitions a. **"Landscape"** means an area, <u>as perceived by people</u> … c. **"Landscape quality objective"** means, for a specific landscape, the formulation by the competent public authorities of the <u>aspirations of the public</u> with regard to the landscape features of their surroundings.
Article 5 – General measures c. to establish procedures for the participation of the general public, local and regional authorities, and other parties with an interest in the definition and implementation of … landscape policies …
Article 6 – Specific measures **C. Identification and assessment** 1. With the active participation of the interested parties, as stipulated in Article 5.c, and with a view to improving knowledge of its landscapes, each Party undertakes: c. To assess the landscape thus identified, taking into account <u>the particular values assigned to them by the interested parties and the population concerned</u>. **D. Landscape quality objectives** Each party undertakes to define landscape quality objectives for the landscapes identified and assessed, <u>after public consultation</u> in accordance with Article 5.c.

Crowdsourcing has been increasingly harnessed to facilitate participatory planning activities, and to support research on LP&P and CES (Bubalo et al., 2019). Brabham (2013) has defined crowdsourcing more narrowly, as "…an online, distributed problem-solving and production model that leverages the collective intelligence of online communities (i.e., crowds) to serve specific organisational goals" (Brabham, 2013). The scope of crowdsourcing in this thesis extends beyond Brabham's definition to be more inclusive which includes crowdsourced geographic information (Harvey, 2013), the emergence of which is driven by the proliferation of geolocatable devices and the participatory websites with geographical and mapping tools. In theory, using online crowdsourcing techniques enable reaching a large number of people in a relatively short time frame and at limited costs, thereby facilitating wider public involvement. The process relies on the perceptual and cognitive abilities of a large, distributed network of volunteers who participate (especially online) in data production and problem-solving tasks.

2.6.2 Crowdsourcing typologies

Gómez-Barrón et al. (2016) proposed a broad differentiation of crowdsourcing methods. These methods are arranged on a spectrum based on the level of participant engagement. At one end of the spectrum, participants exhibit relative passivity, while at the other end, they engage in a higher level of active contribution, and even extending to proactive action when required (Gómez-Barrón et al., 2016). Along this spectrum, they identified three main levels and their related modes of organising people: non-collaborative participation (contributory), collaboration (collaborative), and co-creation (participatory), as illustrated in Figure 2.4. The first level refers to basic participation where autonomous activities and tasks are accomplished independently from other volunteers' contributions. The second level requires communication and relationship among participants to acquire major complex contributions, focusing on collaborative operation. Finally, the most active level of crowdsourcing engagement is through projects that facilitate participatory processes, providing individuals the opportunity to actively decide how the project will be conducted and define necessary outcomes. The three modes of organisation correspond to different sets of motives for contributing. Also, higher levels of involvement or engagement are correlated with the increasing use of volunteers' cognitive abilities, enabling them to tackle more complex problems and tasks.

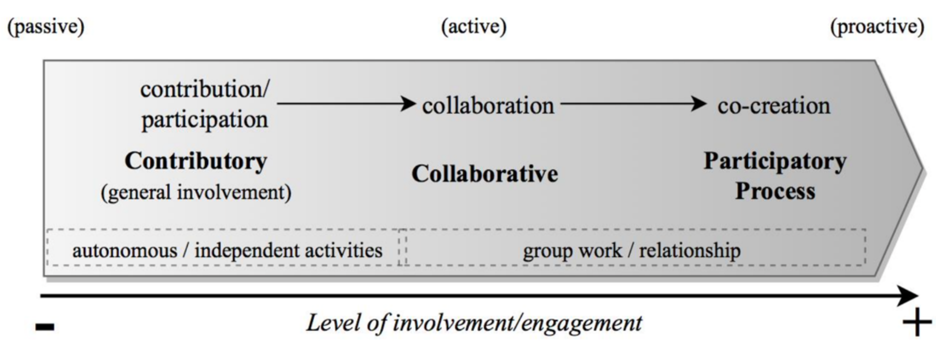


Figure 2.4 Volunteers' level of involvement/engagement and related modes of organisation (source: Gómez-Barrón et al., 2016).

This thesis focuses on landscape perception, considering perception as a special cognitive instrument and landscape as a special cognitive object. The cognitive process can be categorised into four typologies: perceptual, expressive, interpretative, and symbolic, as detailed in Table 2.7. The perceptual and expressive levels of cognitive process align with evolutionary theories, emphasising the commonality in human landscape preferences. In contrast, the interpretative and symbolic levels align with cultural theories, focusing on the diversity of perceptions shaped by individual characteristics. Moreover, the perceptual and interpretative levels contribute to a landscape's narrative function, while the expressive and symbolic levels contribute to its poetic function (Nohl, 2001), as depicted in Figure 2.5. This thesis considers the lower levels of the cognitive process, including the perceptual and interpretative levels.

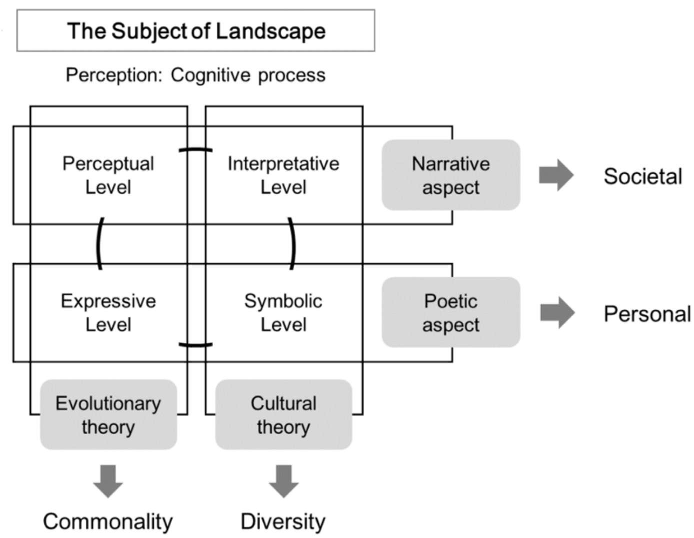


Figure 2.5 Types of the cognitive process of landscape perception, adapted from (Lee and Son, 2017).

Table 2.7. Cognitive typologies of landscape perception (Lee and Son, 2017).

Cognitive Process	**Description**
Perceptual	The beholder of a landscape immediately captures relevant information through the senses (e.g., viewing, hearing, touching, or smelling).
Expressive	All perceived elements and compositions are associated with the beholder's feelings and emotions.
Interpretative	The beholder already understands and interprets the landscape as signs or symptoms. For instance, a sandbank may signify the river's low water power.

Symbolic	Landscape realities become ideas, imaginations, utopian images, which are generated in the head of the beholder.

2.6.3 Crowdsourced geographic information

Crowdsourced geographic information (CGI) has been proved to effectively capture how people perceive and interact with landscapes (Dunkel, 2015), serving as a valuable resource for quantifying and mapping landscape values and CES (Tenerelli et al., 2016; Boris T. van Zanten et al., 2016). The Geograph Britain and Ireland project (https://www.geograph.org.uk/) is an example, inviting people to contribute geographically representative photographs and information for every square kilometre with a nationwide scope. These geolocated photographs collected with awareness or permission of contributors are often viewed as volunteered geographic information (VGI) which refers to intentionally created and shared data (Michael F. Goodchild, 2007).

A number of Geograph photographs were further used by a web-based crowdsourcing campaign, Scenic-Or-Not, to gather people's scenic perceptions of landscapes across Great Britain. Until February 2015, over 1.5 million ratings have been collected for 212,212 geotagged photos available on (http://scenicornot.datasciencelab.co.uk/faq). The campaign employed a spatially even sampling to ensures the even distribution of landscape photos, covering nearly 80% of GB (see Figure 2.6). Participants are asked to rate randomly presented photos for their scenic beauty on an integer scale from 1 (the least scenic) to 10 (the most scenic) without immediately knowing their locations. This dataset provides an otherwise unavailable measure of landscape scenic beauty, capturing the perceptions of the broader public at the national level, which have been utilised to understand the impact of environmental aesthetics on human health (Seresinhe et al., 2015) and happiness (Seresinhe et al., 2019). Research has also used this dataset to verify the estimation of landscape scenic beauty based on social media and OpenStreetMap data (Seresinhe et al., 2018), as well as to train a deep neural network for extracting scenic features and predicting the beauty of scenes for new places (Seresinhe et al., 2017).

The photographic ratings with location information allow for assessing the geographic variation in publicly perceived visual landscape quality (hereafter referred to as 'scenicness'). This CGI data inherently reflects people's subjective emotions, opinions, and values related to landscapes and/or ecosystems, potentially facilitating a deeper understanding of the correlations, similarities and differences in perspectives between experts and the public. The granularity and spatial coverage of this data compensate for the absence of national or regional perception surveys, making it suitable for spatial modelling and mapping in response to the ELC's policy.

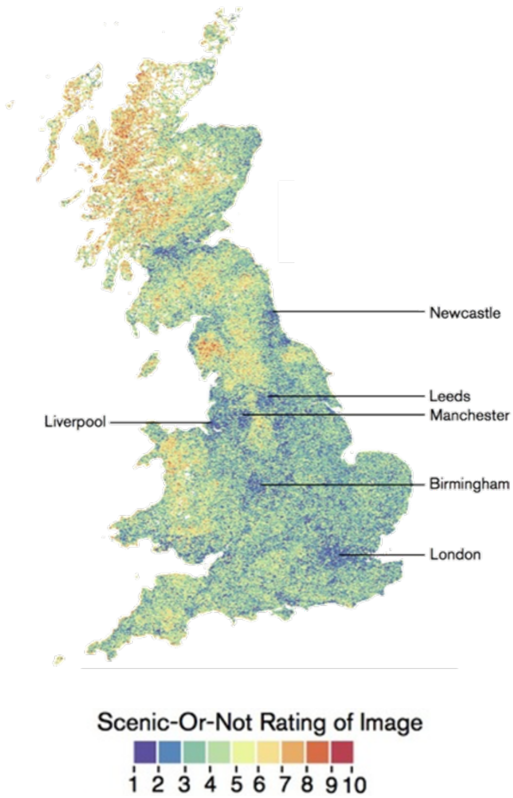


Figure 2.6 Choropleth map showing the scenic ratings across Great Britain. Blue regions mark less scenic areas, mainly around major cities, while more scenic areas, indicated in red, appear in the Scottish Highlands and Northern England. Data scarcity is noted in the Highlands. Image adapted from Journal of Urban Design and Mental Health (2016), retrieved from (https://www.urbandesignmentalhealth.com/journal1-beautifulplacesandwellbeing.html).

2.7 Advanced modelling methods for integrating objectivist and subjectivist landscape assessments

2.7.1 Spatial autocorrelation and heterogeneity

In the field of landscape research, numerous efforts have been made to relate landscape features to preferences (Schirpke et al., 2013; Tenerelli et al., 2016; van Zanten et al., 2016; Zoderer et al., 2019). Global regression techniques, such as multiple linear regression (MLR), have been used to determine the relationships between spatial variables representing landscape features and their corresponding values or qualities (Schirpke et al., 2013). Nevertheless, these non-spatial models are incapable of integrating the spatial dependence structures often observed in geospatial data into their model specification. Consequently, they fail to reveal local-specific relationships or spatial autocorrelation in model residuals. Earlier developments in spatial regression modelling primarily focused on parametric models, including the spatial lag models (SLM), spatial error models (SEM), and spatial Durbin models (SDM), that effectively account for stationary autocorrelation effects (Anselin, 2002).

However, these global models fall short in accounting for spatial heterogeneity or non-stationarity, which have relatively recently been considered when working with spatial data (LeSage and Pace, 2009). Both spatial autocorrelation and spatial heterogeneity, with respect to processes and functions, have long been acknowledged as important effects in the landscape ecology literature (Dale and Fortin, 2014). These spatial effects are also expected in perceptual landscapes but received limited attention in landscape perception and preference studies.

2.7.2 Geographically weighted (GW) framework

Geographically Weighted Regression (GWR) has been increasingly employed as an exploratory tool in spatial analyses of social and environmental data, accounting for both spatial dependence and heterogeneity (Fotheringham et al., 2003). It involves the spatial disaggregation of a classical regression model, in which separate models are calibrated locally by "borrowing data" from nearby locations and weighting these data based on their distance from the observation unit. The resulting GWR estimates can be mapped to facilitate targeted land and ecosystem management efforts (Tenerelli et al., 2016).

Despite its advantages, several potential weaknesses of GWR have been reported in the literature. Firstly, there is an issue of correlation among the coefficient estimates (Wheeler and Tiefelsdorf, 2005). Specifically, high correlation hinders the separation of individual variables effects from those of other variables. Hence, the GWR techniques do not necessarily establish any real cause-effect relationships. Secondly, GWR can be highly susceptible to multicollinearity and overfitting, especially when dealing with sample sizes (Páez et al., 2011). In light of this concern, extended GWR approaches, such as ridge GWR (Wheeler, 2007) and GW lasso (Wheeler, 2009), are devised to remedy the multicollinearity issue. However, Fotheringham and Oshan (2016) asserted, based on a simulation study with 2500 spatial units, that GWR is robust to the multicollinearity issue when the sample size is sufficiently large (A. Stewart Fotheringham and Oshan, 2016). Lastly, a key step to GWR is the optimisation of bandwidth, determined through a measure of model fit, which defines the variation in the local outputs (i.e., the degree of smoothing). However, the use of a "one-size-fit-all" bandwidth across all the covariates in GWR relies on the implicit assumption that each investigated process operates at the same spatial scale, which is unrealistic in a real-world case.

Multiscale geographically weighted regression (MGWR) overcomes this limitation of the GWR framework by allowing separate bandwidths to be estimated for each covariate. These predictor-specific bandwidths enhance the process understanding by revealing the scale of spatial heterogeneity associated with each covariate's relationship to the response. Allowing bandwidth to be distinct for each relationship can also avoid inducing multicollinearity (T.M. Oshan et al., 2019). In theory, MGWR provides a more intuitive and accurate model of real-world processes than GWR, and as such, has been suggested as the default GWR (Comber et al., 2023).

2.7.3 Alternative modelling approaches

On one hand, there exist long-standing objections to geographically weighted modelling, rooted in the view that spatial dependencies in data are a statistical "nuisance" that need to be corrected (Harris, 2019). Further critiques concern the presence of local cluster outliers in the global model, suggesting that any observed spatial effects may indicate the omission of key predictors (McMillen, 2003). If the underlying process is not adequately represented by the model inputs, one remedy is to simply incorporate additional explanatory variables. This might also imply that the theoretical understanding of the investigated process is lacking.

On the other hand, researchers still prefer to use alternative modelling approaches, such as mixed-effect models and Bayesian models, which are considered to provide more robust statistical inference and are commonly seen in landscape and Ecosystem Services (ES) studies. For example, van Zanten et al. (2016) employed a generalised mixed effects (GME) model to estimate the spatial patterns of landscape values using geolocated social media posts as response variables and both landscape features and socioeconomic factors as predictors. They removed spatial autocorrelation by applying eigenvector spatial filtering (Murakami and Griffith, 2019), ensuring a more accurate estimation of the relationships (Boris T. van Zanten et al., 2016). Zoderer et al. (2019) utilised a cumulative link mixed model (CLMM) to explore the casual relationship between different stakeholder groups' perception of ES supply, measured on an ordinal scale, and the spatial information on landscape features (Zoderer et al., 2019). Advanced spatial statistical techniques, such as conditional autoregressive (CAR) models, are also used to account for spatial dependence by including random effects that represent spatially structured random components (Seresinhe et al., 2015; Seresinhe et al., 2018). However, these models may not effectively capture local heterogeneity in data relationship.

The persistence of the GW framework can be attributed to its practical rather than theoretical merits, as it not only serves as an exploratory tool for identifying spatial heterogeneity, but also helps in detecting potential model misspecifications (Comber et al., 2023).

2.7.4 Machine learning modelling

In order to effectively conduct predictive tasks of complex phenomena and processes, such as landscape scenic quality, it is crucial to consider not only spatial effects, but also non-linearity and factor interdependence. This consideration is particularly important when integrating potentially correlated objective and subjective assessments of landscape scenic quality, as explored in the present thesis. Alternative modelling approaches, such as machine learning and, more broadly, artificial intelligence (ML/AI) techniques, require fewer assumptions about the investigated processes, and might be more suitable for capturing complex non-linear interactions and functional forms of relationships from a large amount of data (Bzdok et al., 2018).

ML/AI techniques hold the potential to provide significant benefits, including improved model performance, enhanced representational flexibility, and effective handling of noisy data (Bishop, 1996). XGBoost (eXtreme Gradient Boosting) is one of the most widely used ML algorithms to solve classification or regression problems (Chen and Guestrin, 2016). Studies have demonstrated that well-tuned XGBoost often outperforms alternatives methods (e.g., random forest or deep neural networks) (Joharestani et al., 2019; Shwartz-Ziv and Armon, 2022). Additionally, XGBoost was evidently capable of accounting for possible spatial effects with the aid of a model agnostic explainer, SHAP (SHapley Additive exPlanations) (Lundberg and Lee, 2017) in a simulation study (Li, 2022). The combination of XGBoost and SHAP can also be used to guide the choice of statistical models for rigorous inference, which may help to reinforce some conclusions drawn from statistical models while providing additional insights through differences in results. However, the applications of these techniques remain unexplored in landscape assessments. The primary purpose of this thesis is to integrate potentially correlated objective and subjective assessments of landscape scenic quality. The XGBoost model was used to account for non-linearity effects, and then to map aesthetic values of landscapes across Great Britain (see Chapter 5).

2.8 Research questions

A literature review highlights the need to reconcile expert and public perspectives, as well as objective and subjective approaches in landscape assessments. In recent years, the use of crowdsourced geographic information has significantly increased, offering potential solutions to the lack of large-scale, perception-based landscape evaluations across extensive geographic areas. The hypothesis is that these subjective measures may complement the objective measures of landscape qualities to produce more informed and inclusive evaluation outcomes, albeit with the caveats associated with their qualities and representativeness (Bubalo et al., 2019). The following research questions aim to explore the pragmatic benefits of integrating crowdsourced geotagged data into national-scale landscape assessments:

- RQ. 1: How are crowdsourced perceptions of scenic beauty associated with expert-based measures relevant to landscape wildness quality, and at what scale? (Chapter 3)

- RQ. 2: To what extent do people's photographic ratings for landscape scenic beauty correspond with expert-led, character-based evaluations of scenic quality? (Chapter 4)

- RQ. 3: Can the integration of the subjective perceptions, objective assessments, and character-based evaluations mentioned above be used to effectively map landscape scenic quality? (Chapter 5)

2.9 Conceptual framework

An overarching conceptual framework was developed to guide the analyses in this thesis, in response to increasing policy demands at the national level (Figure 2.7). This framework is grounded in the assumption of a high degree of commonality in human landscape preferences, as emphasised by evolutionary theories (see Section 2.2.1), and the integration of objective and subjective landscape in response to the growing trend towards integrated methods (Medeiros et al., 2021). The framework primarily encompasses three spatial datasets: Scenic-Or-Not, Wildness, and LANDMAP, each representing either expert or public, and objective or subjective, and quantitative or qualitative assessments. The underlying assumption is that spatial association between these three datasets likely exist at a comparable extent and resolution as they are partially informed by landscape physical characteristics (e.g., landform, land cover and land use), albeit with involving certain level of subjectivity.

Chapter 3 explores the spatial association between the four wildness components (i.e., absence, naturalness, remoteness, and ruggedness) and perceived scenicness, as they share certain underlying concepts in common. This exploration also includes the investigation of the spatial variation in these relationships and their spatial scales of processes. Chapter 4 further explores the similarities and differences in subjective perceptions between experts and the public, based on the LANDMAP visual and sensory LCTs and LCAs. The underlying premise is that photographic scenic ratings, which are more closely linked to ground-level perceptions, can complement the objective characterisation of a landscape's perceptual qualities, which are primarily reliant on bird's-eye views. Chapter 5 combines both the wildness index and Scenic-Or-Not ratings to construct predictive models of LANDMAP scenic quality classes, which are, in turn, used to effectively map landscape scenic quality. The main objective is to present a nationwide, integrative landscape assessment, illustrating the variations in scenic quality, in line with the LANDMAP's evaluative levels across Great Britain, while maintaining an appropriate balance between objective and subjective assessments.

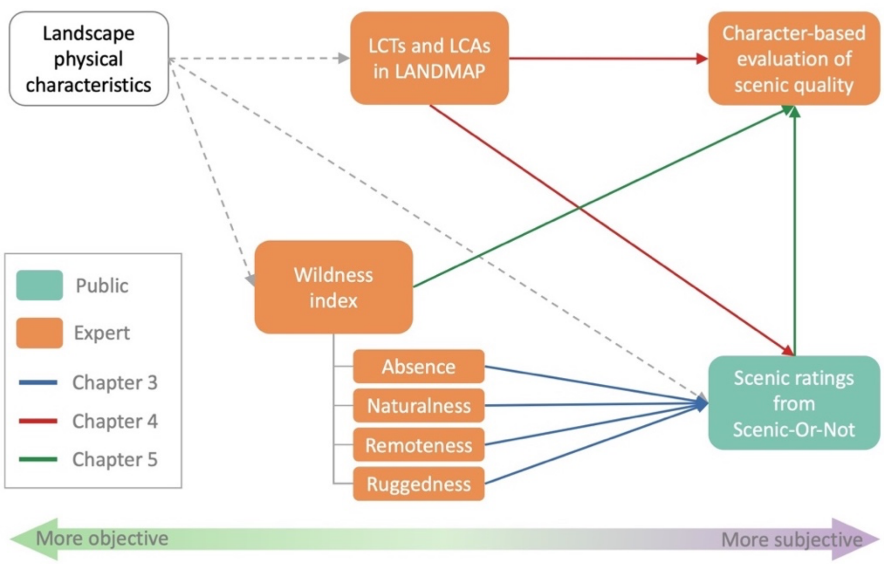


Figure 2.7 Conceptual framework.

References

Adevi, A.A. and Grahn, P. 2012. Preferences for Landscapes: A Matter of Cultural Determinants or Innate Reflexes that Point to Our Evolutionary Background? *Landscape Research*. **37**(1), pp.27–49.

Anselin, L. 2002. Under the hood Issues in the specification and interpretation of spatial regression models. *Agricultural Economics*. **27**(3), pp.247–267.

Aplet, G., Thomson, J. and Wilbert, M. 2000. Indicators of wildness: Using attributes of the land to assess the context of wilderness *In*: S. F. McCool, D. N. Cole, W. T. Borrie and J. O'Loughlin, eds. *USDA Forest Service Proceedings RMRS-P-15-VOL-2*. Missoula, Montana., pp.89–98.

Appleton, J. 1975. *The Experience of Landscape*. London: Wiley Chichester.

Bell, S. 2012. *Landscape: Pattern, Perception and Process*. Routledge.

Bishop, I.D. 1996. Comparing regression and neural net based approaches to modelling of scenic beauty. *Landscape and Urban Planning*. **34**(2), pp.125–134.

Bourassa, S.C. 1990. A Paradigm for Landscape Aesthetics. *Environment and Behavior*. **22**(6), pp.787–812.

Brabham, D. 2013. *CrowdsourCing*. The MIT Press.

Brady, E. and Prior, J. 2020. Environmental aesthetics: A synthetic review. *People and Nature*. **2**(2), pp.254–266.

Bubalo, M., van Zanten, B.T. and Verburg, P.H. 2019. Crowdsourcing geo-information on landscape perceptions and preferences: A review. *Landscape and Urban Planning*. **184**, pp.101–111.

Bzdok, D., Krzywinski, M. and Altman, N. 2018. Machine learning: supervised methods. *Nature Methods*. **15**(1).

Carlson, A. 2001. Aesthetic preferences for sustainable landscapes: seeing and knowing. *In*: *Forests and landscapes: linking ecology, sustainability and aesthetics*.

Carver, S., Comber, A., McMorran, R. and Nutter, S. 2012. A GIS model for mapping spatial patterns and distribution of wild land in Scotland. *Landscape and Urban Planning*. **104**(3–4), pp.395–409.

Carver, S. and Fritz, S. 1999. Mapping remote areas using GIS *In*: M. B. Usher, ed. *Landscape character: Perspectives on management and change*. Natural Heritage of Scotland Series, HMSO.

Chen, T. and Guestrin, C. 2016. XGBoost: A scalable tree boosting system *In*: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., Lü, Y., Murakami, D., Nakaya, T., Wang, Y. and Harris, P. 2023. A Route Map for Successful Applications of Geographically Weighted Regression. *Geographical Analysis*. **55**(1), pp.155–178.

Conrad, E., F. Cassar, L., Jones, M., Eiter, S., Izaovičová, Z., Barankova, Z., Christie, M. and Fazey, I. 2011. Rhetoric and Reporting of Public Participation in Landscape Policy. *Journal of Environmental Policy & Planning*. **13**(1), pp.23–47.

Council of Europe 2000. *European Landscape Convention*. Florence.

Dale, M.R.T. and Fortin, M.J. 2014. *Spatial analysis: A guide for ecologists, second edition*.

Daniel, T.C. 2001. Whither scenic beauty? Visual landscape quality assessment in the 21st century. *Landscape and Urban Planning*. **54**(1–4), pp.267–281.

Daniel, T.C. and Vining, J. 1983. Methodological Issues in the Assessment of Landscape Quality *In*: I. Altman and J. F. Wohlwill, eds. *Behavior and the Natural Environment*. Boston, MA: Springer US, pp.39–84.

Dramstad, W.E., Fry, G., Fjellstad, W.J., Skar, B., Helliksen, W., Sollund, M.L., Tveit, M.S., Geelmuyden, A.K. and Framstad, E. 2001. Integrating landscape-based values - Norwegian monitoring of agricultural landscapes. *Landscape and Urban Planning*. **57**(3–4), pp.257–268.

Dramstad, W.E., Tveit, M.S., Fjellstad, W.J. and Fry, G.L.A. 2006. Relationships between visual landscape preferences and map-based indicators of landscape structure. *Landscape and Urban Planning*. **78**(4), pp.465–474.

Dubey, A., Naik, N., Parikh, D., Raskar, R. and Hidalgo, C.A. 2016. Deep Learning the City: Quantifying Urban Perception at a Global Scale *In*: B. Leibe, J. Matas, N. Sebe and M. Welling, eds. *Eccv*. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp.196–212.

Dunkel, A. 2015. Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning*. **142**, pp.173–186.

Van Eetvelde, V. and Antrop, M. 2009. A stepwise multi-scaled landscape typology and characterisation for trans-regional integration, applied on the federal state of Belgium. *Landscape and Urban Planning*. **91**(3), pp.160–170.

Fairclough, G. 2006. A new landscape for cultural heritage management: Characterisation as a management tool. *Landscapes Under Pressure: Theory and Practice of Cultural Heritage Research and Preservation*., pp.55–74.

Fairclough, G.J., McNab, A., Chadwick, P., Lambrick, G. and Heritage, E. 1999. *Yesterday's world, tomorrow's landscape: the English Heritage Historic Landscape Project 1992-94*.

Fotheringham, A.S., Brunsdon, C. and Charlton, M. 2003. *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons.

Fotheringham, A.S. and Oshan, T.M. 2016. Geographically weighted regression and multicollinearity: dispelling the myth. *Journal of Geographical Systems*. **18**(4).

Fry, G, Tveit, M.S., Ode, Å. and Velarde, M.D. 2009. The ecology of visual landscapes: Exploring the conceptual common ground of visual and ecological landscape indicators. *Ecological Indicators*. **9**(5), pp.933–947.

Gobster, P.H. 1999. An Ecological Aesthetic for Forest Landscape Management. *Landscape Journal*., pp.54–64.

Gobster, P.H. 2008. Yellowstone Hotspot: Reflections on Scenic Beauty, Ecology, and the Aesthetic Experience of Landscape. *Landscape Journal*. **27**(2), pp.291–308.

Gobster, P.H., Nassauer, J.I., Daniel, T.C. and Fry, G. 2007. The shared landscape: what does aesthetics have to do with ecology? *Landscape Ecology*. **22**(7), pp.959–972.

Gobster, P.H., Ribe, R.G. and Palmer, J.F. 2019. Themes and trends in visual assessment research: Introduction to the Landscape and Urban Planning special collection on the visual assessment of landscapes. *Landscape and Urban Planning*. **191**(August), p.103635.

Gómez-Barrón, J.P., Manso-Callejo, M.Á., Alcarria, R. and Iturrioz, T. 2016. Volunteered geographic information system design: Project and participation guidelines. *ISPRS International Journal of Geo-Information*. **5**(7).

Goodchild, M.F. 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal*. **69**(4), pp.211–221.

Harris, P. 2019. A Simulation Study on Specifying a Regression Model for Spatial Data: Choosing between Autocorrelation and Heterogeneity Effects. *Geographical Analysis*. **51**(2).

Harvey, F. 2013. To Volunteer or to Contribute Locational Information? Towards Truth in Labeling for Crowdsourced Geographic Information *In*: *Crowdsourcing Geographic*

Knowledge [Online]. Dordrecht: Springer Netherlands, pp.31–42. Available from: https://link.springer.com/10.1007/978-94-007-4587-2_3.

Joharestani, M.Z., Cao, C., Ni, X., Bashir, B. and Talebiesfandarani, S. 2019. PM2.5 prediction based on random forest, XGBoost, and deep learning using multisource remote sensing data. *Atmosphere*. **10**(7).

Jones, M. 2007. The European Landscape Convention and the question of public participation. *Landscape Research*. **32**(5), pp.613–633.

Julie Martin Associates and Swanwick, C. 2003. *Overview of Scotland's National Programme of Landscape Character Assessment* [Online]. [Accessed 31 March 2023]. Available from: https://www.nature.scot/doc/naturescot-commissioned-report-29-overview-scotlands-national-programme-landscape-character.

Kaplan, R. and Kaplan, S. 1989. *The experience of nature: A psychological perspective.* New York, US: Cambridge University Press.

Kellert, S.R. and Wilson, E.O. 1993. *The Biophilia Hypothesis* [Online]. Washington, DC.: Island Press. Available from: http://journals.sagepub.com/doi/10.1177/027046769501500125.

de la Fuente de Val, G., Atauri, J.A. and de Lucio, J. V. 2006. Relationship between landscape visual attributes and spatial pattern indices: A test study in Mediterranean-climate landscapes. *Landscape and Urban Planning*. **77**(4), pp.393–407.

Lee, K.C. and Son, Y.H. 2017. Exploring landscape perceptions of Bukhansan National Park according to the degree of visitors' experience. *Sustainability (Switzerland)*. **9**(8).

LeSage, J. and Pace, R.K. 2009. *Introduction to spatial econometrics*.

Lesslie, R.G., Taylor, D. and Maslen, M.A. 1993. *National Wilderness Inventory: Handbook of principles, procedures and usage*. Australian Government Pub. Service.

Li, Z. 2022. Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost. *Computers, Environment and Urban Systems*. **96**, p.101845.

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), pp.177–198.

Lundberg, S.M. and Lee, S.I. 2017. A unified approach to interpreting model predictions *In*: *Advances in Neural Information Processing Systems*.

Machado, A. 2004. An index of naturalness. *Journal for Nature Conservation*. **12**(2), pp.95–110.

McMillen, D.P. 2003. Spatial autocorrelation or model misspecification. *International Regional Science Review*. **26**(2).

Medeiros, A., Fernandes, C., Gonçalves, J.F. and Farinha-Marques, P. 2021. Research trends on integrative landscape assessment using indicators – A systematic review. *Ecological Indicators*. **129**.

Meinig, D. 1976. The Beholding Eye - Ten Versions of the Same Scene. *Landscape Architecture*. **66**(1), pp.47–54.

Murakami, D. and Griffith, D.A. 2019. Eigenvector Spatial Filtering for Large Data Sets: Fixed and Random Effects Approaches. *Geographical Analysis*. **51**(1), pp.23–49.

Nassauer, J.I. 2011. Care and stewardship: From home to planet. *Landscape and Urban Planning*. **100**(4).

Nassauer, J.I. 1995. Messy Ecosystems, Orderly Frames. *Landscape Journal*. **14**(2), pp.161–170.

Nohl, W. 2001. Sustainable landscape use and aesthetic perception–preliminary reflections on future landscape aesthetics. *Landscape and Urban Planning*. **54**(1–4), pp.223–237.

Ode, Å., Fry, G., Tveit, M.S., Messager, P. and Miller, D. 2009. Indicators of perceived naturalness as drivers of landscape preference. *Journal of Environmental Management*. **90**(1), pp.375–383.

Ode, Å., Tveit, M. and Fry, G. 2008. Capturing landscape visual character using indicators: Touching base with landscape aesthetic theory. *Landscape Research*. **33**(1), pp.89–117.

Ohta, H. 2001. *A phenomenological approach to natural landscape cognition*.

Orians, G.H. 1980. Habitat selection: General theory and applications to human behavior *In*: J. S. Lockard, ed. *The evolution of human social behavior*. Chicago: Elsevier, pp.49–66.

Oshan, T.M., Li, Z., Kang, W., Wolf, L.J. and Stewart 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).

Páez, A., Farber, S. and Wheeler, D. 2011. A simulation-based study of geographically weighted regression as a method for investigating spatially varying relationships. *Environment and Planning A*. **43**(12).

Prieur, M. 2006. Landscape and social, economic, cultural and ecological approaches *In*: *AA. VV., Landscape and sustainable development. Challenges of the European Landscape Convention, Council of Europe Publishing*., pp.11–28.

Purcell, A.T. and Lamb, R.J. 1998. Preference and naturalness: An ecological approach. *Landscape and Urban Planning*. **42**(1).

Schirpke, U., Tappeiner, G., Tasser, E. and Tappeiner, U. 2019. Using conjoint analysis to gain deeper insights into aesthetic landscape preferences. *Ecological Indicators*. **96**(January 2018), pp.202–212.

Schirpke, U., Tasser, E. and Tappeiner, U. 2013. Predicting scenic beauty of mountain regions. *Landscape and Urban Planning*. **111**(1), pp.1–12.

Scottish Natural Heritage, S.N.H. 2002. Policy Statement: Wildness in Scotland's Countryside.

Seresinhe, C.I., Moat, H.S. and Preis, T. 2018. Quantifying scenic areas using crowdsourced data. *Environment and Planning B: Urban Analytics and City Science*. **45**(3).

Seresinhe, C.I., Preis, T., MacKerron, G. and Moat, H.S. 2019. Happiness is Greater in More Scenic Locations. *Scientific Reports*. **9**(1).

Seresinhe, C.I., Preis, T. and Moat, H.S. 2015. Quantifying the Impact of Scenic Environments on Health. *Scientific Reports*. **5**(1), p.16899.

Seresinhe, C.I., Preis, T. and Moat, H.S. 2017. Using deep learning to quantify the beauty of outdoor places. *Royal Society Open Science*. **4**(7).

Sevenant, M. and Antrop, M. 2010. The use of latent classes to identify individual differences in the importance of landscape dimensions for aesthetic preference. *Land Use Policy*. **27**(3), pp.827–842.

Shwartz-Ziv, R. and Armon, A. 2022. Tabular data: Deep learning is not all you need. *Information Fusion*. **81**.

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

Taylor, J.G., Zube, E.H. and Sell, J.L. 1990. Landscape assessment and perception research methods *In*: R. B. Bechtel, R. W. Marans and W. M. Michelson, eds. *Methods in environmental and behavioral research*. Malabar, Fla. : R.E. Krieger Pub. Co., pp.361–393.

Tenerelli, P., Demšar, U. and Luque, S. 2016. Crowdsourcing indicators for cultural ecosystem services: A geographically weighted approach for mountain landscapes. *Ecological Indicators*. **64**, pp.237–248.

Tuan, Y.-F. 1990. *Topophilia: A study of environmental perception, attitudes, and values*. Columbia University Press.

Tudor, C. 2014. An approach to landscape character assessment.

Tveit, M., Ode, Å. and Fry, G. 2006. Key concepts in a framework for analysing visual landscape character. *Landscape Research*. **31**(3), pp.229–255.

Tveit, M.S. 2009. Indicators of visual scale as predictors of landscape preference; a comparison between groups. *Journal of Environmental Management*. **90**(9), pp.2882–2888.

Tveit, M.S., Ode Sang, Å. and Hagerhall, C.M. 2018. Scenic Beauty *In*: *Environmental Psychology*. Chichester, UK: John Wiley & Sons, Ltd, pp.45–54.

Ulrich, R.S. 1986. Human responses to vegetation and landscapes. *Landscape and Urban Planning*. **13**, pp.29–44.

Warren, C.R. 2009. Wilderness *In*: *International Encyclopedia of Human Geography*. Elsevier, pp.254–259.

Wheeler, D. and Tiefelsdorf, M. 2005. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems*. **7**(2).

Wheeler, D.C. 2007. Diagnostic tools and a remedial method for collinearity in geographically weighted regression. *Environment and Planning A*. **39**(10).

Wheeler, D.C. 2009. Simultaneous coefficient penalization and model selection in geographically weighted regression: The geographically weighted lasso. *Environment and Planning A*. **41**(3).

Zahniser, H. 1992. *Where wilderness preservation began: Adirondack writings of Howard Zahniser*. North Country Books Incorporated.

van Zanten, B.T., Van Berkel, D.B., Meentemeyer, R.K., Smith, J.W., Tieskens, K.F. and Verburg, P.H. 2016. Continental-scale quantification of landscape values using social media data. *Proceedings of the National Academy of Sciences*. **113**(46), pp.12974–12979.

Zoderer, B.M., Tasser, E., Carver, S. and Tappeiner, U. 2019. An integrated method for the mapping of landscape preferences at the regional scale. *Ecological Indicators*. **106**, p.105430.

Zube, E.H., Sell, J.L. and Taylor, J.G. 1982. Landscape perception: Research, application and theory. *Landscape Planning*. **9**(1), pp.1–33.

Chapter 3 Linking wildness with scenicness

Abstract

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.

Keywords: landscape character assessments; crowdsourcing; wildness; Scenic-Or-Not

3.1 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 et al., 2019). However, aesthetic preferences vary widely across social and cultural contexts (Zube and Pitt, 1981; Dramstad et al., 2006), 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 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 (Conrad, Christie, et al., 2011; Butler and Berglund, 2014). 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 et al., 2016), values (Boris T 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 et al., 2015) and happiness (Seresinhe et al., 2019), enabling a clearer understanding of public perceptions regarding landscape composition and scenic beauty (Seresinhe et al., 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 et al., 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 et al., 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 (Fritz et al., 2000; Carver et al., 2002; Comber et al., 2010; Carver and Washtell, 2012). 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 et al., 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 (Schirpke et al., 2013; Frank et al., 2013; Boris T 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 et al., 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 (Luo and Wei, 2009; Y.-F. Su et al., 2012; Hong and Jeon, 2017; Sun et al., 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 (Cleveland, 1979; Bowman, 1984; Brunsdon et al., 1996). 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 (Yang, 2014; A Stewart Fotheringham et al., 2017), 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.

3.2 Data and methods

3.2.1 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 et al., 2016; Oteros-Rozas et al., 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.

3.2.2 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 and 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 3.1). 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 3.2 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 3.1 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

Table 3.2 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°)	+10 mins/100 m of ascent	$T = a * \Delta S + 6 * \Delta \mathrm{H}$	Naismith's rule
		Slight downhill (-5° < slope < 0°)	5	$T = a * \Delta S$	Langmuir's correction
		Moderate downhill (-12° < slope < -5°)	−10 min/300 m of descent	$T = a * \Delta S + 2 * \Delta \mathrm{H}$	
		Steep downhill (slope < -12°)	+10 min/300 m of descent	$T = a * \Delta S - 2 * \Delta \mathrm{H}$	
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 metre.
ΔH is the vertical elevation difference between cells in metre.
a is the horizontal cost factor according to different land cover types.

3.2.3 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 et al., 2020). The median values of both response and explanatory variables within the cells were determined for each of the 11,786 grid cells. Figure 3.1 shows the spatial pattern of the aggregated data for the scenicness response and the standardised covariates.

3.2.4 Data analysis

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

$$y_i = \beta_0 + \sum_{j=1}^{m} \beta_j x_{ij} + \varepsilon_i \tag{3.1}$$

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 (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:

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^{m} \beta_j(u_i, v_i) x_{ij} + \varepsilon_i \tag{3.2}$$

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

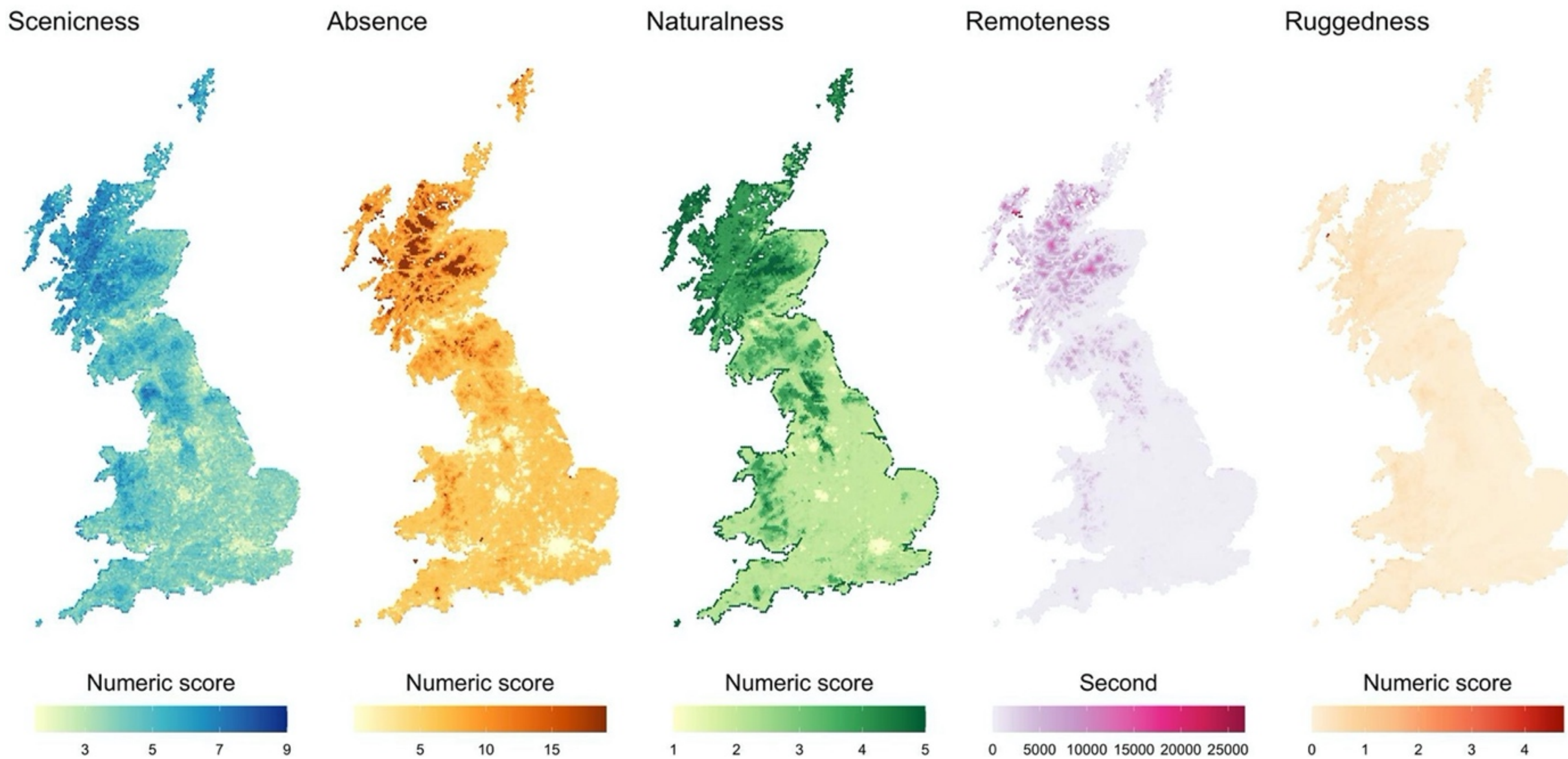


Figure 3.1 The unstandardised Scenic-Or-Not ratings (scenicness) and the four wildness components (i.e. absence, naturalness, remoteness and ruggedness) for Great Britain aggregated over a hexagonal grid with a cell width of 5 km.

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 (Yang, 2014; A Stewart Fotheringham et al., 2017). 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 et al., 1999; Mei et al., 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 (Yang, 2014; A Stewart Fotheringham et al., 2017). In a MGWR model, an individual bandwidth is determined for each predictor variable. This allows the scale of relationship non-stationarity to vary for each target-to-predictor variable relationship, as described in Equation ((3.3):

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^{m} \beta_{bwj}(u_i, v_i)x_{ij} + \varepsilon_i \tag{3.3}$$

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 (T. Oshan et al., 2019).

3.3 Results

3.3.1 Exploratory analysis

The pairwise Pearson correlation analysis is shown in Figure 3.2. 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 and Tiefelsdorf, 2005). However, more recent research has suggested that collinearity is unproblematic where the correlation is < 0.8 or > -0.8 (Comber and Harris, 2018). The robustness of GWR to the effects of multicollinearity has been also demonstrated, particularly with a large sample size (Páez et al., 2011b; A Stewart Fotheringham and Oshan, 2016).

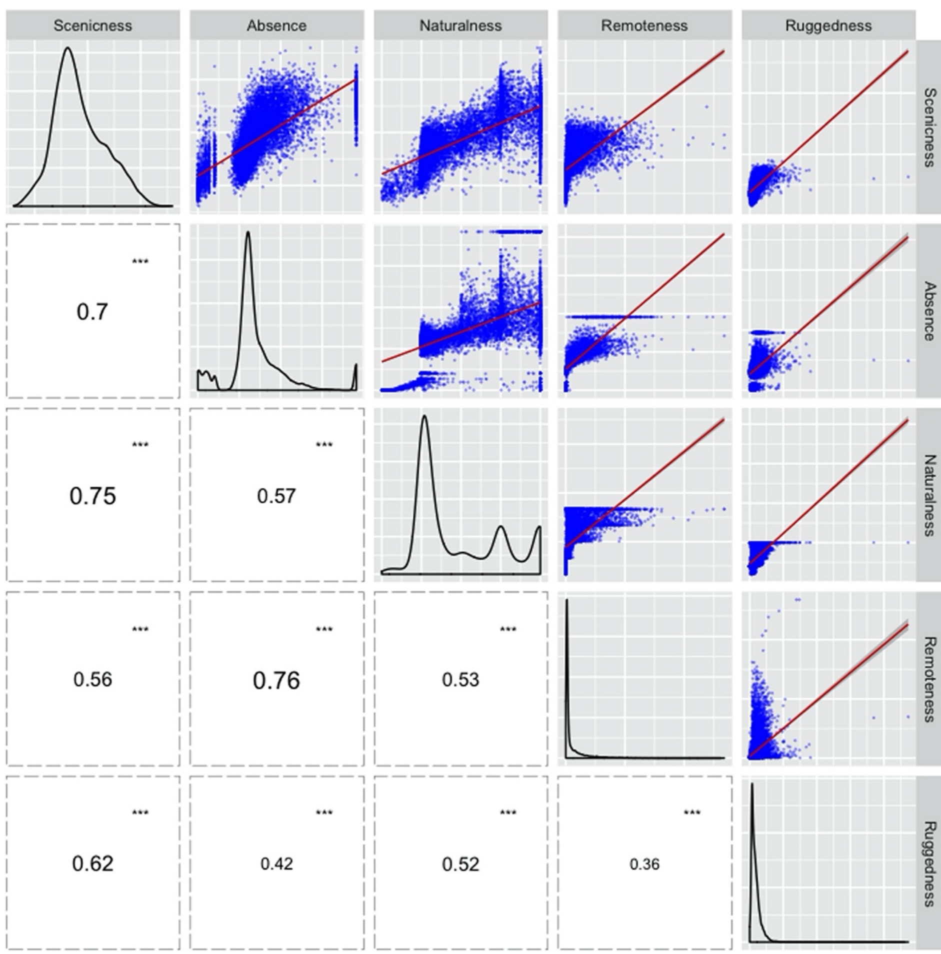


Figure 3.2 Pearson pairwise correlation, scatterplots and distributions of the input data (significance indicated by *** < 0.001, ** < 0.01, * < 0.05).

3.3.2 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 3.3). 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 et al., 1980). A marginally improved model fit with all covariates was found, as indicated by the adjusted R-squared and corrected Akaike information criterion (AICc) values (see Table 3.3). 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 3.3 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 ($\text{Moran's I} = 267,\ p < 0.001$; Jarque-Bera statistic $= 15074,\ p < 0.001$). The map of residuals (Figure 3.3) 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 (Figure 3.3) 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.

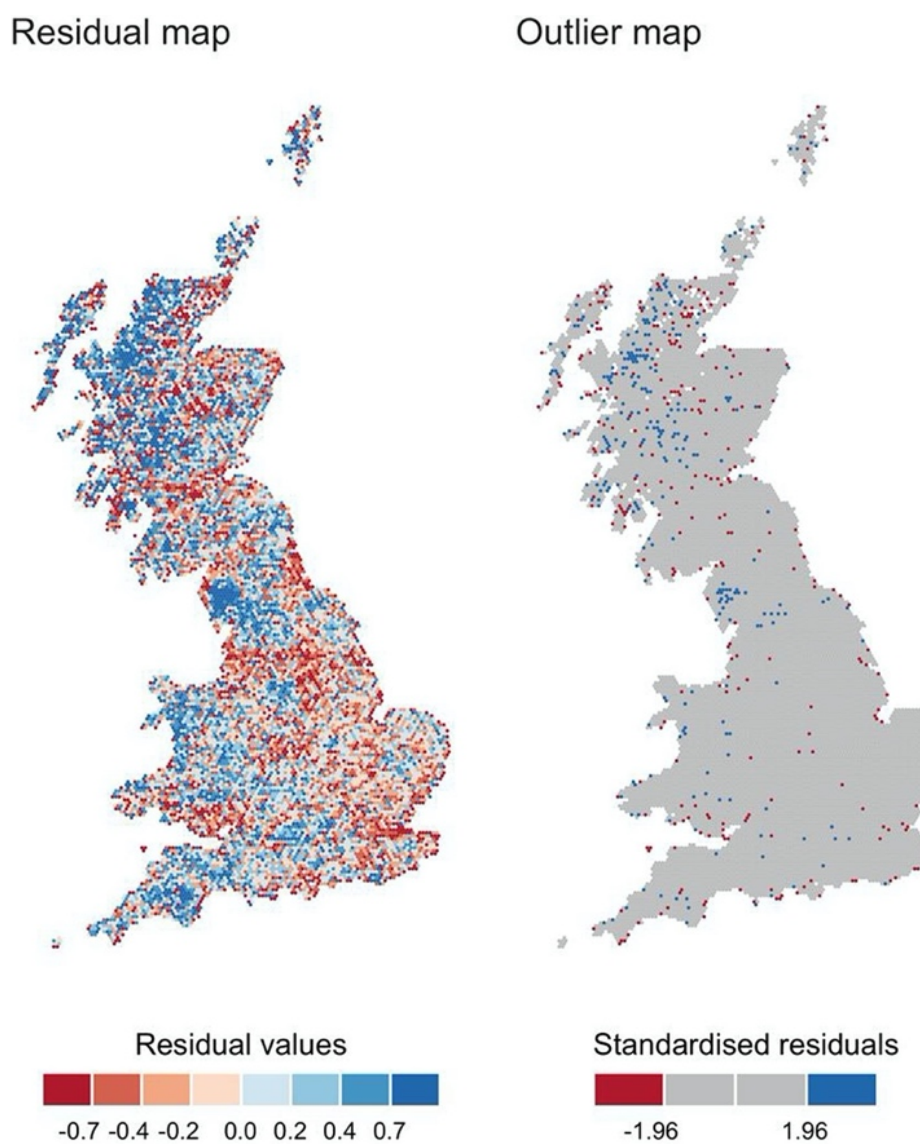


Figure 3.3 The quantile-classified residual map (left) and the outlier map (right) highlights areas where the global model overestimated (red) and under-estimated (blue) landscape scenic beauty.

3.3.3 Standard GWR and multiscale GWR

As collinearity may be present in local subsets under the GW framework (Wheeler and Tiefelsdorf, 2005) despite a global absence, the GWR and MGWR analyses were coupled with the local collinearity diagnostic tests using the mgwr Python package (T. Oshan et al., 2019). Figure 3.4 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 are indicative of significant collinearity amongst the predictor variables (Belsley et al., 1980; Gollini et al., 2015). This collinearity may be caused by the single GWR bandwidth, which can increase collinearity between variables (Oshan and Fotheringham, 2018). All of the local MGWR models were found to have CNs of less than 3.

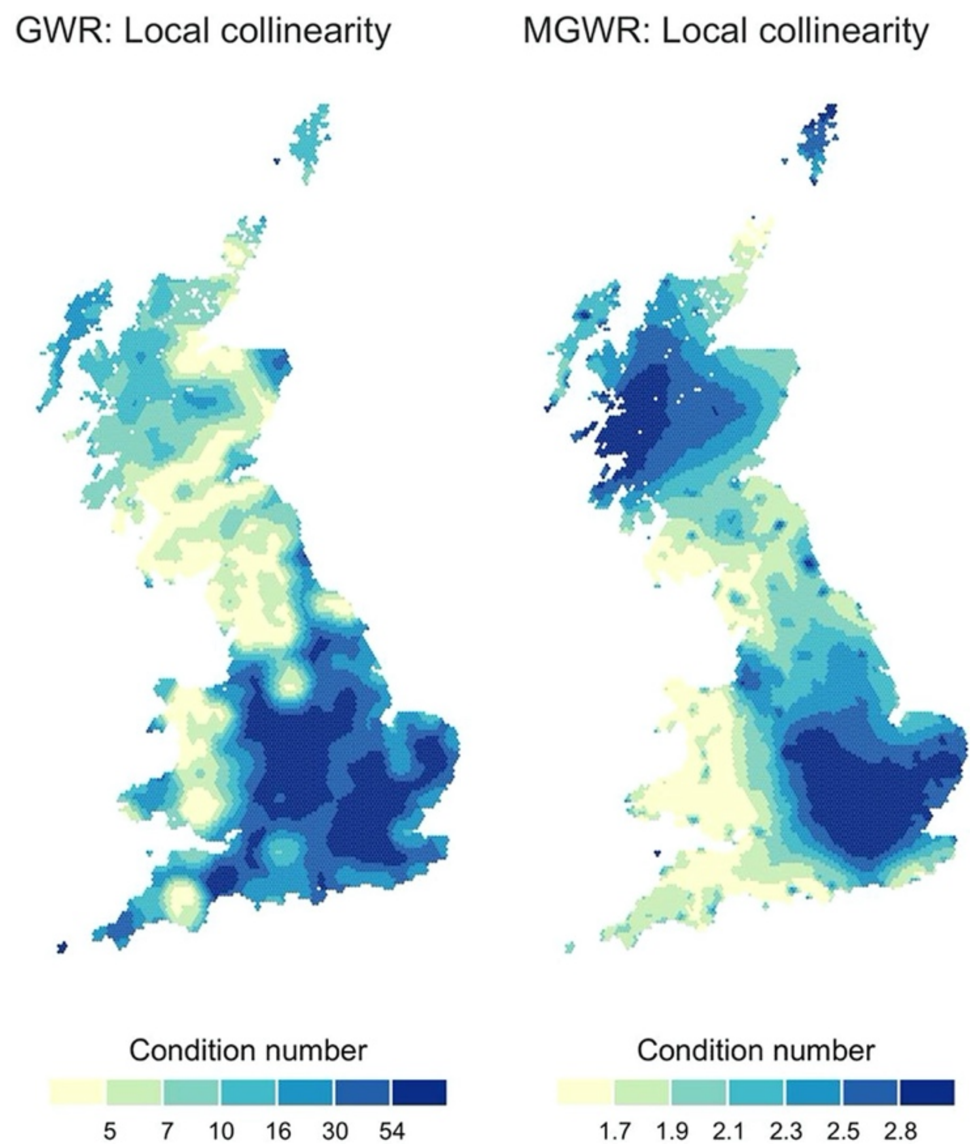


Figure 3.4 The diagnostic tests of the local collinearity for the GWR (left) and the MGWR (right) models using quantile breaks.

Bandwidth selections for both the GWR and MGWR models were optimised using a cross-validation approach under a Gaussian weighting kernel. Table 3.4 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 R2 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.

Figure 3.5 and 3.6 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 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 (Figure 3.3). 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 number of locations have significant coefficient estimates obtained from the MGWR calibration.

The MGWR results shown in Table 3.4 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 Figure 3.6 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.

Table 3.3 The coefficient estimates and associated p-values of the MLRs with and without remoteness.

	MLR without Remoteness					MLR with Remoteness				
Variable	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 R2 = 0.709, AICc = 23,027					Adjusted R2 = 0.710, AICc = 23,001				

Table 3.4 The coefficient estimates arising from the GWR and MGWR models (1Q = 1st quartile, Med = median, 3Q = 3rd quartile, IQR = interquartile range).

	GWR	Bandwidth: 15.2 km			MGWR				
Parameter	1Q	Med	3Q	IQR	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

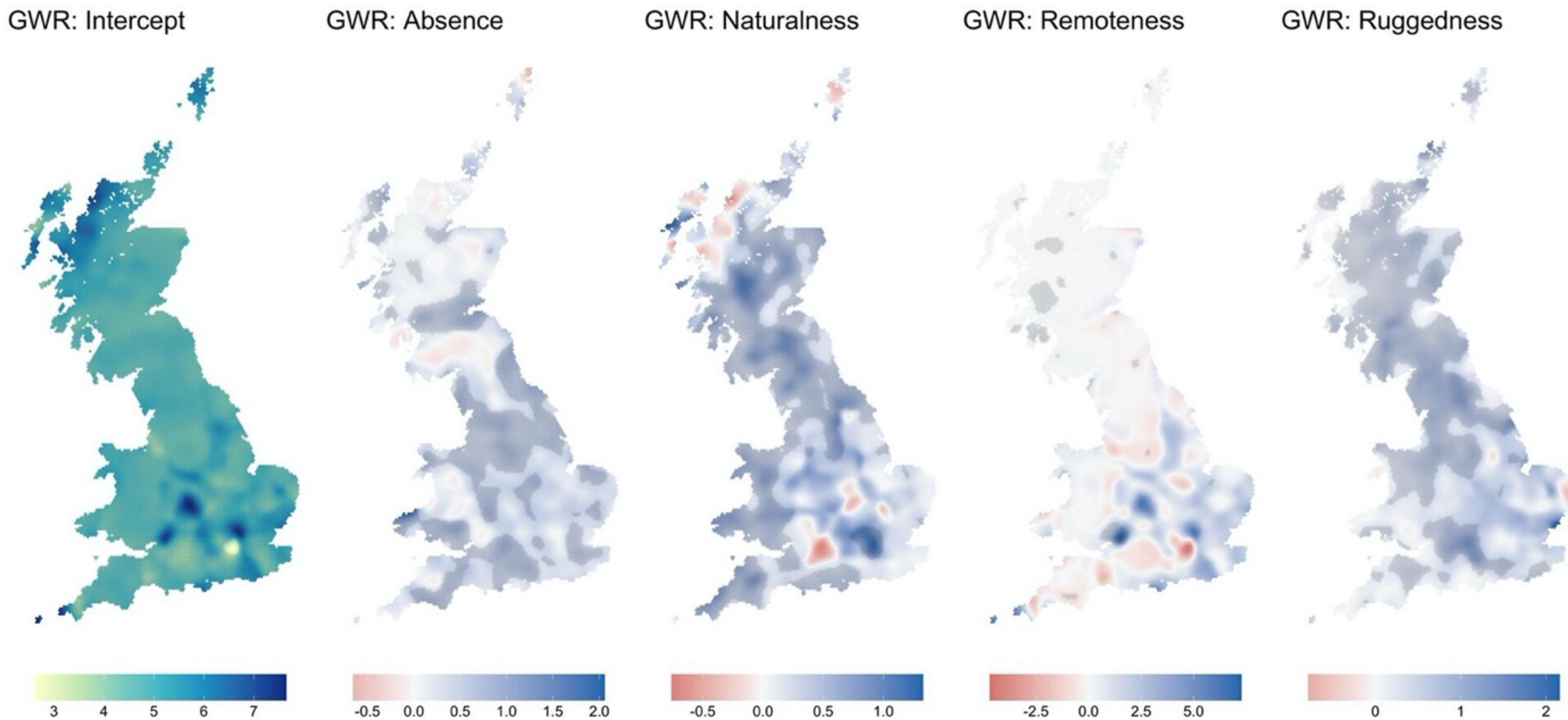


Figure 3.5 The GWR coefficient estimates for the intercept and each wildness covariate with the significance of coefficient estimates denoted by black shaded outlines.

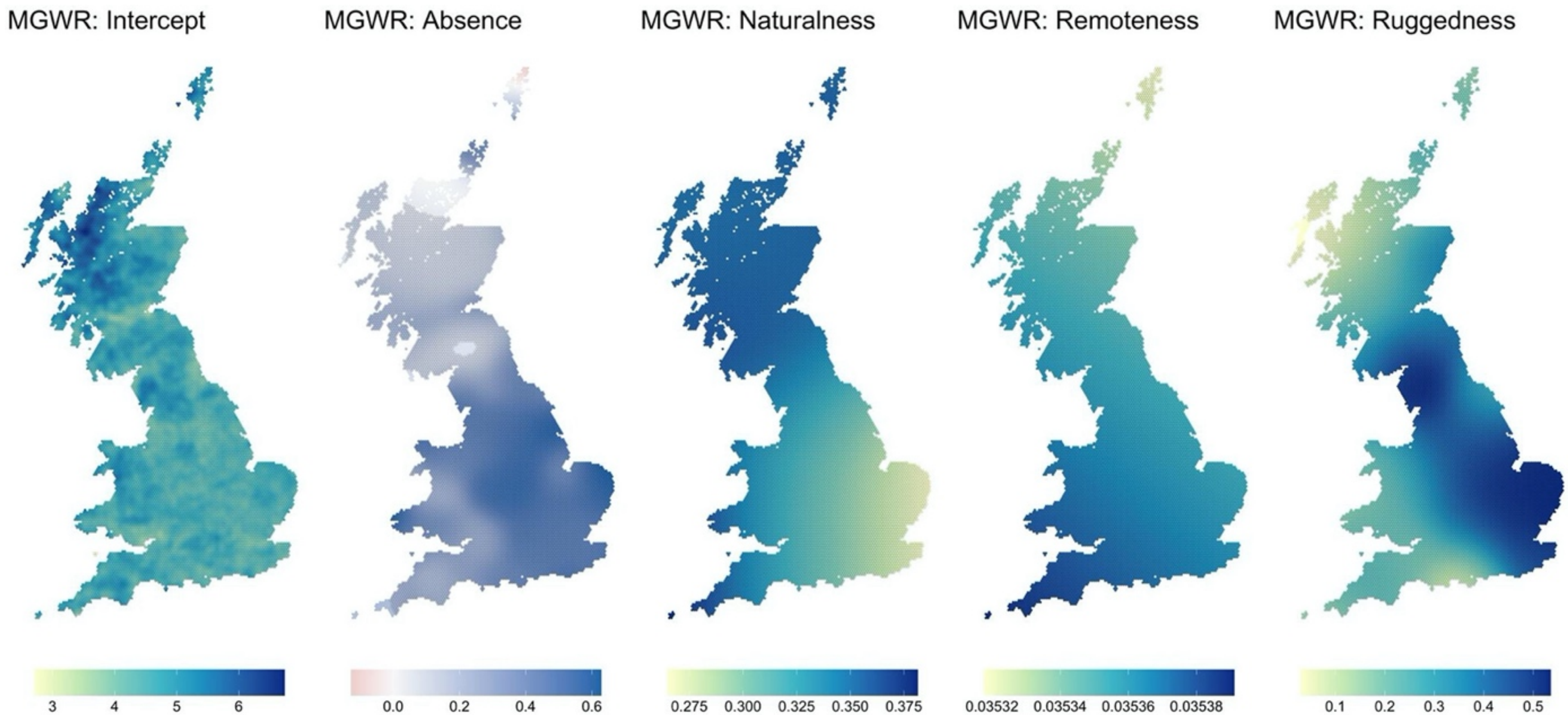


Figure 3.6 The MGWR coefficient estimates for the intercept and each wildness covariate with the significance of coefficient estimates denoted by black shaded outlines.

3.4 Discussion

3.4.1 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 (Lu et al., 2017; A Stewart Fotheringham et al., 2017; Wolf et al., 2018; Murakami et al., 2018; Comber et al., 2020) 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 et al., 2018) these findings highlight strategies for future landscape enhancement and conservation throughout the United Kingdom.

3.4.2 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) (Stan Openshaw, 1984; Stan. Openshaw, 1984). 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 and 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 (James F Palmer and 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 et al., 2019), such efforts sensitively connect both the human and the natural components of landscapes.

3.5 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.

Reference

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), pp.3–25.

Belsley, D.A., Kuh, E. and Welsch, R.E. 1980. *Regression Diagnostics*. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Bowman, A.W. 1984. An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*. **71**(2), pp.353–360.

Brunsdon, C., Fotheringham, A.S. and Charlton, M. 1999. Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Science*. **39**(3), pp.497–524.

Brunsdon, C., Fotheringham, A.S. and Charlton, M.E. 1996. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*. **28**(4), pp.281–298.

Butler, A. and Berglund, U. 2014. Landscape Character Assessment as an Approach to Understanding Public Interests within the European Landscape Convention. *Landscape Research*. **39**(3), pp.219–236.

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

Carver, S., Comber, A., McMorran, R. and Nutter, S. 2012. A GIS model for mapping spatial patterns and distribution of wild land in Scotland. *Landscape and Urban Planning*. **104**(3–4), pp.395–409.

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

Carver, S., Tricker, J. and Landres, P. 2013. Keeping it wild: Mapping wilderness character in the United States. *Journal of Environmental Management*. **131**, pp.239–255.

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

Cleveland, W.S. 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*. **74**(368), pp.829–836.

Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., Lü, Y., Murakami, D., Nakaya, T., Wang, Y. and Harris, P. 2020. The GWR route map: a guide to the informed application of Geographically Weighted Regression. , pp.1–34.

Comber, A., Carver, S., Fritz, S., McMorran, R., Washtell, J. and 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), pp.142–152.

Comber, A. and Harris, P. 2018. Geographically weighted elastic net logistic regression. *Journal of Geographical Systems*. **20**(4), pp.317–341.

Comber, A., Mooney, P., Purves, R.S., Rocchini, D. and Walz, A. 2016. Crowdsourcing: It Matters Who the Crowd Are. The Impacts of between Group Variations in Recording Land Cover T. C. Matisziw, ed. *PLOS ONE*. **11**(7), p.e0158329.

Conrad, E., Christie, M. and Fazey, I. 2011. Is research keeping up with changes in landscape policy? A review of the literature. *Journal of Environmental Management*. **92**(9), pp.2097–2108.

Daniel, T.C. 2001. Whither scenic beauty? Visual landscape quality assessment in the 21st century. *Landscape and Urban Planning*. **54**(1–4), pp.267–281.

Dramstad, W.E., Tveit, M.S., Fjellstad, W.J. and Fry, G.L.A. 2006. Relationships between visual landscape preferences and map-based indicators of landscape structure. *Landscape and Urban Planning*. **78**(4), pp.465–474.

Dunkel, A. 2015. Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning*. **142**, pp.173–186.

Fotheringham, A.S. and Oshan, T.M. 2016. Geographically weighted regression and multicollinearity: dispelling the myth. *Journal of Geographical Systems*. **18**(4), pp.303–329.

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

Fotheringham, A.S., Yang, W. and Kang, W. 2017. Multiscale geographically weighted regression (MGWR). *Annals of the American Association of Geographers*. **107**(6), pp.1247–1265.

Frank, S., Fürst, C., Koschke, L., Witt, A. and Makeschin, F. 2013. Assessment of landscape aesthetics—Validation of a landscape metrics-based assessment by visual estimation of the scenic beauty. *Ecological Indicators*. **32**, pp.222–231.

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

Gliozzo, G., Pettorelli, N. and Haklay, M. (Muki) 2016. Using crowdsourced imagery to detect cultural ecosystem services: a case study in South Wales, UK. *Ecology and Society*. **21**(3), p.art6.

Gollini, I., Lu, B., Charlton, M., Brunsdon, C. and Harris, P. 2015. GWmodel: An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. *Journal of Statistical Software; Vol 1, Issue 17 (2015)*.

Hong, J.Y. and Jeon, J.Y. 2017. Exploring spatial relationships among soundscape variables in urban areas: A spatial statistical modelling approach. *Landscape and Urban Planning*. **157**, pp.352–364.

Kerebel, A., Gélinas, N., Déry, S., Voigt, B. and Munson, A. 2019. Landscape aesthetic modelling using Bayesian networks: Conceptual framework and participatory indicator weighting. *Landscape and Urban Planning*. **185**, pp.258–271.

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), pp.177–198.

Lu, B., Brunsdon, C., Charlton, M. and Harris, P. 2017. Geographically weighted regression with parameter-specific distance metrics. *International Journal of Geographical Information Science*. **31**(5), pp.982–998.

Luo, J. and 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), pp.51–64.

Mei, C.-L., Xu, M. and Wang, N. 2016. A bootstrap test for constant coefficients in geographically weighted regression models. *International Journal of Geographical Information Science*. **30**(8), pp.1622–1643.

Morton, R.D., Rowland, C.S., Wood, C.M., Meek, L., Marston, C.G. and Smith, G.M. 2014. Land Cover Map 2007 (25m raster, GB) v1.2.

Murakami, D., Lu, B., Harris, P., Brunsdon, C., Charlton, M., Nakaya, T. and Griffith, D.A. 2018. The importance of scale in spatially varying coefficient modeling. *Annals of the American Association of Geographers*., pp.1–21.

Naismith, W.W. 1892. Excursions: Cruach Ardran, Stobinian, and Ben More. *Scottish Mountaineering club journal*. **2**(3), p.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), pp.673–690.

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

Openshaw, S. 1984. Ecological fallacies and the analysis of areal census data. *Environment & Planning A*. **16**(1), pp.17–31.

Oshan, T., Li, Z., Kang, W., Wolf, L. and 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), p.269.

Oshan, T.M. and Fotheringham, A.S. 2018. A Comparison of Spatially Varying Regression Coefficient Estimates Using Geographically Weighted and Spatial-Filter-Based Techniques. *Geographical Analysis*. **50**(1), pp.53–75.

Oteros-Rozas, E., Martín-López, B., Fagerholm, N., Bieling, C. and 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**, pp.74–86.

Páez, A., Farber, S. and 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), pp.2992–3010.

Palmer, J.F. and Hoffman, R.E. 2001. Rating reliability and representation validity in scenic landscape assessments. *Landscape and Urban Planning*. **54**(1–4), pp.149–161.

Schirpke, U., Tasser, E. and Tappeiner, U. 2013. Predicting scenic beauty of mountain regions. *Landscape and Urban Planning*. **111**(1), pp.1–12.

Seresinhe, C.I., Preis, T., MacKerron, G. and Moat, H.S. 2019. Happiness is Greater in More Scenic Locations. *Scientific Reports*. **9**(1), p.4498.

Seresinhe, C.I., Preis, T. and Moat, H.S. 2015. Quantifying the Impact of Scenic Environments on Health. *Scientific Reports*. **5**(1), p.16899.

Seresinhe, C.I., Preis, T. and Moat, H.S. 2017. Using deep learning to quantify the beauty of outdoor places. *Royal Society Open Science*. **4**(7), p.170170.

Su, Y.-F., Foody, G.M. and 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), pp.172–180.

Sun, R., Xie, W. and Chen, L. 2018. A landscape connectivity model to quantify contributions of heat sources and sinks in urban regions. *Landscape and Urban Planning*. **178**, pp.43–50.

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. and 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**, pp.128–137.

Wang, L., Ai, T., Shen, Y. and Li, J. 2020. The isotropic organization of DEM structure and extraction of valley lines using hexagonal grid. *Transactions in GIS*. **24**(2), pp.483–507.

Wheeler, D. and Tiefelsdorf, M. 2005. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems*. **7**(2), pp.161–187.

Wolf, L.J., Oshan, T.M. and Fotheringham, A.S. 2018. Single and multiscale models of process spatial heterogeneity. *Geographical Analysis*. **50**(3), pp.223–246.

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

van Zanten, B.T., Van Berkel, D.B., Meentemeyer, R.K., Smith, J.W., Tieskens, K.F. and Verburg, P.H. 2016. Continental-scale quantification of landscape values using social media data. *Proceedings of the National Academy of Sciences*. **113**(46), pp.12974–12979.

Zoderer, B.M., Tasser, E., Carver, S. and Tappeiner, U. 2019. An integrated method for the mapping of landscape preferences at the regional scale. *Ecological Indicators*. **106**, p.105430.

Zube, E.H. and Pitt, D.G. 1981. Cross-cultural perceptions of scenic and heritage landscapes. *Landscape Planning*. **8**(1), pp.69–87.

Chapter 4 Comparing scenic evaluations between experts and non-experts

Abstract

The debate over the conceptual constructs of landscape aesthetics, specifically whether landscape quality is inherently related to landscape physical characteristics or is subjectively "in the eye of the beholder," has continued for years. Solutions accommodating both the biophysical and perceptual aspects of landscapes are thus desirable for landscape planners and policymakers. In response to policy shifts that emphasise both expert and public landscape perspectives, this study investigates the relationships between formal and informal landscape evaluations. It analyses crowdsourced data describing landscape aesthetic quality (Scenic-Or-Not) and authoritative landscape quality assessments (the Landscape Assessment Decision Making Process (LANDMAP) of Wales). Some agreement was found regarding landforms most likely to be perceived as scenic or unattractive by experts and non-experts, which aligns with previous landscape perception studies. However, contested landscape typologies are identified formal and informal landscape aesthetic evaluations are compared. Several limitations and implications for current formal landscape assessment paradigms (GIS based and vertical) are discussed and several approaches for capturing on-the-ground perceptions are suggested including recent extensions to GIS derived viewsheds (e.g., vertical voxel viewsheds).

Keywords: landscape character assessments; crowdsourcing; Scenic-Or-Not

4.1 Introduction

For many years there has been a vigorous debate on whether landscape quality relates to inherent physical landscape characteristics or is in the eye of the beholder (Lothian, 1999; Daniel, 2001). Given these philosophical contrasts, a successful marriage of these objective and subjective paradigms has not yet been achieved, making landscape evaluation challenging. There have been numerous efforts to devise techniques for landscape assessment (Dunkel, 2015; Antrop and Van Eetvelde, 2017; Bubalo et al., 2019), and solutions that balance the perceptual and biophysical aspects of landscapes are still being sought (Daniel, 2001). Perceptual approaches to landscape aesthetic assessment are unavoidably problematic due to differentiated preferences driven by varying cultural backgrounds (Zube and Pitt, 1981) and social stratum (Dramstad et al., 2006), as well as the cost of eliciting collective perspectives. Thus, bio-physical approaches that rely heavily on expert knowledge remain favoured choices in landscape management and planning practices, such as nature preservation and designation (Daniel, 2001; Simensen et al., 2018; Gosal and Ziv, 2020). Inevitably, such expert-based approaches have also faced criticism for their lack of objectivity, transparency, and replicability (even among experts) (Terkenli, 2001; Daniel, 2001).

The development and ratification of the European Landscape Convention's (ELC) landscape policy set a broader definition of landscape as, "an area... perceived by people," placing the public as central to any understanding of landscape. This document aims at pressing on the establishment of procedures for public participation in landscape management, protection, and planning practices (Council of Europe, 2000). The identification and classification of landscapes were underlined, and the definition of landscape quality objectives derived from public consultations were required (Santé et al., 2020). The focus of landscape assessments in the signatory countries has since shifted from expert to local/nonexpert knowledge in response to this guidance (Jones, 2007). Despite emphasis on the public involvement, the formulation of landscape quality objectives has not been standardised and the concept of quality indicators in the ELC continues to be questioned (Santé et al., 2020). As such, it is expedient for practitioners to avoid dealing with diverse perspectives and perceptions from multiple people (Conrad, Christie, et al., 2011).

The implementation of the ELC landscape policies, the UK Landscape Character Assessment (LCA) framework (Swanwick, 2002), has at its centre a hierarchical classification system of landscape character. The LCA initially uses a typology of landscape character based on the visual distinctiveness and continuity of combinations of geology, landform, soils, vegetation, land use, field patterns, and human settlement that is applied through spatial hierarchical mapping. Next, judgements are imparted about landscape character, leading to decisions concerning the management, planning, and protection of the various landscape types and areas. The process of characterisation in the first phase is considered objective while assigning quality or value in the second phase is subjective.

Notwithstanding, both processes in the framing of landscape characters and values are dominated by expert perspectives as the mainstay of formal landscape evaluation (Butler, 2016). The expert appraisal of landscape aesthetic values as part of the LCA has been largely criticised as a subjective and opaque process, where the landscape perceptions held by the public have hitherto been insufficiently included (Conrad, F. Cassar, et al., 2011).

A few landscape characteristics that are well-recognised by experts to contribute to scenic quality are used in the LCA classification scheme. For instance, physical factors such as topography, water, land cover, and human artefacts are seen as determinants of scenic quality, and the landscape is classified into relatively homogenous units of land (i.e., character areas) based on visual interpretations of these data. The characterisation of landscape description units (LDUs) based on static maps with the aid of GIS is claimed to be an objective process, utilising landscape perceptions from a birds-eye perspective rather than that of a human (Groom G. et al., 2006). However, this overlooks the subjectivity involved in choosing predetermined criteria for LDUs (Crofts, 1975). Moreover, the subsequent on-site assessment heavily relies on the subjective judgments of experts at a few representative locations from which the valuation is formulated and generalised to the entire character area. Thus landscape- intrinsic qualities are assessed at limited, predefined locations and are subject to interpreter preferences, shaped by their professional backgrounds and experiences (Dramstad et al., 2006). Such qualitative evaluations may overlook any interactive effects of landscape components, for example, visual screening effects caused by vegetative patterns or terrain (Arriaza, J. F. Cañas-Ortega, et al., 2004), and scenic quality may be inappropriately extrapolated over a large area. There have been very few studies that have examined how professional landscape evaluations relate to those held by citizens.

Crowdsourced spatial information has been increasingly used to understand human/landscape interactions from local to continental scales (Dunkel, 2015; Tenerelli et al., 2016; Boris T van Zanten et al., 2016). Data generated by citizen sensors (Michael F Goodchild, 2007) have the potential to be more closely related to on-the-ground perceptions and less bounded by professional biases (Bubalo et al., 2019). They also have the potential to contribute to formal landscape assessments. In the UK, a web-based crowdsourcing project, Scenic-Or-Not [http://scenicornot.datasciencelab.co.uk/ (accessed on 12 February 2, 2021)] collects people's landscape preferences at a granularity of 1 km. Participants are invited to rate the scenicness of randomly presented photographs from each 1 km^2 grid on a 1 to 10 scale. Empirical studies using these data have investigated the impact of scenic beauty on human health (Seresinhe et al., 2015) and happiness (Seresinhe et al., 2019), explored the composition of landscape scenic beauty as perceived by the public (Seresinhe et al., 2017), verified landscape scenic estimation based on Flickr and OpenStreetMap data (Seresinhe et al., 2018) and constructed a language model to predict landscape scenic beauty (Chesnokova, Nowak and Ross S Purves, 2017). In landscape evaluation, using photographs as proxies to elicit public landscape preferences is generally considered an acceptable approach for landscape aesthetic assessments, since each ground-level image approximates what people perceive at each specific location, despite a continuing debate on its reliability and validity (Daniel, 2001; Unwin, 2006; Gyllin and Grahn, 2015). Scenic-Or-Not's wide and granular coverage provides a valuable resource that could be used within LCA practice, and to respond to the ELC's policy to involve public perception.

It is not the intention of this study to aspire to any ultimate resolution for the landscape classification system, but instead to seek the benefits of synergy to supplement the formal practice of LCA. The central aim is to leverage the granular coverage of crowdsourced spatial data concerned with scenic beauty and evaluate the validity of the LDUs of the formal paradigm by investigating the variability of public opinions underneath these distinct character areas "objectively" delineated by experts. Based on these units, this study attempts not only to understand which perspectives on landform typology are viewed differently between landscape architects and the public but also to identify which perspectives are unlikely to be disputed.

4.2 Data and methods

4.2.1 Study area

Wales is a relatively small country (approximately 21,000 km^2) in the United Kingdom, bordered by England to its east and surrounded by the Irish Sea and the Bristol Channel to its north, west and south. It is famed for the mountainous and coastal landscapes with three National Parks—Snowdonia, the Brecon Beacons and the Pembrokeshire Coast—bringing numerous economic value and benefits of tourism (see Figure 4.1). The main population areas are situated in South Wales, including cities such as Cardiff, Swansea, and Newport, with another significant population area in North East Wales.

Figure 4.1 The study area.

4.2.2 Data

4.2.2.1 LANDMAP visual and sensory aspect

In the early 1990s, the methodological deficiencies of conventional landscape character mapping in terms of the vagueness of the qualitatively evaluative criteria and the robustness of quality assurance were evident, hindering the justification of policy-making, e.g., designation of "quality" landscapes. Consequently, Wales commenced developing its national landscape assessment methodology (known as the Landscape Assessment Decision Making Process, LANDMAP), providing greater consistency and defensible information on landscape for policy- and decision-making.

The LANDMAP data serves as the Welsh landscape baseline, comprising five landscape aspects—geological landscape, landscape habitats, visual and sensory, historic landscape, and cultural landscape—each of which considers a different theme but collectively covers the scope of a landscape's character. Each aspect has its own unique map layer and survey records carried out by the local authority and landscape architects which have been, in turn, joined as five nationwide themed datasets.

The visual and sensory aspects were used in this study where the landscape was classified into the distinct character areas as perceived through human senses. These areas were drawn upon the spatial overlays using the physical attributes of landform and land cover information accessed through the Geo-Portal for Wales (http://lle.gov.wales/catalogue/item/LANDMAPVisualSensory/ (accessed on 12 February 2021)). This dataset consists of 1,991 distinct character areas which were characterised in the form of discrete polygons and allocated on a hierarchy of four levels. The level 1 typology was defined in accordance with broad landform and land cover and was consecutively collated into levels 2 and 3 as landform and land cover, respectively (Table 4.1).

Table 4.1 The level 1–3 classification categories used for characterising the visual and sensory aspect areas in the Landscape Assessment Decision Making Process (LANDMAP) approach, adapted from (Weledol and Landmap, 2016). The grey background is used to easily distinguish the descendants of each level-2 typology.

Level 1 **Broad landform and land cover**	**Level 2** **Landform**	**Level 3** **Land cover**
Upland	Exposed upland or plateau	Barren or rocky upland Upland moorland Upland grazing Wooded upland and plateau Mosaic upland and plateau
	Upland valleys	Open upland valleys Open or wooded mosaic upland valleys Wooded upland valleys
	Hills, lower plateau, and scarp slopes	Hillside and scarp slopes moorland Hillside and scarp slopes grazing Wooded hillside and scarp slopes Hillside and scarp slopes mosaic Open hillside and scarp slopes Hill and lower plateau moorland Hill and lower plateau grazing Wooded hill and lower plateau Hill and lower plateau mosaic Open hill and lower plateau
Lowland	Lowland valleys	Open lowland valleys Mosaic lowland valleys Wooded lowland valleys
	Rolling lowland	Open rolling lowland Mosaic rolling lowland Wooded rolling lowland
	Flat lowland or levels	Flat open lowland farmland

		Flat wooded lowland Flat lowland mosaic Lowland wetland
	Coastal	Intertidal Dunes and dune slack Cliffs and cliff tops Other coastal wildland Small island
Development	Built land	Village Dispersed settlement Urban
	Developed unbuilt land	Amenity land Informal open space Excavation Derelict or waste ground Road corridor
Water	Coastal waters	Sea Estuary
	Inland water (including the associated edge)	River Lake Ria

4.2.2.2 Scenic-Or-Not

This website allows people to evaluate places in Britain by rating photos collected from the Geograph project (https://www.geograph.org.uk/ (accessed on 12 February 2021), which is the early volunteered geographic information (VGI) campaign in the UK. This project invites people to contribute geographically representative photographs and information for every square kilometre. The Scenic-Or-Not dataset contains 212,212 photos at distinct locations, each of which has been rated by at least three people on an integer scale of 1 to 10, with 10 being the most scenic and each square kilometre contains one photo. Following the homogenous assumption of scenic quality implicitly made by the landscape characterisation approach, all scenic ratings within each aspect area were aggregated to calculate the mean and entropy of scenic scores. A subset of the Scenic-Or-Not dataset covering the study area was employed to represent the public opinions on landscape aesthetics, containing a total number of 138,312 scenic scores associated 19,063 images. Hereafter, "scenicness" is used to simply refer to these scenic ratings.

Table 4.2 The importance definitions of LANDMAP evaluation (Natural Resources Wales, 2017).

Evaluation score	Definition of importance
Outstanding	International or national
High	Regional and county
Moderate	Local
Low	Little or no importance

4.2.3 Methods

4.2.3.1 Shannon entropy

Given the Scenic-Or-Not images with different numbers of ratings but at least 3 votes on a discrete 1–10 scale, these integer scores allow the calculation of not only the central tendency of scenicness but also the variability or dispersion of collective opinions. The Shannon entropy metric was borrowed from the information field and used to quantify the dispersion of opinions regarding each specific landform (Shannon, 1948). This metric was originally developed to determine the average minimum number of bits required to fully encode a message in relation to the statistical distribution of possible messages. It has been widely employed in many disciplines (e.g., statistics, physics, and communication). The mathematical equation is as follows:

$$H(X) = -\sum_{i} p_i \log_2 p_i \tag{4.1}$$

where X represents the collective scenic scores for each image, and p_i is the probability of occurrence of the i^{th} discrete outcome ($i = 1, \cdots, 10$ herein).

4.2.3.2 Generalised linear model

The associations between the landscape typologies and scenic quality assigned by experts and non-experts were analysed using general linear models (GLMs) where the dummy coding was used with the same reference category for comparison. For the expert-based evaluation, each of the ordinal levels was dichotomised as a binary response variable y_{ik} to fit the binary logistic regression model as follows:

$$\log\left(\frac{P(y_{ik} = 1|x_{ij})}{1 - P(y_{ik} = 1|x_{ij})}\right) = \beta_0 + \sum_j \beta_j x_{ij} + \epsilon_i \tag{4.2}$$

where i indexes the i^{th} observation, and k denotes the four memberships of scenic quality; x_{ij} represents the j^{th} dummy-coded categorical variables ($j = 1, \cdots, 10$ herein). The logit link function is employed as the left side of Equation (4.2) to model the log of the odds, resting on the assumption of a logistic distribution for the error term. However, it is not suited for modelling the continuous response such as average of scenicness herein. Thus, the normal GLM, employing the identity link function and assuming a normal distribution for y_i, known as an ordinary linear regression model, was used with the form shown in Equation (4.3):

$$y_i = \beta_0 + \sum_j \beta_j x_{ij} + \epsilon_i \tag{4.3}$$

As mentioned, the difference in the equation is on the left side and the coefficients β_j are commonly estimated by the ordinary least squares (OLS) rather than the maximum likelihood (ML) method, viewed as a special case of GLM thereof (Myers and Montgomery, 1997). All statistics and analyses were implemented via a set of packages in the open-source RStudio Environment (v. 1.3).

4.3 Results

4.3.1 Exploratory analysis

Table 4.3 shows the contingency table of the landscape typologies (11 level-2 classes) and the corresponding four ordinal levels of scenic quality rated by experts. From the marginal distribution, 19% of the 1,986 distinct aspect areas were categorised as 'built land,' probably indicating the highly fragmented but clearly human-made patterns for map-based characterisation, and the majority of these aspect areas were evaluated as the middle of scenic spectrum (moderate scenic quality: 38%, high scenic quality: 36%).

Given the aspect areas with low scenic quality, 56% of the corresponding landforms turns out to be the 'built land' with a probability of 0.19 when considering the moderate scenic quality. At the opposite spectrum, the rate of the 'coastal' typology (2.72%) was higher than other landforms. Despite the same marginal distributions (13%), the respective conditional distributions of the 'exposed upland or plateau' and 'hills, lower plateau, and scarp slopes' rated as high scenic quality were 13% and 19%. A chi-square test of independence showed that there was a significant association between landform and scenic quality, $\chi^2(30, N = 1{,}986) = 886.1, p < .001$.

The point-based Scenic-Or-Not data were aggregated over the aspect areas of the formal landscape assessment to make these data comparable with the areal unit-based data. Following the implicit assumption of homogenous character adopted by the LCA approach regarding landscape perceptions, the equal weighting for each scenic vote, regardless of the corresponding image content, was considered to calculate the underlying central tendency and variability of public scenic perceptions for a given aspect area. There were 1,716 aspect areas intersected by the locational Scenic-Or-Not data.

Table 4.3 Contingency (cross-tabulation) table of the level-2 landscape typologies and the expert-evaluated scenic quality. A total of 1,991 aspect areas were classified with 5 unassessed areas.

LANDMAP Level-2 class	Low (%)		Moderate (%)		High (%)		Outstanding (%)		Total (%)	
Coastal waters	0	(0%)	1	(0.05%)	9	(0.45%)	8	(0.40%)	18	(1%)
Coastal	1	(0.05%)	11	(0.55%)	79	(3.98%)	54	(2.72%)	145	(7%)
Inland water	0	(0%)	16	(0.81%)	23	(1.16%)	19	(0.96%)	58	(3%)
Exposed upland or plateau	20	(1%)	102	(5.14%)	97	(4.88%)	48	(2.42%)	267	(13%)
Upland valleys	5	(0.25%)	70	(3.52%)	68	(3.42%)	21	(1.06%)	164	(8%)
Lowland valleys	5	(0.25%)	92	(4.63%)	123	(6.19%)	19	(0.96%)	239	(12%)
Flat lowland or levels	11	(0.55%)	51	(2.57%)	33	(1.66%)	7	(0.35%)	102	(5%)
Hills, lower plateau, and scarp slopes	7	(0.35%)	99	(4.98%)	136	(6.84%)	16	(0.81%)	258	(13%)
Rolling lowland	6	(0.30%)	122	(6.14%)	75	(3.77%)	11	(0.55%)	214	(11%)
Developed unbuilt land	75	(3.78%)	56	(2.82%)	10	(0.50%)	6	(0.30%)	147	(7%)
Built land	168	(8.46%)	142	(7.15%)	61	(3.07%)	3	(0.15%)	374	(19%)
Sum (%)	298	(15%)	762	(38%)	714	(36%)	212	(11%)	1,986	

Figure 4.2 illustrates the distribution and density characteristics of the mean and Shannon entropy measures of collective scenic ratings for each landform typology. The former nearly follows a normal distribution, ranging from 1–9 with a mean of 4.41 and a median of 4.45, while the latter exhibits a negatively skewed distribution ranging from 0 to 3.25 with a mean of 2.6 and a median of 2.78. The boxplots for each landform type illustrate a difference in the collective scenic perceptions among these typologies in terms of their central tendency and variability. 39% of the total variation in average scenic score was

accounted for by these classification typologies, confirmed by a Welch one-way ANOVA, $Welch's\ F(10, 264.45) = 76.59, p < .001, est.\ \omega^2 = .389$.

4.3.2 Variability of public perceptions on scenic beauty

The entropy metric provides an insight into the underlying variability of the public opinions on landscape scenic beauty within a given landform. Several features are worthy of observation here. At first glance, all distributional patterns exhibited negative skewness in different degrees, indicating the deficiency in using these presumably homogeneous LDUs to delineate landscape perceptions. The typologies associated with upland turned out to exhibit sharply peaked distributions and concentration around a higher entropy of 3 bits. The "exposed upland or plateau" typology exemplified these highly sharp and compact unimodal distributions while the peak splitting was observed in the case of "upland valleys." Likewise, those concerning lowland, except "coastal," demonstrated similar unimodal shapes but with slightly gentle curves such as "rolling lowland" and "lowland valleys." Noteworthy is that the "flat lowland or levels" and "coastal" typologies with fewer effects in topographic relief had a wider range in the entropy measures of scenic scores. Some characteristics possibly overlooked and subsumed within one typology were briefly reported in the discussion.

A few landforms concerning development and water typologies, such as "built land," "developed unbuilt land," and "inland water" had relatively low central values with wider dispersion, compared with those regarding upland and lowland. This suggests that the public opinions on the scenic quality of these landforms were generally less variable than the other typologies, but there were larger variations in the entropy metrics as well. A possible interpretation is that a consensus of scenic quality on these landforms exists; nevertheless, the consent could be context-specific and vary drastically from place to place.

Notably, the "inland water" falls into a bimodal distribution, reflecting two different groups of variability where the major group displays greater variation than the minor one. Although the presence of water has been empirically evidenced as a preferable landscape element in the previous landscape perception studies (Dramstad et al., 2006; Brown and Brabyn, 2012), radically diverging views on scenic beauty within water areas was also evidenced, which could be explored further by breaking this landform down into the level-3 typologies to investigate whether different types of water bodies such as lakes and rivers exert an influence over the observed distribution that is, nonetheless, beyond the scope of this study.

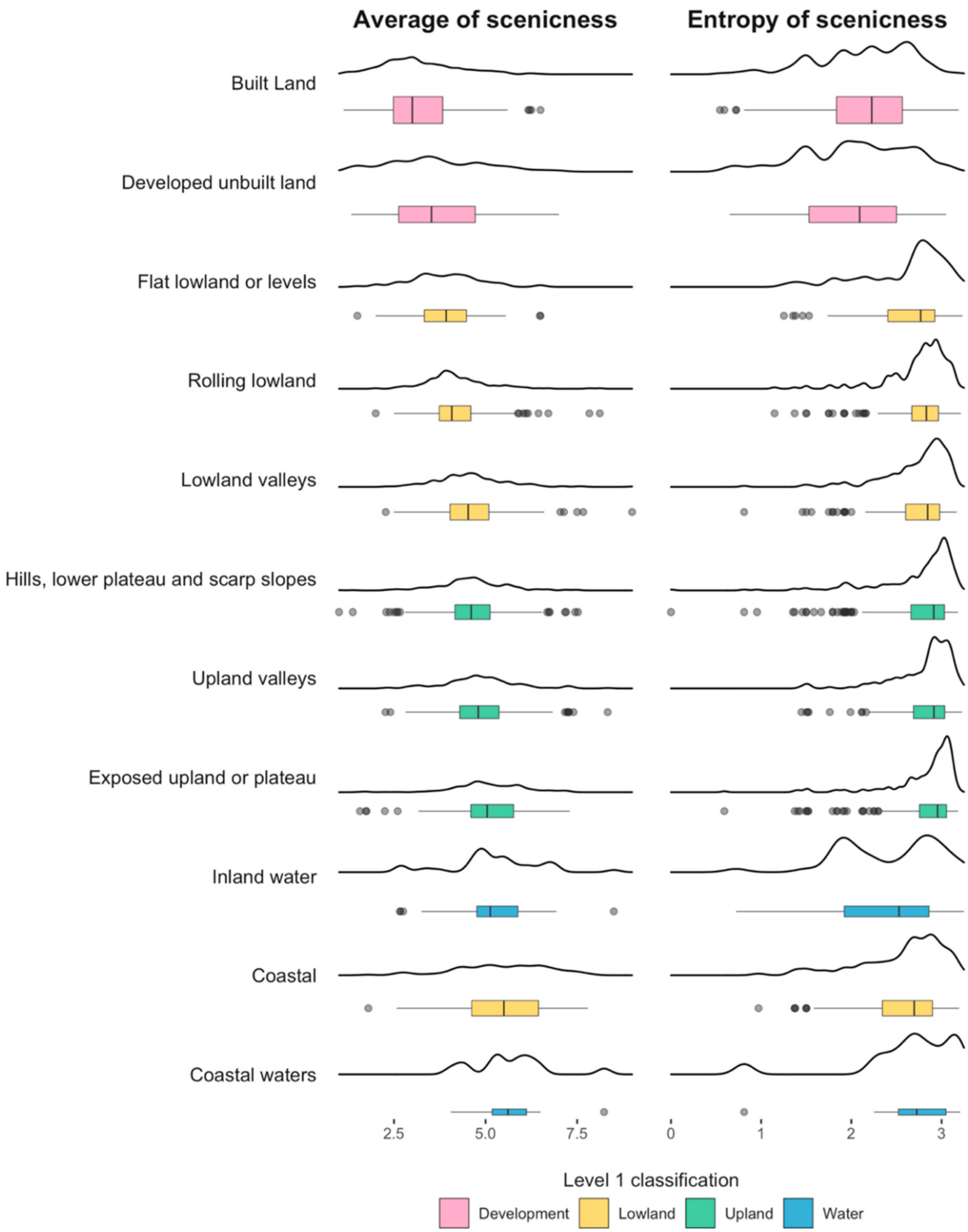


Figure 4.2 Boxplots of the mean (left) and Shannon entropy (right) measures of public scenic ratings for each level-2 landform typology show the underlying central tendency and variability of opinions based on the intersected 1716 aspect areas/observations.

Shading reflects the root of the hierarchical classification scheme (that is, level-1 typology) and the width of the boxplot is proportional to the sample size.

4.3.3 Summary of expert perspectives

The relationships between landscape typologies and scenic quality assigned by the experts and non-experts were further assessed by a set of generalised linear regressions and a simple linear regression. For ranking, the same dummy coding was used with the reference category of 'built land,' which allows for the calculation of the odds ratio and confidence intervals associated with different landscape typologies compared to the identical baseline. Figure 4.3 illustrates the outcomes of each GLM on a log scale to visually compare the degree of uncertainty associated with the point estimate that was denoted by the error bar of 95% confidence interval and standard error. The point estimates at 5% significance levels were denoted by the colour blue. The associated exponentiated coefficients, indicating the odds ratios between the specific landform and the baseline one, are summarised in Table 4.4 Table 4.4for ease of interpretation.

In Figure 4.3, the model for the low scenic quality shows the best performance among the four logistic models, as indicated by the goodness-of-fit measures of pesudo R^2 and Akaike information criterion ($\text{Nagelkerke's } \mathrm{R}^2 = .407; \mathrm{AIC} = 1176.508$) (Akaike, 1973; Nagelkerke, 1991). This appears to be plausible given that the abundant aspect areas of 'developed unbuilt land' and 'built land' were consistently associated with low scenic quality. By contrast, the model for moderate scenic quality exhibited the lowest performance ($\text{Nagelkerke's } \mathrm{R}^2 = .082; \mathrm{AIC} = 2542.847$) where only the odds ratios associated with four of the landforms (i.e., 'coastal water,' 'coastal,' 'flat lowland or levels,' and 'rolling lowland') were statistically significant, as denoted by the colour blue. Concerning high and outstanding scenic quality, all landform types were statistically significant at 5% significance levels. A glimpse of performances over the four models suggests that these landform typologies explain the probabilities for the middle scenic qualities less than those for both ends of the spectrum.

Considering the low scenic quality, the effects of all the landform types except 'developed unbuilt land' were negative and statistically significant, suggesting a decreased likelihood of being rated as low when compared to the 'built land.' In other words, 'built land' is more likely linked to low quality than other typologies based on expert professional views (see Table 4.4). This is prima facie evidence of the subjectivity in the characterisation process of the desk study despite the claimed objectivity. Among the significant typologies, the respective odds of being evaluated as moderate within the landforms of 'rolling lowland' and 'flat lowland or levels' were around 2.2 and 1.6 times greater than within areas of 'built land.' In contrast, the relative odds within the 'built land' were around 10.4 and 7.5 times greater than within the 'coastal water' and 'coastal' areas, respectively. The coastal landforms were not only the most likely to be evaluated as having high scenic quality, but also to have an increased likelihood of being rated with outstanding quality. Additionally, 'hills, lower plateau, and scarp slopes' and 'lowland valleys' were respectively 5.72 and 5.44 times more likely than 'built land' to be evaluated as having high scenic quality. Notably, these two landforms were relatively easier to access compared to those most likely associated with outstanding beauty. The typologies associated with water bodies (i.e., 'coastal waters,' 'coastal,' and 'inland water') were consistently the most likely to be evaluated as having outstanding scenic quality, corresponding to the human preference for water presence found in previous studies (Dramstad et al., 2006; Brown and Brabyn, 2012). The second most likely landforms to be assessed as outstanding were upland landforms. The odds of 'exposed upland or plateau' and 'upland valleys' being rated as outstanding were respectively 27 and 18 times greater than that of 'built land.'

Table 4.4 Results of the logistic regressions for the four levels of scenic quality evaluated by experts to the level-2 classes which were dummy-coded and the 'built land' class was used as the reference category ($^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$).

Categorical variable (LANDMAP Level-2 class)	Odds ratio			
	Low	Moderate	High	Outstanding
Intercept	0.816*	0.612***	0.195***	0.008***
Coastal waters	0.033***	0.096*	5.131**	98.933***
Coastal	0.013***	0.134***	6.142***	73.385***
Inland water	0.010***	0.622	3.372***	60.248***
Exposed upland or plateau	0.102***	1.010	2.928***	27.105***
Upland valleys	0.042***	1.217	3.635***	18.161***
Lowland valleys	0.029***	1.023	5.441***	10.680***
Flat lowland or levels	0.154***	1.634*	2.454***	9.112**
Hills, lower plateau, and scarp slopes	0.037***	1.017	5.720***	8.176**
Rolling lowland	0.038***	2.167***	2.769***	6.701**
Developed unbuilt land	1.276	1.005	0.375**	5.262*
Built land (reference)	-	-	-	-

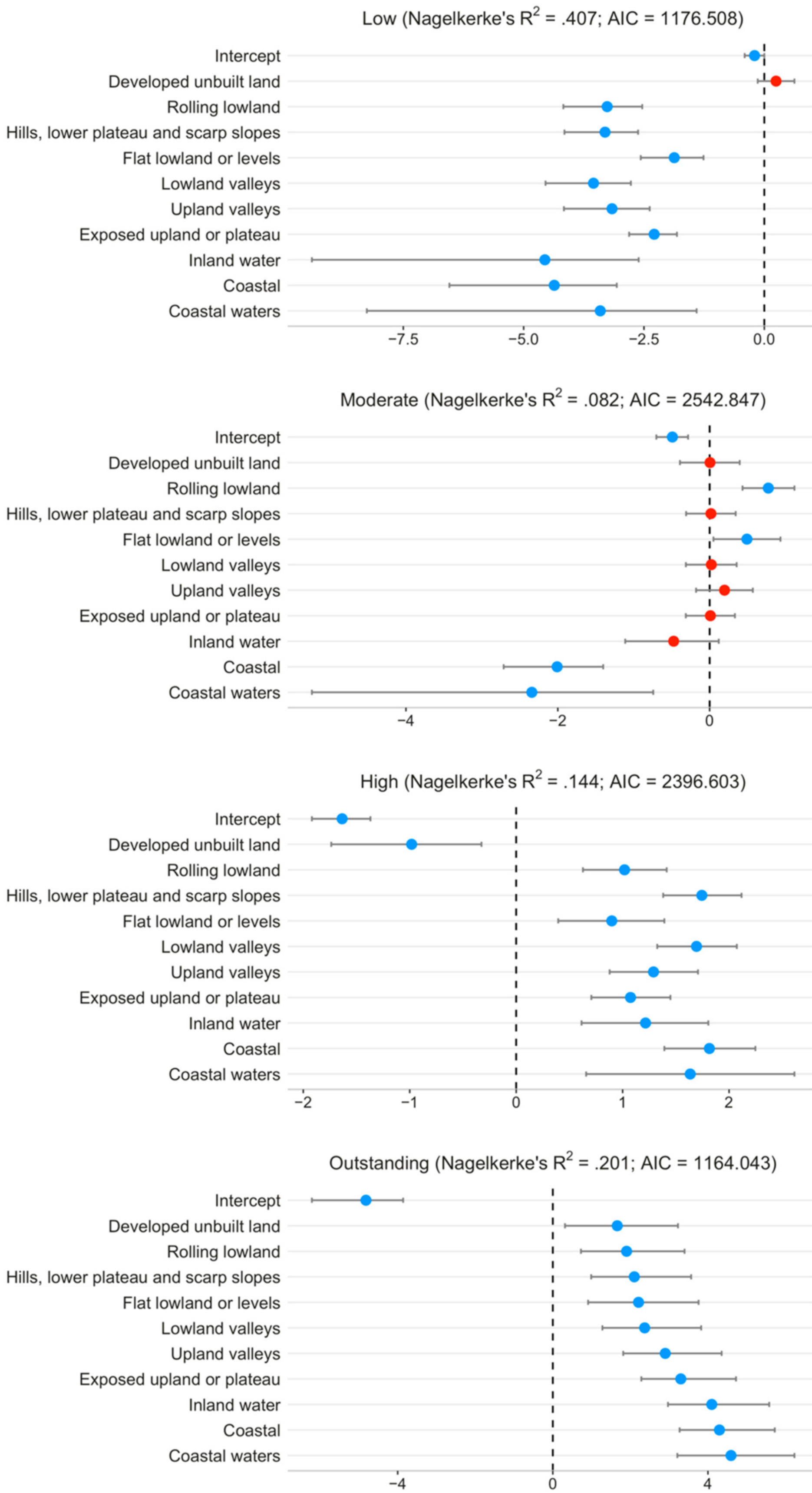

Low (Nagelkerke's R2 = .407; AIC = 1176.508)
Intercept
Developed unbuilt land
Rolling lowland
Hills, lower plateau and scarp slopes
Flat lowland or levels
Lowland valleys
Upland valleys
Exposed upland or plateau
Inland water
Coastal
Coastal waters
−7.5
−5.0
−2.5
0.0
Moderate (Nagelkerke's R2 = .082; AIC = 2542.847)
−4
−2
0
High (Nagelkerke's R2 = .144; AIC = 2396.603)
−2
−1
0
1
2
Outstanding (Nagelkerke's R2 = .201; AIC = 1164.043)
−4
0
4

Figure 4.3 The effect of various landforms on different levels of scenic quality were graphed on log scales, allowing a visual comparison of the magnitudes of confidence intervals and standard errors. Dots represent unstandardised point estimates (that is, log odds ratios) derived from the binary logistic linear regressions for the landform typologies concerning different levels of scenic quality evaluated by experts where the referent class was "built land." The vertical dash line represents the line of null effect, denoting there is no difference from the baseline. The goodness-of-fit of the model is indicated by Nagelkerke's R^2 and AIC measures. The error bar denotes 95% confidence intervals, indicating the uncertainty of the estimate. While the confidence interval crosses the line of null effect, the point estimate is statistically significant, denoted by dot colour (blue: $p < 0.05$; red: $p \geq 0.05$).

4.3.4 Summary of non-expert perspectives

Since the average scenic scores aggregated over the aspect areas exhibit a normal distribution, this average scenicness is regressed on the dummy-coded covariate, using an OLS model. The result is summarised in Table 4.5, and the coefficient estimates yielded are statistically significant at the 0.001 level. This suggests there are significant discrepancies in collective scenic perceptions between different landforms that could be informed by the map-based characterisation procedure. The intercept corresponds to the average scenic scores for the reference category ('built land') and the individual coefficient estimate of each landform class denotes the expected difference in the mean of scenic ratings compared to the baseline one. For example, the 'upland valleys' and 'lowland valleys' predict average scenicness around 1.7 and 1.4 greater than that of the baseline category, respectively.

4.3.5 Comparison of perspectives between experts and non-experts

The marginal effects of the landform typologies on the public scenic ratings and the expert evaluations of scenic quality could be used to further produce different ranking orders. These relative rank positions enable the comparison of the orders from both perspectives. Figure 4.4 illustrates the changes in order between both sides across four levels of scenic quality. The larger rank-order differences (i.e., change greater than or equal to 3 position) are highlighted in green (≥ 3) and red (≤ -3) and the rest are shaded in grey, showing which typology may be contesting from both perspectives. A relatively small amount of significant change in rankings is seen at the outstanding scenic quality, suggesting there exists mutual consent at the highest end of the spectrum. There is, however, a large amount of significant disparity in ranking at the level of high scenic quality. Moreover, the landform typologies associated with water bodies are generally ranked in the top three whilst those associated with man-made landscapes are bottom-ranked.

Table 4.5 Results of the simple linear regression, examining the relationship between the dummy-coded level-2 classes and the 'built land' class is used as the reference category and the average scenic ratings that are aggregated over the units of the visual and sensory aspect areas (*$p < .05$, **$p < .01$, ***$p < .001$).

Categorical variable (LANDMAP Level-2 class)	Number	Coefficient estimate	Standard error	t-value	p-value
Intercept	-	3.204	0.058	55.089	0.000***
Coastal waters	14	2.443	0.278	8.800	0.000***
Coastal	117	2.257	0.110	20.440	0.000***
Inland water	40	2.020	0.171	11.827	0.000***
Exposed upland or plateau	256	1.908	0.086	22.164	0.000***
Upland valleys	157	1.652	0.100	16.558	0.000***
Lowland valleys	216	1.400	0.090	15.499	0.000***
Flat lowland or levels	91	0.690	0.121	5.684	0.000***
Hills, lower plateau, and scarp slopes	241	1.408	0.088	16.088	0.000***
Rolling lowland	201	1.015	0.092	11.001	0.000***
Developed unbuilt land	78	0.455	0.129	3.528	0.000***
Built land (reference)	305	-	-	-	-

$R^2 = 0.331$; AIC $= 4935.927$

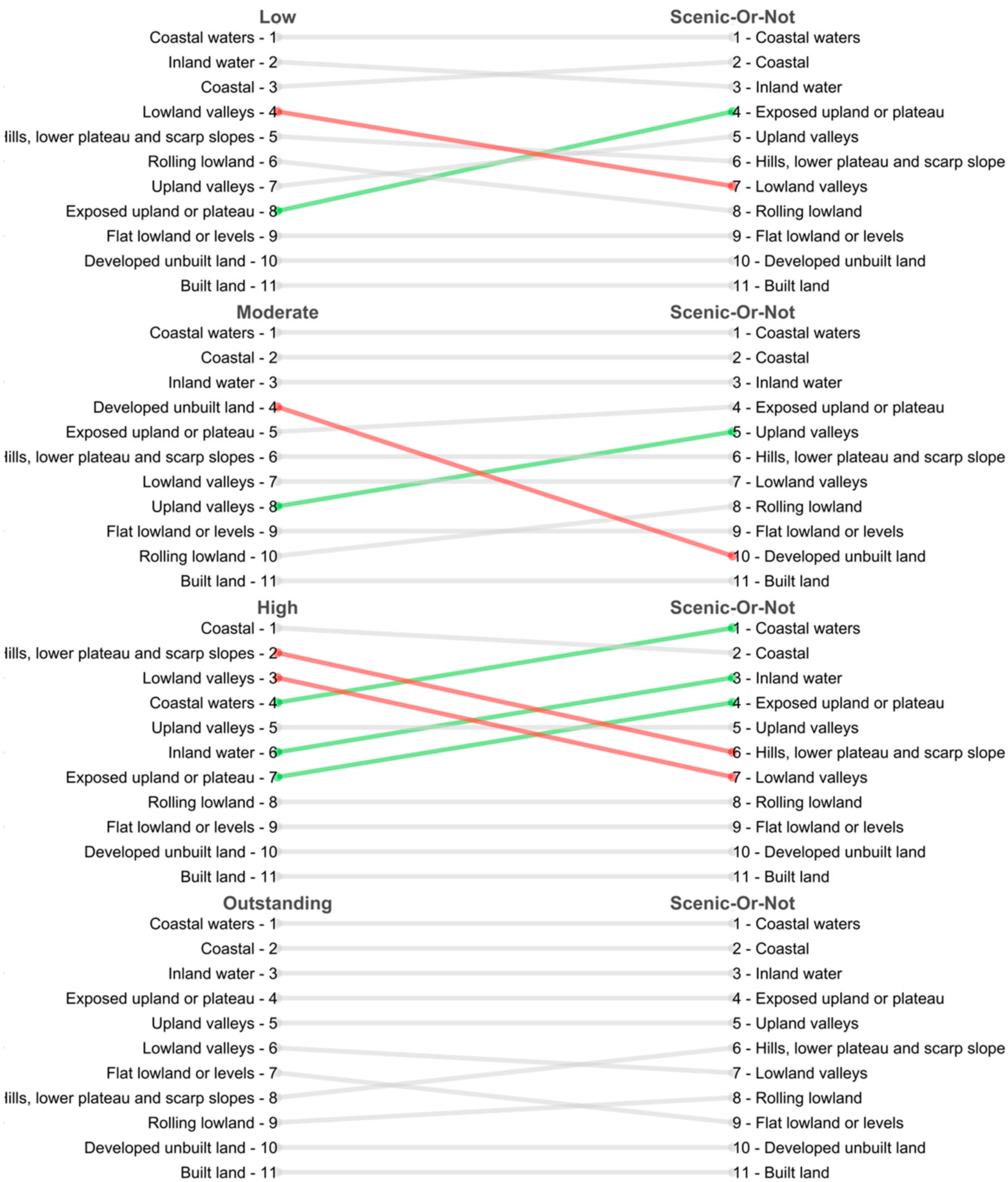


Figure 4.4 Landform rankings of scenic quality were compared between experts and non-experts, based on the results of the four GLMs and the OLS. The larger rank differences are highlighted where changes of greater than 3 or less than 3 are coloured green and red, respectively, and the rest are shown in grey.

4.4 Discussion

4.4.1 Implications for Landscape Character Assessment

The exploratory analysis provides some insights into the central tendency and variability in the public perceived scenicness within 'aspect areas.' Defined as regions sharing similar physical attributes, these 'aspect areas' are perceived as homogeneous in quality by the landscape architects. Broadly speaking, areas of water received the highest scores on average scenicness, followed by those of upland, lowland, and development, with a sole exception of "coastal" areas that recorded higher scenic values than the upland areas but are categorised as lowland. Concerning the underlying variability indicated by entropy metrics, upland areas were found to be highly variable, followed by lowland, water, and development areas.

It has been acknowledged in the field of landscape assessment that some landscape characteristics—terrain, water, ground cover, and human artefacts—are permanently recognised as contributing factors of scenic quality (Wherrett, 1996). In line with the literature and irrespective of professional expertise, the water-related typologies were ranked in the top three among all the landforms, while those associated with human artefacts were on the opposite end (Real et al., 2000a; Arriaza, J.F. Cañas-Ortega, et al., 2004). Meanwhile, the underlying scenic ratings of the characterised units associated with water and human artefacts features tend to exhibit less variability, reflecting consensual values in scenic quality. Also, this observation may suggest the formal classification scheme and evaluation for these characters could effectively correspond to the perceptions on the ground, compared to other landforms. This is, nevertheless, not the case for the landforms with a complex interaction of terrain and ground cover, resulting in the diversity of physical landscape characteristics as well as the variability of landscape perceptions. The "exposed upland or plateau" typology particularly exemplifies this type of landform where there is a consistently high amount of variation in the scoring, depicting the extreme variability in opinions, given the high centre and narrow spread of entropy measures. A possible justification is that an increase in relief implies diversity and complexity involved therein, where the topographical and meteorological effects were probably magnified. This may lead to positive or negative influences on landscape perceptions (e.g., the grandeur of mountains or harsh weather conditions). However, these landscape intricacies could be overlooked and subsumed within one type by the simplicity of the current spatial framework.

The most plausible explanation of this finding might be that the classification scheme rests on the assumption that visual quality is an amalgamation of landform and land cover elements where a variable relationship presumably exists. As the landform characteristics increase in dimension, for instance, the increase of relative elevation in magnitude (e.g., from flatlands through hills to mountains), the land cover pattern becomes less important as an element of visual quality. In contrast, with the decrease in landform dimension, the importance of the variety of ground cover increases for the maintenance of high scenic value (Zube, 1984). However, the complex interrelation of landforms and land covers on human perception, such as the visual screening effects caused by vegetation or adjacent hills may hardly be accounted for by the current practice. Hence, the classification method would remain an appropriate basis of appraisal for low-lying landscapes, but any interactive effects of the individual elements could be overlooked (Crofts, 1975).

The comparison of aesthetic assessments between the experts and non-experts also has implications for the formal landscape assessment paradigm in the context of public participation. Regardless of professional expertise, water bodies (that is, "inland water" and "coastal waters") were generally perceived to possess comparatively high scenic quality among all the landforms by the public, which was also the case in the expert evaluations. Likewise, landscapes dominated by human-imposed changes (that is, "developed unbuilt land" and "built land") were more likely rated as having low scenic quality from both sides. Furthermore, the pairwise ranking comparisons of subjectivity in terms of the geomorphological effects on scenic quality reveals that mutual consent is located towards the upper end of the scenic spectrum. There was a relatively small amount of significant change in rankings at the highest end of the scale, and less consensus was found in the middle of the spectrum. It may be noteworthy that at the opposite end of the spectrum, the "exposed upland or plateau" typology tends to be more appreciated by the laymen, rising four places to fourth. In contrast, the ranking of the "lowland valleys" typology drops from fourth to seventh position. Given low scenic quality generally typified by slightly undulating topography and monotonous patterns of vegetation, this may suggest the scenic quality of these two landforms could be under- (or over-) estimated by specialists. The virtue of the hierarchical classification approach allows identification of the most likely contesting descendant by digging deeper into the finer classes (that is, level 3 typologies). The textual information associated with those scenic ratings can also be retrieved in the sourced repository that gives researchers insight into what landscape features might be ignored; however, that is beyond the scope of this study.

4.4.2 Limitations and outlook

The critical issues in crowdsourced data are centred around their reliability and veracity (Brunsdon and Comber, 2021). The demography of the participants of Scenic-Or-Not is untold, and the reliability of scenic ratings may not be asserted due to the inherent biases concerning participant inequality (Bubalo et al., 2019). For instance, without demographic characteristics of participants, the representativeness of the scenic ratings for the general public is questionable for meeting the research context (Goodspeed, 2017). Biases, introduced by a small group of prolific contributors (Koblet and Purves, 2020), could not be filtered out in parallel as well. Albeit with these issues, researchers have continuously sought the best practice to harness the benefits of these increasingly growing sources, facilitating the understanding of dynamic landscape perceptions and preferences (Goodspeed, 2017; Bubalo et al., 2019).

Furthermore, there is no guarantee for the geographical representation of photographs employed in the scenic rating campaign despite being sourced from Geograph.org where a quality control protocol regarding the image and location information of each contribution exists. Additionally, for the Scenic-Or-Not, every square kilometre contains only one photo where landscapes with diverse and complex characteristics may hardly be captured fully and nothing is known about its actual mechanism behind image selection (Chang Chien et al., 2020). Moreover, unlike site assessment where 360° views from a given standpoint could be evaluated, the participants of Scenic-Or-Not only evaluated a particular vista framed in the given photo without knowing any local context information, which is very much at the mercy of image composition and has been recognised as the pitfall of photographic preference surveys (Unwin, 2006). Hence, this study was notably based on the untenable assumption that the crowdsourcing initiative adequately provides a discerning measure for public landscape perceptions.

It should be noted that the scenic ratings evenly distributed across the entire country were confined within the expert delineated boundaries as the statistical inferences were based on the units of the aspect areas. This parallels the compromise solution of handling public perceptions in the earlier work, conducted at the county level (Scott, 2002). An interlinked challenge, therefore, remains for landscape planners and architects to test the validity of expert-led landscape characterisation as reflective of on-the-ground experiences. The empirical results of this national study may be reconciled with the existing evaluations to achieve a further improved practice of LANDMAP systematically.

Landscape quality evaluations of either areas or points as discussed herein fail to capture the quality of the scenery, experienced and perceived by a viewer from the point the viewer stands in all directions. Recent developments in GIS-based viewshed analyses (e.g., vertical voxel viewsheds), taking the viewer's contexts (for example, viewpoints, distance-decay effects) into account to model visual landscape experiences on the ground, shed light on measuring people's experiences of landscape characters through an automated process. These metrics could be integrated with landscape preference judgements and subsequently converted to landscape quality (Carver and Washtell, 2012;

Brabyn, 2015). Recent studies further investigated the diverse contents of crowdsourced data, ranging from geographical to textual and imagery dimensions, to gain a variety of perceptual details and contextual information of landscapes (Callau et al., 2019; Koblet and Purves, 2020). They demonstrate potential ways to alleviate the challenges of current LCA, efficiently eliciting multiple perspectives—that are not expert-dominated —and involve other sensory features – that are not predominantly visual – of landscapes. In pursuance of more informed landscape policy- and decision-making, the incorporation of such supplementary information with respect to public perspectives into a practical landscape assessment should remain a major research avenue to explore.

4.5 Conclusions

Despite calls for the inclusion of public opinions into a formal paradigm of landscape assessments, an integrated solution of the two has not yet been achieved due to the methodological limitations and the deficiency of large-scale surveys. The crowdsourcing paradigm provides a viable solution for efficiently eliciting large-scale public perspectives on landscape aesthetics. The present study has been one of relatively few attempts to investigate a potential synergy of crowdsourced data to supplement a practical landscape assessment, albeit with the acknowledgement of inevitable biases embedded in these data. The authoritative data, delineating the bespoke landscape characters concerning overall perceptual quality, has not sufficiently addressed situations of topographical diversity, such as screening effects from adjacent hills. The results show some mutual consent in landforms perceived as scenic or unattractive by experts and non-experts which are consistent with the previous landscape perception studies and suggest some potentially contested landscape typologies from both sides. It is concluded that there are potential opportunities to develop landscape metrics for the assessment of visual landscape perception to better reflect the perceived landscape character on the ground by utilising the established GIS viewshed approaches. These landscape metrics could be further combined with landscape preference information to contribute to a gradual improvement of the authoritative spatial framework in the evaluation of landscape scenic quality.

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*.

Antrop, M. and Van Eetvelde, V. 2017. *Landscape Perspectives*. Dordrecht: Springer Netherlands.

Arriaza, M., Cañas-Ortega, J. F., Cañas-Madueño, J.A. and Ruiz-Aviles, P. 2004. Assessing the visual quality of rural landscapes. *Landscape and Urban Planning*. **69**(1), pp.115–125.

Brabyn, L. 2015. Modelling landscape experience using “experions”. *Applied Geography*. **62**, pp.210–216.

Brown, G. and Brabyn, L. 2012. An analysis of the relationships between multiple values and physical landscapes at a regional scale using public participation GIS and landscape character classification. *Landscape and Urban Planning*. **107**(3), pp.317–331.

Brunsdon, C. and Comber, A. 2020. Opening practice: supporting reproducibility and critical spatial data science. *Journal of Geographical Systems*. (0123456789).

Bubalo, M., van Zanten, B.T. and Verburg, P.H. 2019. Crowdsourcing geo-information on landscape perceptions and preferences: A review. *Landscape and Urban Planning*. **184**, pp.101–111.

Butler, A. 2016. Dynamics of integrating landscape values in landscape character assessment: the hidden dominance of the objective outsider. *Landscape Research*. **41**(2), pp.239–252.

Callau, A.À., Albert, M.Y.P., Rota, J.J. and Giné, D.S. 2019. Landscape characterization using photographs from crowdsourced platforms: content analysis of social media photographs. *Open Geosciences*. **11**(1), pp.558–571.

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

Chang Chien, Y.M., Carver, S. and Comber, A. 2020. Using geographically weighted models to explore how crowdsourced landscape perceptions relate to landscape physical characteristics. *Landscape and Urban Planning*. **203**(August), p.103904.

Chesnokova, O., Nowak, M. and Purves, R.S. 2017. A crowdsourced model of landscape preference *In*: *13th International Conference on Spatial Information Theory (COSIT 2017)*. L'Aquila, Italy, pp.1–13.

Conrad, E., Christie, M. and Fazey, I. 2011. Is research keeping up with changes in landscape policy? A review of the literature. *Journal of Environmental Management*. **92**(9), pp.2097–2108.

Conrad, E., F. Cassar, L., Jones, M., Eiter, S., Izaovičová, Z., Barankova, Z., Christie, M. and Fazey, I. 2011. Rhetoric and Reporting of Public Participation in Landscape Policy. *Journal of Environmental Policy & Planning*. **13**(1), pp.23–47.

Council of Europe 2000. *European Landscape Convention*. Florence.

Crofts, R.S. 1975. The Landscape Component Approach to Landscape Evaluation. *Transactions of the Institute of British Geographers*. (66), pp.124–129.

Daniel, T.C. 2001. Whither scenic beauty? Visual landscape quality assessment in the 21st century. *Landscape and Urban Planning*. **54**(1–4), pp.267–281.

Dramstad, W.E., Tveit, M.S., Fjellstad, W.J. and Fry, G.L.A. 2006. Relationships between visual landscape preferences and map-based indicators of landscape structure. *Landscape and Urban Planning*. **78**(4), pp.465–474.

Dunkel, A. 2015. Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning*. **142**, pp.173–186.

Goodchild, M.F. 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal*. **69**(4), pp.211–221.

Goodspeed, R. 2017. Research note: An evaluation of the Elo algorithm for pairwise visual assessment surveys. *Landscape and Urban Planning*. **157**, pp.131–137.

Gosal, A.S. and Ziv, G. 2020. Landscape aesthetics: Spatial modelling and mapping using social media images and machine learning. *Ecological Indicators*. **117**(May), p.106638.

Groom G., Wascher D., Potschin M. and Haines-Young R. 2006. Landscape character assessments and fellow travellers across Europe: A review *In*: *Proceedings of the European IALE Conference, 29 March - 2 April 2005*.

Gyllin, M. and Grahn, P. 2015. Semantic Assessments of Experienced Biodiversity from Photographs and On-Site Observations – A Comparison. *Environment and Natural Resources Research*. **5**(4), p.46.

Jones, M. 2007. The European landscape convention and the question of public participation. *Landscape Research*. **32**(5), pp.613–633.

Koblet, O. and Purves, R.S. 2020a. *Extracting Perceived Landscape Properties from Text Sources*. Zurich.

Koblet, O. and Purves, R.S. 2020b. From online texts to Landscape Character Assessment: Collecting and analysing first-person landscape perception computationally. *Landscape and Urban Planning*. **197**(January), p.103757.

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), pp.177–198.

Myers, R.H. and Montgomery, D.C. 1997. A tutorial on generalized linear models. *Journal of Quality Technology*. **29**(3), pp.274–291.

Nagelkerke, N.J.D. 1991. A note on a general definition of the coefficient of determination. *Biometrika*. **78**(3), pp.691–692.

Natural Resources Wales 2017. *LANDMAP Guidance Note 1 : LANDMAP and Special Landscape Areas*.

Real, E., Arce, C. and Sabucedo, J.M. 2000. Classification of landscapes using quantitative and categorical data, and prediction of their scenic beauty in north-western Spain. *Journal of Environmental Psychology*. **20**(4), pp.355–373.

Santé, I., Tubío, J.M. and Miranda, D. 2020. Public participation in defining landscape planning scenarios and landscape quality objectives (LQO): Landscape Guidelines for Galicia (NW Spain) case study. *Land Use Policy*. **94**(February), p.104559.

Scott, A. 2002. Assessing public perception of landscape: the LANDMAP experience. *Landscape Research*. **27**(3), pp.271–295.

Seresinhe, C.I., Moat, H.S. and Preis, T. 2018. Quantifying scenic areas using crowdsourced data. *Environment and Planning B: Urban Analytics and City Science*. **45**(3), pp.567–582.

Seresinhe, C.I., Preis, T., MacKerron, G. and Moat, H.S. 2019. Happiness is Greater in More Scenic Locations. *Scientific Reports*. **9**(1), p.4498.

Seresinhe, C.I., Preis, T. and Moat, H.S. 2015. Quantifying the Impact of Scenic Environments on Health. *Scientific Reports*. **5**(1), p.16899.

Seresinhe, C.I., Preis, T. and Moat, H.S. 2017. Using deep learning to quantify the beauty of outdoor places. *Royal Society Open Science*. **4**(7), p.170170.

Shannon, C.E. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal*. **27**(3), pp.379–423.

Simensen, T., Halvorsen, R. and Erikstad, L. 2018. Methods for landscape characterisation and mapping: A systematic review. *Land Use Policy*. **75**, pp.557–569.

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

Tenerelli, P., Demšar, U. and Luque, S. 2016. Crowdsourcing indicators for cultural ecosystem services: A geographically weighted approach for mountain landscapes. *Ecological Indicators*. **64**, pp.237–248.

Terkenli, T.S. 2001. Towards a theory of the landscape: The Aegean landscape as a cultural image. *Landscape and Urban Planning*. **57**(3–4), pp.197–208.

Unwin, K.I. 2006. The relationship of observer and landscape in landscape evaluation. *Transactions of the Institute of British Geographers*. (66), p.130.

Weledol, M. and Landmap, S. 2016. LANDMAP Methodology Visual and Sensory.

Wherrett, J. 1996. Visualization techniques for landscape evaluation: Literature review. *The Macaulay Land Use Research Institute*. [Online]. [Accessed 16 February 1996]. Available from: https://macaulay.webarchive.hutton.ac.uk/visualisationlitrev/chapters.html.

van Zanten, B.T., Van Berkel, D.B., Meentemeyer, R.K., Smith, J.W., Tieskens, K.F. and Verburg, P.H. 2016. Continental-scale quantification of landscape values using social media data. *Proceedings of the National Academy of Sciences*. **113**(46), pp.12974–12979.

Zube, E.H. 1984. Themes in landscape assessment theory. *Landscape Journal*. **3**(2), pp.104–110.

Zube, E.H. and Pitt, D.G. 1981. Cross-cultural perceptions of scenic and heritage landscapes. *Landscape Planning*. **8**(1), pp.69–87.

Chapter 5 Integrating objective and subjective landscape assessments

Abstract

This study integrates citizen generated assessments of landscape scenicness with measures of landscape wildness to construct models of landscape aesthetic quality. In so doing it integrates different landscape paradigms related to citizen and expert evaluations, and to objectivist landscape measurements and subjectivist perceptions, in order to train ensemble machine learning models using extreme gradient boosting (XGBoost). These were constructed and validated using the LANDMAP classification of Wales and then applied to Great Britain to construct predictive maps of landscape aesthetic quality at 1 km resolution. Two distinct models were developed to represent the assessments conducted by two separate expert groups, each responsible for the LANDMAP classification in the northern and southern parts of Wales, respectively. Both models achieved an overall classification accuracy of over 67% and their predictions of landscape aesthetic quality were compared. Both models identified the same regions, in the South of Scotland in particular, that are not currently under any form of landscape designation, but potentially could be. Some methodological implications are discussed, including the opportunity to integrate citizen derived perceptions of landscape quality and spatial data describing the formal wildness measures into current evaluations of landscape character, such as the Landscape Character Assessment (LCA) in the UK. A number of areas of further work are suggested including some detail about how current LCA practice could be advanced and in so doing make it more consistent, tractable, democratic and accountable.

Keywords: Landscape Character Assessments, Crowdsourcing, Wildness, Scenic-Or-Not.

5.1 Introduction

Landscapes possess many attributes values, with common examples including aesthetic, recreational and economical value, not all of which are material and tangible, such as those exclusively held by people. Landscape scenic beauty (and more broadly, environmental aesthetics) is a closely intertwined product of human-nature interactions and has been one of the most emphasised landscape values and as such widely is recognised as providing benefits to social well-being and a healthy society (Selman and Swanwick, 2010).

Research regarding landscape aesthetics and aesthetic quality has typically taken one of two different but not mutually exclusive routes based on objectivist or subjectivist approaches (Lothian, 1999). The objectivist paradigm mainly relates to expert-based approaches that aim to identify physical landscape properties related to assumed human preferences. This is exemplified by the landscape metrics commonly used in the landscape ecology literature that are used to model and quantify landscape scenic quality (Palmer, 2004). In contrast, the subjectivist paradigm focuses on perceptual or judgemental responses of observers. For example, the preference surveys employed in some landscape studies use photos as proxies in order to investigate the link between respondent preferences, normally measured on a Likert-type scale, and physical landscape features (James F Palmer and Hoffman, 2001) and these measures can be associated with respondent attributes and socio-cultural backgrounds. Both paradigms are complementary and some theoretical work has been undertaken to, for instance, link landscape evolutionary theories (Appleton, 1975; Kaplan and Kaplan, 1989) with cultural preference ones (Allen Carlson, 2001), where landscape beauty is conceptualised to be in the eye of the beholder (M Tveit et al., 2006).

This paper seeks to integrate both expert, objectivist and citizen perceptions of landscape in order to construct predictive models of landscape character, and thereby to support landscape assessments and policy-making. The models are applied over new areas to predict landscape character in those places, and to suggest potential areas suitable for designation. The aim was to demonstrate a novel method for assessing landscape character, one that integrates landscape perceptions held by citizens and professionals. The rationale and potential data sources for this work are described in the next section.

5.2 Background

5.2.1 Formal Landscape Assessments

The adoption of the European Landscape Convention (ELC) in 2000 resulted in a definition of landscape as "an area, as perceived by people" (Council of Europe, 2000), which started to inform broader notions of Ecosystem Service (Boris T van Zanten et al., 2016). The Millennium Ecosystem Assessment (MEA) categorised landscape aesthetics as 'non-material' cultural services (MA, 2005), reflecting a wider sense of landscape value and a greater emphasis on public involvement in landscape initiatives. Despite this development, a commonly accepted definition of landscape 'quality' has not emerged because of its intangible, and often subjective nature (Santé et al., 2020). Consequently, the views of the public are often bypassed and technical, expert led evaluations continue to prevail in landscape assessments (Conrad, F. Cassar, et al., 2011; Butler and Berglund, 2014).

Due to the devolved nature of environmental policy matters, each country within the United Kingdom upholds its own landscape policies. These policies align with the principles of the ELC, which they have signed and ratified. Implementation of these policies is guided by the Landscape Character Assessment (LCA) framework (Swanwick, 2002). The essence of this approach is to identify areas with a uniform character and to describe the associated attributes or qualities that distinguish a given area via a hierarchical classification, which can be applied at a number of different scales (Simensen et al., 2018). Its process is primarily comprised of two iterative stages: desk-based characterisation and fieldwork evaluation, involving both objective and subjective assessments (Terkenli et al., 2021). The initial characterisation is considered to be objective although the subjectivity of the prescriptive hierarchical classification system is frequently overlooked (Crofts, 1975), while the valuation procedure may be subjective due to variations in professional or institutional perspectives, and with qualitative assessments that typically have a more descriptive rather than comparative focus.

In Wales, a detailed national landscape assessment methodology was devised: the LANDMAP (Landscape Assessment Decision Making Process, LANDMAP). This aims to provide baseline characteristics describing landscape character, quality and value, recorded in different mapped layers to inform wider practices (Scott, 2002). The development of a national classification scheme for different landscape types, supports consistent land use planning and policy, such as designations and visual impact assessments at national and local levels. The amalgamation of these aspects – character, quality and value – using GIS layers supports transparency and accountability in decision-making. However, the values and definitions of quality landscape may not adequately represent public perspectives, and professional valuations of the landscape may vary.

5.2.2 Crowdsourcing

In recent decades, the proliferation of location-aware, connected devices and services has enabled public landscape experiences and locally relevant landscape perceptions to be captured (Boris T van Zanten et al., 2016; Callau et al., 2019). Examples include Geograph (https://m.geograph.org.uk) in which the public upload and share pictures of Ordnance Surveys 1 km grid square. Users can include captions as well as location information and tags. The project aims to establish an archive of photographs at eye-level over every square kilometre in Great Britain and Ireland. Such georeferenced information, often referred to as volunteered geographic information (VGI) (Michael F Goodchild, 2007), can be of immense value to both academia and practitioners who seek to supplement and enrich land management and planning (Dunkel, 2015; Bubalo et al., 2019).

The existence of Geograph supported a further internet-based crowdsourcing campaign—Scenic-Or-Not (https://scenicornot.datasciencelab.co.uk). This randomly selects Geograph photos and asks users to rate their scenic beauty on an integer scale from 1 (the least scenic) to 10 (the most scenic). Over 1.5 million contributions have been made since February 2015, enabling quantitative analyses of perceived landscape environmental aesthetics (Seresinhe et al., 2018) and the relationship between landscape scenicness and wildness (Chang Chien et al., 2020). In landscape studies, the reliability and validity of photographic media is generally accepted as a fair representation of portions of landscape's visible condition (James F Palmer and Hoffman, 2001; Palmer, 2004). The Scenic-Or-Not, thus, offers an otherwise unavailable measure of landscape scenic beauty, as held by the wider public, despite data sparsity in some areas, related to the Geograph coverage. The granularity and spatial coverage of the data allows it to be used in support of national and regional landscape policy and planning.

5.2.3 Integrating Landscape Wildness

Wildness (or more widely wilderness) and aesthetics share some conceptual common ground in landscape assessment, with similar dependencies on landscape spatial form, components and structure. Approaches using GIS-based wildness mapping, e.g. (Carver et al., 2002; Carver et al., 2013), can provide spatially explicit and quantitative insights into areas of potential conservation value (Carver et al., 2012). In a recent study, 80% of the variation in the crowdsourced measures of landscape aesthetics were explained by the formal wildness measures (Chang Chien et al., 2020). This suggests that there are opportunities for such wildness measure to be used to complement perception-based data in an integrated framework, thereby linking both subjectivist and objectivist paradigms and potentially providing a more holistic approach to landscape studies and assessments.

5.2.4 Study aims

A comparative and systematic assessment of landscape aesthetic value, in the manner of LANDMAP in Wales, has so far not been developed in England and Scotland. The valuation of landscape aesthetics encapsulated in LANDMAP may be transferable to these countries. To this end, a data-driven framework using machine learning techniques was developed. The gradient boosting model framework has been found to have high prediction accuracy, flexibility, scalability, transparent outputs, and ease of implementation(Chen and Guestrin, 2016). It can handle data sparsity and is not affected by multicollinearity.

This study constructs a model of LANDMAP classes using data from the Scenic-Or-Not initiative and wildness spatial layers, described below, utilising the XGBoost algorithm. After calibration and tuning, the model is then applied to Scenic-Or-Not and wildness data to predict the LANDMAP classes for England and Scotland. As the LANDMAP data of landscape aesthetics in Wales was undertaken by two separate expert groups, two XGBoost models were constructed and applied to predict the spatial distribution of landscape scenic quality across England and Scotland (as well as Wales), where such assessments are lacking: there is no GB wide assessment of landscape aesthetic quality.

5.3 Data and methods

5.3.1 Data

LANDMAP is comprised of five spatial datasets concerning the geological, ecological, visual and sensory, historic and cultural facets of the landscape. These are used to construct areas whose scenic quality was then evaluated by one of two consulting groups, referred to as Consultant A and B (Figure 5.1). The scenic quality of each area was labelled using an ordinal scale from Low (little or no importance), Moderate (local importance), High (regional and county importance) to Outstanding (international or national importance). In this study the scenic quality was used as the target variable in the models.

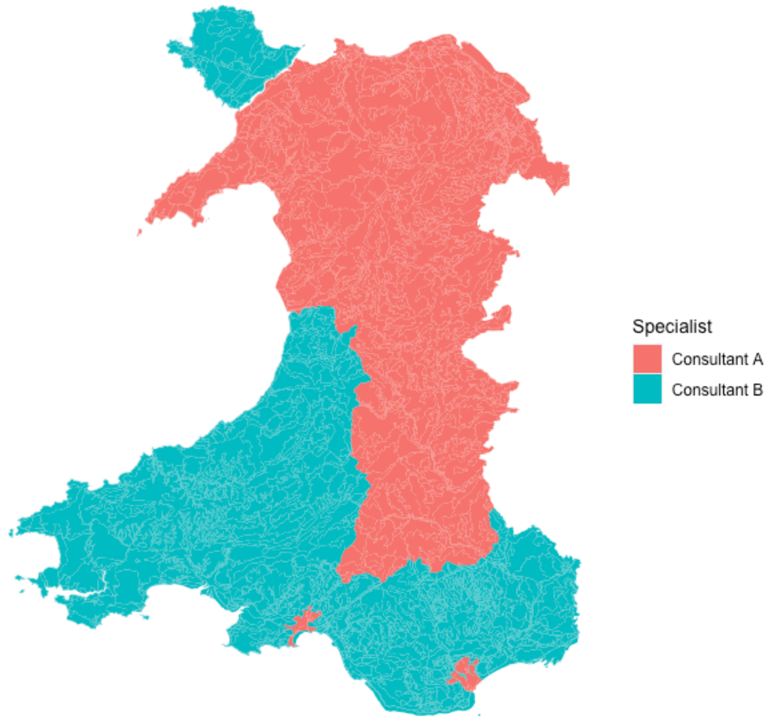


Figure 5.1 The LANDMAP project split of Wales into two subsets for assessment by Consultant A and B.

The Scenic-Or-Not data set includes ratings for 212,212 of the Geograph images covering nearly 80% of the Ordnance Survey (OS) 1 km grid squares of Great Britain. The data includes each grid square contains one picture for which there is at least three ratings from unknown users. These subjective measures were referred to hereafter as 'scenicness'. Grid cells with missing values and photos are concentrated in remote areas such as the Scottish Highlands.

Formal wildness measures were also examined. Wildness has been calculated from four components: the absence of modern human artefacts, perceived naturalness of land cover, remoteness from mechanised access, and rugged and challenging terrain. These are briefly described in Table 5.1 and more detailed description can be found in (Carver et al., 2012; Chang Chien et al., 2020). Each component was given equal weighting to create a wildness layer with a resolution of 25 m.

In the LANDMAP data, some character areas were represented by non-contiguous multi-polygons. These were divided into individual areas to serve as units of analysis. The point-based scenicness scores and grid-based wildness layer were aggregated to these areas to capture central tendency and dispersion. However, these discrete areas did not align well with the spatial structure of continuous wildness indices, leading to a skewed distribution. On the other hand, the aggregate scenicness demonstrated a normal distribution. Given these observations, the resistant measures were chosen for the wildness covariates and the sensitive measures were employed for the scenicness covariates. For each of these character areas, the median and interquartile range (IQR) of the intersecting wildness indices, and the mean and entropy of intersecting scenicness were determined. Thus, the LANDMAP evaluations of scenic quality were modelled using four predictor features: the mean of scenicness, the median of wildnesss, the entropy of scenicness, and the IQR of wildness.

The 2,968 LANDMAP areas included 1,315 whose aesthetic values were determined by Consultant A and 1,653 by Consultant B, with uneven distributions among different aesthetic quality classes (Table 5.2). The data for each consultant was subject to an 80/20 split into training and validation subsets to support the creation of two models, one for each consultant. This was done using a within class stratification to ensure representative split across classes and consultants.

Table 5.1 The components of the wildness layer.

Wildness component	Description
Absence of modern human artefacts	The visual absence of man-made structures—including linear, non-natural vegetation, built, engineering, and novel industrial features—within a 360-degree panoramic view of each location. This is achieved by quantifying the visible proportions of these structures using a novel viewshed approach that accounts for both the horizontal and vertical aspects, as well as distance decay effects.
Perceived naturalness of land cover	A reclassification of land cover data, in which each class was allocated a naturalness score of 0-5 based on its degree of human intervention. Under the assumption of a visibility limit, the area weighted mean naturalness score for a given location was calculated within a 250-metre radius for each location.
Remoteness from mechanised access	Remoteness refers to the time needed to walk to a destination from the nearest road access. It is essentially the cumulative cost surface based on an assumed speed of 5 km per hour over flat terrain with cross slope correction.

Rugged and physically challenging nature of the terrain	Physical variations in terrain morphology, as well as weather conditions caused by the nature of the terrain, primarily calculated from 2 standard deviations of terrain curvature within a 250-metre radius of each location.

Table 5.2 The distribution of the LANDMAP aesthetic quality classes across the two consulting groups.

Consultant	NA	Low	Moderate	High	Outstanding	Total
A	9	164 (13%)	473 (36%)	549 (42%)	120 (9%)	1,315
B	1	273 (17%)	575 (35%)	563 (34%)	241 (15%)	1,653

5.3.2 Analysis

EXtreme Gradient Boosting (XGBoost) is an efficient and scalable variant of the Gradient Boosting Machine (GBM), using classification and regression trees (CART) to predict the outcome variable (Chen and Guestrin, 2016). The individual leaf scores are summed up as a final score and evaluated through an additive function, as shown in Equation (5.1):

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^{K} f_k(x_i), f_k \in \mathcal{F} \tag{5.1}$$

where $\hat{y}_i$ represents the estimated output of the gradient boosting tree model, x_i is the feature corresponding to sample i, K represents the number of trees, and f_k is an independent tree structure with leaf scores in the functional space $\mathcal{F}$ which denotes the space of all possible CARTs.

The difference between GBM and XGBoost lies in model formalisation and computational power—parallelising the tree formation. Moreover, it comprises an objective function, which combines the loss function and a regularization term that controls model complexity and avoid over-fitting (Chen and Guestrin, 2016), which is given by:

$$Obj(\Theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega\,(f_k) \tag{5.2}$$

For a multi-class classification problem such as this, the first term l in Equation (5.2) is a cross-entropy (CE) loss function that compares the predicted probability distribution ($\hat{y}_{ij}$) with the expected probability distribution (y_{ij}) where y_{ij} denotes the probability for the i^{th} instance j^{th} class. The CE loss is computed given a list of ground-truth labels for a set of samples that was encoded as a $1\text{-}of\text{-}M$ binary indicator matrix – Y – as well as a matrix of probability estimates – P – with $p_{ij} = \Pr\,(y_{ij} = 1)$. The mathematic express is as follows:

$$L_{\log}(Y, P) = -\log\Pr(Y|P) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}\log(p_{ij}) \tag{5.3}$$

where $y_{ij} = 1$ if sample i has label j taken from a set of j labels, M is the number of classes and N is given data samples. Cross-entropy loss increases as the predicted probability diverges from the actual label.

The second term Ω is the ensemble regulariser, penalising both the number of the tree leaves and the $L2$ norms of its leaf weight vector, as defined by:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T} w_j^2 \tag{5.4}$$

where T and w are the number of leaves and the score on each leaf, respectively, and γ and λ are constants to control the degree of regularisation. Apart from the use of regularisation, shrinkage and descriptor subsampling provide two additional techniques used to prevent over-fitting.

5.3.3 Hyperparameter tuning and model training

Hyperparameters are manually set before training machine learning algorithms. The best set of these is determined by tuning. Hyperparameter tunning is needed to deal with the potential for bias variance trade-off, which plays a vital role in the learning process (Geman et al., 1992). The XGBoost model has a number of tuneable hyperparameters as shown in Table 5.3 and three tuning strategies are often employed – grid search, random search, and Bayesian optimisation. A grid search is simply an exhaustive searching through manually specified combinations of hyperparameters, with the best combination identified by some performance metric, such internal measures of model fit. The search space can be very large and computationally demanding and requiring too much time to be impractical.

A random search is similar to a grid search but the number of hyperparameter combinations that are evaluated is reduced by examining the hyperparameter space, using the statistical distribution for each hyperparameter (Bergstra and Bengio, 2012). Both grid and random searches can be inefficient since the optimisation procedure is uninformed by any past evaluations, within the search.

The Bayesian optimisation provides a solution to these issues (Bergstra et al., 2011). Here, it was undertaken using the MIBayesOpt R package. The Bayesian framework is comprised of two main components: a "surrogate" model of the objective function and an acquisition function. The idea is to optimise a probabilistic model for the objective function where the fitness evaluation of the former is much computationally cheaper than the latter (Snoek et al., 2012). An initial proxy function is specified based on the assumption of a Gaussian process (GP) (Rasmussen and Williams, 2006) and the posterior distribution over function is continuously updated with the added set of observations. The expected improvement (EI) function drawn from the posterior is commonly used as the acquisition function and hence chosen to indicate where to evaluate the function next in this study (Močkus, 1975).

During the training phase, a procedure of repeated k-fold cross validation was employed to increase the predictive performance of the classifiers (Kuhn and Johnson, 2013). The training set was randomly partitioned into 10 folds of roughly equal size, each of which was held out in turn for validation while the other 9 folds were used for training. Meanwhile, each fold maintained the approximate proportion of each class identified in the previous step to alleviate the potential classification deterioration. The entire procedure was iterated 10 times with different permutations of the training set and the mean performance across all folds were averaged.

Table 5.3. The XGBoost hyperparameters to be tuned.

Hyperparameter	**Definition**
`nrounds`	Boosting iterations
`eta`	Learning rate by which to shrink the feature weights
`max_depth`	Maximum tree depth
`gamma`	Minimum loss reduction
`Subsample`	Subsample percentage

`colsample_bytree`	Subsample ratio of columns
`min_child_weight`	Minimum sum of instance weight

XGBoost provides variable or feature importance metrics as part of its output to facilitate understanding of the model, clustered using k-means. The gain metric describes the average training loss gained when including a feature and provides the most relevant metric to interpret the relative importance of each feature.

5.3.4 Accuracy assessment

Model accuracy was quantified using confusion matrices, comparing predicted with observed classes in the test subsets. Several accuracy measures can be derived from the confusion matrix including Overall accuracy (the proportion of correctly classified observations) whose utility is reduced in case with imbalanced class distributions (Haibo He and Garcia, 2009) and Kappa accuracy which incorporates adjustments for random allocation agreement, and is more usually reported as an overall measure of accuracy (Foody, 2002).

5.3.5 Mapping outputs

The two models were applied to predict the evaluative level of scenic quality adopted in LANDMAP across the entirety of Great Britain, with a spatial resolution of 1 km. The Scenic-Or-Not data are reported over these units and the wildness attributes were aggregated to them in the following way: for each dataset, the central tendency and the variability within the 1 km were calculated (the mean and entropy for scenicness and the median and IQR for wildness) were calculated. The predictive models for both consultants were applied to these to predict the level of scenic quality for each grid square in Great Britain. The resulting maps capture the integration of different landscape assessment paradigms, and a bivariate choropleth map was created to compare the predictions from the two models.

5.4 Results

5.4.1 Hyperparameter optimisation and model building

The optimisation of hyperparameter tuning is important for predictive performance. A Bayesian approach to this was taken using the hold-out method and the *softmax* function was used to estimate the normalised probability [0, 1] for each class. The best combination of parameters for the 2 models are shown in Table 5.4.

Table 5.4 The results of the Bayesian optimisation of model hyperparameters.

Hyperparameter	Consultant A	Consultant B
nrounds	209.910	1433.996
eta	0.451	0.892
max_depth	12	4
gamma	0	0
subsample	0.758	0.500
colsample_bytree	1	1
min_child_weight	0.2	0.6

5.4.2 Classification performance

The confusion matrices for each classification are shown in Figure 5.2. The Overall classification accuracies are 67.3% for Model A (Consultant A) and 74.5% for Model B (Consultant B), and the Kappa statistics for these models were 0.5 and 0.64 respectively. This suggest some inconsistencies in the relationships between the input spatial features and the LANDMAP classes, with the prediction performance of Model B being "substantially good" (Landis and Koch, 1977; Cicchetti and Sparrow, 1981), explaining the variation of the target variable better than Model A.

In addition to these assessments, an evaluation was conducted where Model A was used to predict the test data sets of Consultant B and vice versa. These tests result in the Kappa statistics of 0.52 and 0.54 respectively, which served as an internal validation within the Wales dataset, examining the interoperability of the two models before considering spatial extrapolation to the UK. The results of this reciprocal prediction test further shed light on the comparative performance of the two models. The performance disparities may be due to geographical differences, variations in expert training and institutional background, as well as the variability of public scenic perceptions. However, the models' performance is therefore generally in line with existing landscape perception literature, albeit with some potential for improvement and further refinement.

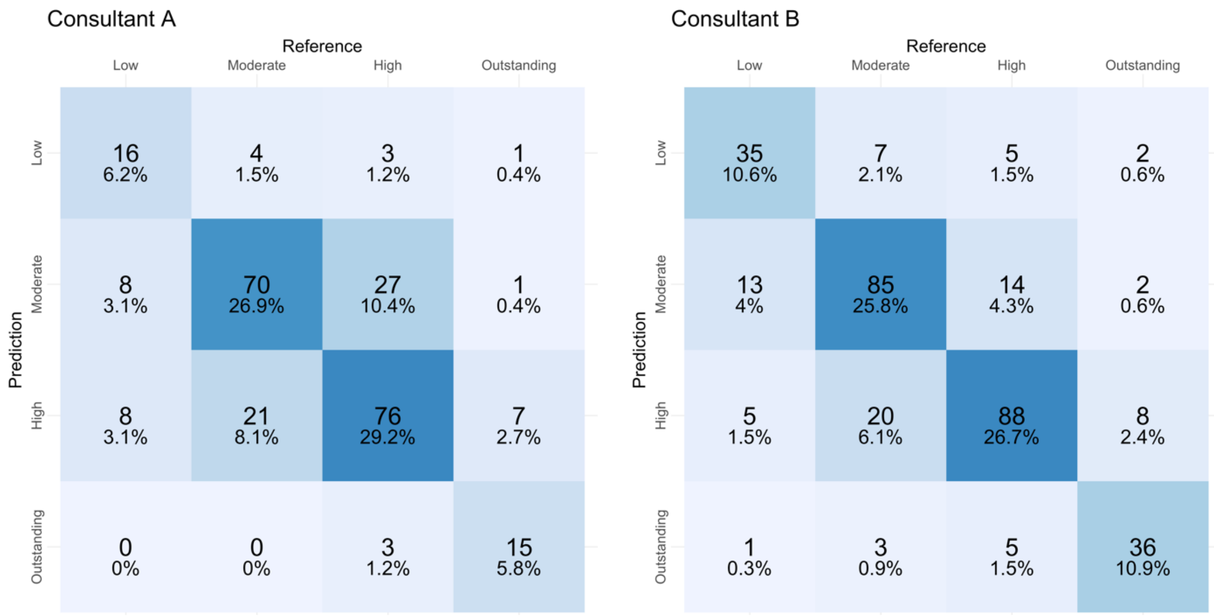


Figure 5.2 The confusion matrices for Consultant A (left) and B (right) with the number and proportion of correspondences, and shading reflecting these.

5.4.3 Relative importance of variables

The relative importance of the variables was assessed through the gain metric and provides some understanding of the modelled processes. The results are shown in Figure 5.3. For both models, the median of wildness was the most significant feature, followed by the mean of scenicness, with the IQR of wildness and the entropy of scenicness in a different order in the two models. Thus, there is considerable commonality in the models but with much less variation in the importance of the 4 predictor variables in the Consultant A model. Potentially these reflect the differences in the types of area assessed and classified by the two groups as described in Table 5.2 as well as potential differences in their landscape assessments.

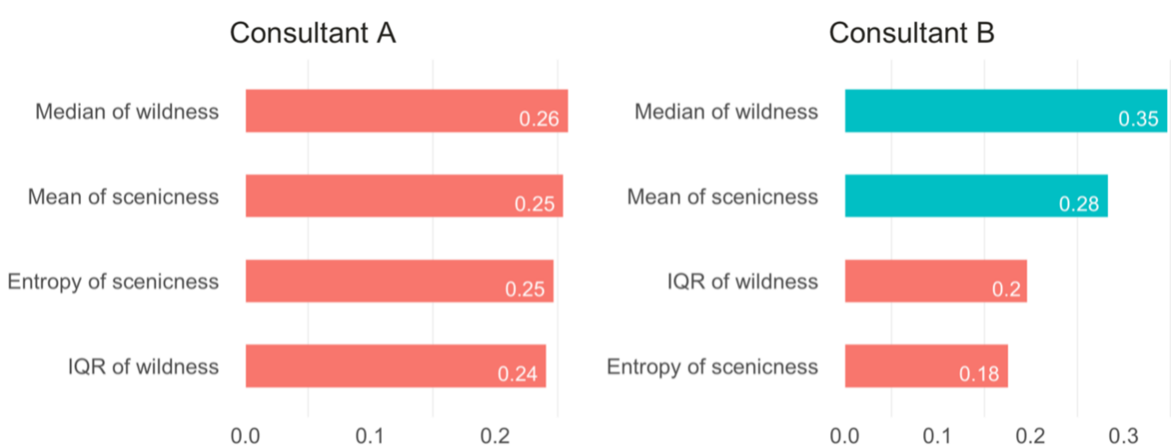


Figure 5.3 The variable importance metrics for each model.

5.4.4 Spatial distributions of landscape scenic significance

The two XGBoost classification models were applied to predict the landscape scenic significance for each 1 km square grid cell and the results are shown in Figure 5.4. Areas of Outstanding scenic quality were mostly concentrated in the uplands of Scotland, the North of England and Wales, regions which are closely associated with higher levels of topographic variation, greater visual scales, and landscape diversity than those are not found in flat, lowland regions. However, there are some notable local differences in the levels of aesthetic landscape quality predicted by the two models, particularly in the differentiation of High and Outstanding and landscape quality. The spatial distribution of Low scenic quality was far less pronounced in the predictions of the Consultant A model, whereas these areas are obviously associated with urban regions in the predictions of the Consultant B model.

5.4.5 Comparison of conceived maps regarding landscape aesthetics

The bivariate choropleth map in Figure 5.5 directly compares the model predictions. It uses a 2-dimensional colormap to allow differences between the two classifications to be unpicked: areas that were classified as Outstanding scenic quality in both models are shown in green and were generally located in upland areas, whereas those of Low scenic quality in both models are shaded in magenta. These were mainly located in densely populated urban areas, including the London and Birmingham metropolitan areas. There are clearly substantial similarities between the two models, especially at either end of the quality scale (Low and Outstanding), suggesting the presence of a degree of consensus on landscape aesthetic quality for these classes. There is more confusion across the Moderate and High classes, as might be expected. Additionally, very few of the existing designated areas (National Parks, AONBs see https://landscapesforlife.org.uk/about-aonbs/aonbs/overview), particularly those in a lowland context and situated in the south England, are identified as having High or Outstanding scenic quality in both models. One plausible explanation for this is that these protected areas are in close proximity to urban areas and their designation reflects a different set of priorities, for example related to access in support of recreation and relaxation. A further noteworthy observation is the large undesignated areas with Outstanding scenic value in Scotland, particularly around the southern uplands of Scotland.

Figure 5.4 Predicted scenic quality from the models trained on landscape quality assessment produced by Consultant A (left) and Consultant B (right).

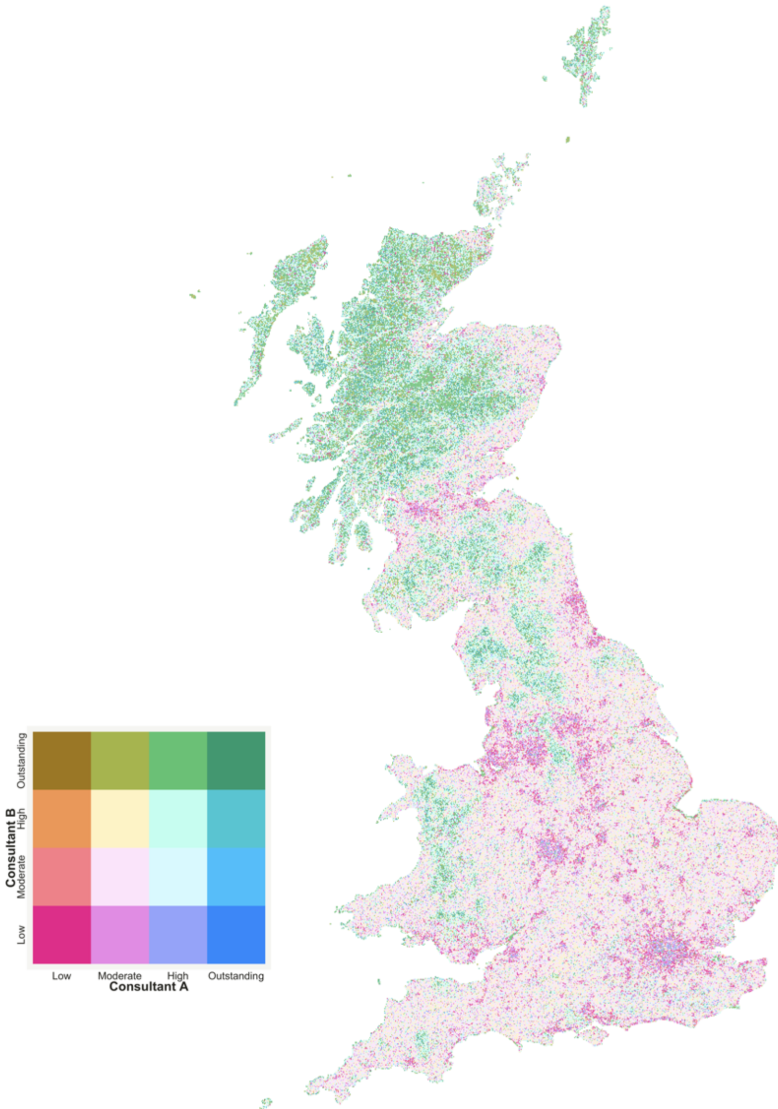


Figure 5.5 A bivariate map of predictive outcomes of scenic quality from the two models.

5.5 Discussion

5.5.1 Integration of objective and subjective assessments

This study sought to take an integrated and holistic approach to modelling landscape character, and to potentially complement current LCA practices. It combined citizen-generated perceptions of landscape scenicness with objective measures of landscape wildness to predict landscape aesthetic quality, using models trained on data – expert assessments of aesthetic quality in Wales – generated by two different groups. The models were then used to predict landscape quality over England, Scotland and Wales. The spatial distributions of the two model predictions are broadly similar, especially at the Low and Outstanding ends of the aesthetic quality spectrum, with model differences potentially reflecting differences in the areas upon which each model was trained (as shown in Figure 5.1). Despite the inherent limitations of crowdsourced data, as discussed below, a web-based preference survey of georeferenced landscape photos capturing public perceptions of landscape has been shown to mitigate the need for labour-intensive and time-consuming practices of traditional participative methods (e.g., *in-situ* questionnaires and interviews). This suggests that there are opportunities for landscape planners to adapt similar strategies by collecting and incorporating public opinions quickly and cheaply on a large scale, to augment the expert-led approaches of current assessments, such as LCAs.

The novelty of this study lies in the systematic integration of physical and socially derived measures of landscapes characteristics to construct models of expert-based valuations of landscape aesthetic quality, using state-of-art machine learning approaches. The Overall classification accuracies of 67.3% and 74.5% and the Kappa statistics of 0.50 and 0.64 are comparable to the previous studies and reflect some of the vagaries associated with human perceptions and cultural preferences (Palmer, 2004; Warnock and Griffiths, 2015; Chesnokova, Nowak and Ross S Purves, 2017). The methods employed in this study demonstrate how current paradigms in landscape character and aesthetic assessments could be augmented and used to construct consensus across experts and non-experts. They also indicate a potential approach that addressed the lack of any systematic and reproducible evaluations of landscape quality and aesthetics for England and Scotland, rather than the current exclusive focus on opaque Landscape Character Assessments. Such approaches and the resulting maps can be seen as a fusion of these three existing paradigms – the LANDMAP classification, the physical landscape measures related to wildness and the public perceptions of scenicness – and represent an integration of different types of views about the landscape. If different expert perceptions are explored (as in Figure 5.4 and 5.5), then a consensus of important regions of outstanding scenery, can be determined, across different evaluations.

5.5.2 Implication for landscape management and planning

The application of the predictive models provides crucial insights into the distribution of landscape aesthetic quality across Great Britain. Specifically, these models could help fill the void of landscape baseline in England and Scotland, regions that currently lack such landscape baselines. The resulting spatial outputs embody the collective influence of the predictor features and offer adequate detail for national-scale management and planning and could potentially supplement the existing national/regional landscape character framework.

In the broader context of landscape conservation and enhancement, the findings present several critical implications for making nuanced, context-specific strategies in landscape management. Regions of outstanding scenic quality were highlighted, including some of the existing protected landscapes such as the Cairngorms, the Lake District, and the Snowdonia National Parks. These areas, characterized by their high aesthetic scenery, necessitate concerted preservation efforts to maintain their unique landscapes. Conversely, regions exhibiting low scenic quality, most of which are densely populated urban areas, including the London and Birmingham metropolitan areas, point towards areas that could benefit from targeted landscape enhancement interventions.

Moving to a more specific geographic lens, these models could guide discussions about potential locations for Scotland's new national park – a significant concern on the Scottish government's landscape policy agenda. Following public consultation, seven potential sites—Harris, Wester Ross, Glen Affric, Ben Nevis/Glen Coe, Cheviots, Galloway, and Coastal and Marine—have been identified. The models offer authorities crucial insights into the scenic quality of these sites, a crucial consideration in the decision-making process for designation.

Particularly noteworthy is the case of Galloway, an area marked by extensive coniferous woodland, a land use which may not traditionally be perceived as aesthetically pleasing. Galloway is currently home to several distinct landscape designations, including the Galloway Forest Park, the Galloway International Dark Sky Park, and a UNESCO Biosphere Reserve. Despite these recognitions, the area's existing designation status is considered insufficient for ensuring its adequate protection. Rapid land use changes—reflected in large-scale afforestation, hydro development, shifts in agriculture, and renewable energy development—pose significant threats to Galloway's exceptional natural beauty. The spatial outputs can significantly contribute to understanding the landscape's aesthetic quality and potential future management strategies for the area, particularly in the context of its suitability for designation as a National Park.

Given the "very real concerns" about the costs associated with establishing any new national parks, the Scottish government is eager to maximise the potential of existing designations and associated economic development opportunities. However, the existing designation status of Galloway is considered insufficient for ensuring its adequate protection. In light of rapid land use changes—reflected in large-scale afforestation, hydro development, shifts in agriculture, and renewable energy development—

Galloway’s exceptional natural beauty warrants dedicated efforts and resources for its preservation.

5.5.3 Limitations of data and method

There are some limitations to this study, especially regarding the use of crowdsourced data and the associated, well recognised issues related to its data quality. The Scenic-Or-Not project employs only one Geograph photograph to represent the local landscape characteristics of each 1 km, resulting in the potential for bias around the representativeness of the scenicness rating (James F Palmer and Hoffman, 2001) and of the contributors. There may also be issues of consistency across public contributors, with individual participant’s assessments of the same scene varying (Daniel and Boster, 1976). There are also issues associated with the perceived quality in the photographic view not always matching *in situ* assessments and their multisensory ‘engagement’ (Palmer & Hoffman, 2001).

There are also potential limitations associated with the scales of analysis undertaken in this study. The machine learning models were trained on data assembled over the LANDMAP areas, and then applied over 1 km grid cells, and thus potentially affected by the differences in the model scale and the inference scale. Such issues are reflected in the modifiable areal unit problem (S. Openshaw, 1984), which posits that any statistical correlations and relationships may vary when data are brought together over different scales of analysis.

A final set of considerations relates to the areas over which the models were trained. First, none of the selection criteria used in the delineation of the LANDMAP areas were incorporated in this analysis. These were unknown to this study, but a greater understanding how individual areas were delineated and constructed could be used to refine the predictive models. Second, the model constructed over the areas classified by Consultant A were mainly mid and North Wales, where the landscape is predominantly mountainous, and has a profound effect on the observer’s experience, such as the topographic screening and impacts on views and aspect. Thus, there is uncertainty associated with extrapolating over the area with different characteristics, such as lowland and flat coastal plains. No such local context was included in the models but the inclusion of local frameworks to capture local structure amongst variables could support a more nuanced model (explanation). For instance, the shapley additive explanations (SHAP) is a fast practical method based on game theory that attempts to enhance interpretability by computing the importance values for each input feature for individual predictions (Lundberg, Erion, Chen, DeGrave, Prutkin, Nair, Katz, Himmelfarb, Bansal and S.-I. Lee, 2020). Such approaches could enhance the relatability of landscape predictions.

5.5.4 Outlook for LCA

This study has shown how more inclusive landscape assessment practices could evolve to incorporate multiple perspectives, perceptions and evaluations into the LCA processes (Wascher, 2005; Fairclough et al., 2018). It has demonstrated how information and data from diverse sources (including citizen science, user-generated content, social media, etc) can be integrated with geographical information to provide extensive and consistent landscape assessments.

It is hoped that approaches to landscape assessment and research will continue to develop around approaches that seek to include understandings of how people think about, conceptualise and value their landscapes, focusing on their individual perceptions, experiences, cognitions and behaviours (Dunkel, 2015; Koblet and Purves, 2020). These embedded micro-level views may also, in the future, contain not just visual but aural and olfactory facets of landscapes, for example. Whilst this poses potential methodological challenge to handle these diverse sources of information, requiring interdisciplinary techniques (e.g., text mining and computer vision, etc), such inputs and approaches could be intertwined with the existing LCA to improve the validity of the assessment itself and evaluations of the integrity of a landscape (Koblet and Purves, 2020). However, this more inclusive approach to LCA is unlikely to happen unless practitioners are willing to expand their thinking, rather than preserving their primacy, and are prepared to handle diverse and potentially conflicting views to their own within assessments in support of decision-making (Butler, 2016).

References

Appleton, J. 1975. *The Experience of Landscape*. London: Wiley Chichester.

Bergstra, J., Bardenet, R., Bengio, Y. and Kégl, B. 2011. Algorithms for hyper-parameter optimization. *Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011, NIPS 2011*., pp.1–9.

Bergstra, J. and Bengio, Y. 2012. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*. **13**, pp.281–305.

Bubalo, M., van Zanten, B.T. and Verburg, P.H. 2019. Crowdsourcing geo-information on landscape perceptions and preferences: A review. *Landscape and Urban Planning*. **184**, pp.101–111.

Butler, A. 2016. Dynamics of integrating landscape values in landscape character assessment: the hidden dominance of the objective outsider. *Landscape Research*. **41**(2), pp.239–252.

Butler, A. and Berglund, U. 2014. Landscape Character Assessment as an Approach to Understanding Public Interests within the European Landscape Convention. *Landscape Research*. **39**(3), pp.219–236.

Callau, A.À., Albert, M.Y.P., Rota, J.J. and Giné, D.S. 2019. Landscape characterization using photographs from crowdsourced platforms: content analysis of social media photographs. *Open Geosciences*. **11**(1), pp.558–571.

Carlson, A. 2001. Aesthetic preferences for sustainable landscapes: seeing and knowing. *Forests and Landscapes. New York: CABI Publishing*., pp.31–42.

Carver, S., Comber, A., McMorran, R. and Nutter, S. 2012. A GIS model for mapping spatial patterns and distribution of wild land in Scotland. *Landscape and Urban Planning*. **104**(3–4), pp.395–409.

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

Carver, S., Tricker, J. and Landres, P. 2013. Keeping it wild: Mapping wilderness character in the United States. *Journal of Environmental Management*. **131**, pp.239–255.

Chang Chien, Y.M., Carver, S. and Comber, A. 2020. Using geographically weighted models to explore how crowdsourced landscape perceptions relate to landscape physical characteristics. *Landscape and Urban Planning*. **203**(August), p.103904.

Chen, T. and Guestrin, C. 2016. XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. **13-17-Augu**, pp.785–794.

Chesnokova, O., Nowak, M. and Purves, R.S. 2017. A crowdsourced model of landscape preference *In*: *13th International Conference on Spatial Information Theory (COSIT 2017)*. L'Aquila, Italy, pp.1–13.

Cicchetti, D. V. and Sparrow, S.A. 1981. Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*. **86**(2), pp.127–137.

Conrad, E., F. Cassar, L., Jones, M., Eiter, S., Izaovičová, Z., Barankova, Z., Christie, M. and Fazey, I. 2011. Rhetoric and Reporting of Public Participation in Landscape Policy. *Journal of Environmental Policy & Planning*. **13**(1), pp.23–47.

Council of Europe 2000. *European Landscape Convention*. Florence.

Crofts, R.S. 1975. The Landscape Component Approach to Landscape Evaluation. *Transactions of the Institute of British Geographers*. (66), pp.124–129.

Daniel, T.C. and Boster, R.S. 1976. Measuring landscape aesthetics: The scenic beauty estimation method. *USDA Forest Service Research Paper*. **167**.

Dunkel, A. 2015. Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning*. **142**, pp.173–186.

Fairclough, G., Herlin, I.S. and Swanwick, C. 2018. *Routledge handbook of landscape character assessment: current approaches to characterisation and assessment*. Routledge.

Foody, G.M. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*. **80**(1), pp.185–201.

Geman, S., Bienenstock, E. and Doursat, R. 1992. Neural Networks and the Bias/Variance Dilemma. *Neural Computation*. **4**(1), pp.1–58.

Goodchild, M.F. 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal*. **69**(4), pp.211–221.

Haibo He and Garcia, E.A. 2009. Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*. **21**(9), pp.1263–1284.

Kaplan, R. and Kaplan, S. 1989. *The experience of nature: A psychological perspective.* New York, US: Cambridge University Press.

Koblet, O. and Purves, R.S. 2020. From online texts to Landscape Character Assessment: Collecting and analysing first-person landscape perception computationally. *Landscape and Urban Planning*. **197**(January), p.103757.

Kuhn, M. and Johnson, K. 2013. *Applied predictive modeling*. Springer.

Landis, J.R. and Koch, G.G. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics*. **33**(1), p.159.

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), pp.177–198.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N. and Lee, S.-I. 2020. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*. **2**(1), pp.56–67.

MA 2005. *Millennium ecosystem assessment*.

Močkus, J. 1975. On Bayesian methods for seeking the extremum *In*: *Optimization techniques IFIP technical conference*. Springer, pp.400–404.

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

Palmer, J.F. 2004. Using spatial metrics to predict scenic perception in a changing landscape: Dennis, Massachusetts. *Landscape and Urban Planning*. **69**(2–3), pp.201–218.

Palmer, J.F. and Hoffman, R.E. 2001. Rating reliability and representation validity in scenic landscape assessments. *Landscape and Urban Planning*. **54**(1–4), pp.149–161.

Rasmussen, C. and Williams, C. 2006. *Gaussian processes formachine learning*. Taylor & Francis Group.

Santé, I., Tubío, J.M. and Miranda, D. 2020. Public participation in defining landscape planning scenarios and landscape quality objectives (LQO): Landscape Guidelines for Galicia (NW Spain) case study. *Land Use Policy*. **94**(February), p.104559.

Scott, A. 2002. Assessing public perception of landscape: the LANDMAP experience. *Landscape Research*. **27**(3), pp.271–295.

Selman, P. and Swanwick, C. 2010. On the meaning of natural beauty in landscape legislation. *Landscape Research*. **35**(1), pp.3–26.

Seresinhe, C.I., Moat, H.S. and Preis, T. 2018. Quantifying scenic areas using crowdsourced data. *Environment and Planning B: Urban Analytics and City Science*. **45**(3), pp.567–582.

Simensen, T., Halvorsen, R. and Erikstad, L. 2018. Methods for landscape characterisation and mapping: A systematic review. *Land Use Policy*. **75**, pp.557–569.

Snoek, J., Larochelle, H. and Adams, R.P. 2012. Practical Bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems*. **4**, pp.2951–2959.

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

Terkenli, T.S., Gkoltsiou, A. and Kavroudakis, D. 2021. The Interplay of Objectivity and Subjectivity in Landscape Character Assessment: Qualitative and Quantitative Approaches and Challenges. *Land*. **10**(1), p.53.

Tveit, M., Ode, Å. and Fry, G. 2006. Key concepts in a framework for analysing visual landscape character. *Landscape Research*. **31**(3), pp.229–255.

Warnock, S. and Griffiths, G. 2015. Landscape characterisation: the living landscapes approach in the UK. *Landscape Research*. **40**(3), pp.261–278.

Wascher, D.M. 2005. *European landscape character areas: typologies, cartography and indicators for the assessment of sustainable landscapes*. Landscape Europe.

van Zanten, B.T., Van Berkel, D.B., Meentemeyer, R.K., Smith, J.W., Tieskens, K.F. and Verburg, P.H. 2016. Continental-scale quantification of landscape values using social media data. *Proceedings of the National Academy of Sciences*. **113**(46), pp.12974–12979.

Chapter 6 Discussion and conclusion

This chapter provides a critical review of the research results, data, and methods. It follows the structure of Chapter 3 – Chapter 5, with each research question being discussed in a dedicated section. The fourth section delves into the limitations of the data and methodological issues. The fifth section is centred on the potential implications of this research. The concluding remarks provide a comprehensive summary and highlight the future prospects in landscape assessments.

6.1 Linking wildness with scenicness

6.1.1 Answering research question 1:

How are crowdsourced perceptions of scenic beauty associated with expert-based measures relevant to landscape wildness quality, and at what scale? (Chapter 3)

In Chapter 3, a range of statistical methods, including correlation analysis, global and local regression modelling, was applied to examine the spatial relationships between the crowdsourced scenic ratings and the GIS-based measures of wildness quality. Upon aggregating these variables into larger, regular grids, they revealed statistically significant correlations. These findings align with earlier research (e.g., Palmer, 2004), demonstrating a high correlation between GIS-described landscape patterns and human perceptions. The findings serve to strengthen the conceptual common ground between wildness and scenicness, bridging two related fields of study.

However, it is important to acknowledge the role of the scale of analysis, in this case, the grid size, in shaping the findings. Different levels of abstraction introduced by the chosen scale can influence the relationships within the spatial data. A detailed discussions on this scale effect is provided in Section 6.4.2. For this study, a hexagon was used as the grid shape. This choice was made because a hexagonal grid structure provides distinct advantages in terms of consistent and equidistant neighbour traversal, compared to triangle or square grids (Brodsky, 2018). Further, the selection of a 5 km grid size was driven by three primary considerations: Firstly, there are certain reservations concerning the spatial accuracy and representativeness of the Scenic-Or-Not data, as further discussed in Section 6.4.1. The process of data aggregation might help to reduce biases associated with these concerns. Secondly, for a national scale case study, this grid size is deemed capable of capturing an appropriate level of spatial variability. Thirdly, dealing with large sample sizes through the MGWR framework poses significant computational complexity, an issue that cannot be overlooked. This computational burden can be mitigated by the process of data aggregation. Nonetheless, it is essential to bear in mind that the effect of data aggregation could result in difference in inferences and conclusion drawn from the study.

6.1.2 Bandwidth selection and the role of remoteness in landscape perception

Chapter 3 presents a contrast of the results from GWR and MGWR. This comparison highlights the sensitivity of bandwidth selection within the geographically weighted (GW) regression framework, which can lead to contradictory sets of local coefficient estimates produced by different local model specifications (Li et al., 2020). This is evident in the case of "remoteness" where the appropriate bandwidth for remoteness, as determined by MGWR (1,944.2 km) — a distance that spans the entire study area, is significantly larger than the 'average' bandwidth obtained in GWR (15.2 km). The incongruent scales result in the drastic difference in the local coefficient estimates. For example, the local coefficient estimates for remoteness, derived from GWR, show spatial non-stationarity, with a majority of local parameter estimates failing to reach statistical significance (see Figure 3.5). This may be attributed to the 'one-size-fits-all' bandwidth selection employed in GWR calibration. In contrast, the MGWR coefficient estimates for remoteness exhibit a nearly consistent relationship across space due to the smoothing effect of the very large bandwidth. Consequently, these estimates, though appearing much smaller in magnitude when compared to the other covariates, nevertheless attain an increased level of statistical significance (see Figure 3.6). This revelation contributes to a nuanced understanding of the stationary or global effect, albeit marginal, of remoteness on scenicness.

Meanwhile, this study emphasises the importance of bandwidth selection in GW modelling. This can be interpreted as follows: remoteness, mainly pertaining to landscape accessibility, does not significantly affect the aesthetic characteristics of a landscape. Rather, it serves as a contextual factor in landscape perception. For example, an outstanding yet remote landscape might be inaccessible for visitation, thereby inhibiting the formation of aesthetically pleasing perceptions of such a landscape. Meanwhile, previous research on landscape value modelling has shown that accessibility account for a significant amount of the observed variation in photo concentrations across various photo-sharing platforms (Boris T. van Zanten et al., 2016). These findings could potentially inform future theoretical development that non-visual factors (such as remoteness or roadlessness and accessibility) could play a marginal effect on an appraisal of visual quality. Hence, Chapter 3 make a distinct contribution by distinguishing the difference in the spatial scales of processes between the non-visual factor—remoteness—and the other visual landscape characteristics—absence, naturalness, and ruggedness.

The MGWR outputs—the bandwidth estimates, local coefficients and their significance—are clearly more sensible than those calibrated from GWR. It also advances the knowledge regarding the different levels of spatial heterogeneity exhibited by the four separate processes.

6.1.3 Superiority and applicability of MGWR

In the results, the goodness-of-fit of the MGWR model was slightly higher than that of GWR (adjusted $R^2 = 0.831$ and 0.818, respectively), both of which outperformed the MLR (adjusted $R^2 = 0.710$). Moreover, obvious decreases in AICc of MGWR over GWR (AICc = 18,313 and 18,430, respectively) were observed. Both were much lower than that of MLR (AICc = 23,001). This comparative result is in alignment with prior research (Li et al., 2020) as MGWR relaxes the unrealistic restriction of the same bandwidth for each modelled process and the intercept in GWR, allowing each relationship at the same location to have a different spatial weighting matrix (A. Stewart Fotheringham et al., 2017). This adjustment mitigates the susceptibility to local collinearity typically associated with GWR, enhancing the reliability and precision of the model. All these findings denoted that MGWR was superior to GWR in explaining the relationships between scenicness and four wildness components. Hence, this empirical study supports the advocacy that MGWR is now regarded as the default local model specification, with GWR being used only in specific circumstances (Comber et al., 2023).

The calibration of the MGWR model employs a more complex backfitting algorithm, and thus, requires more time to converge. Nonetheless, recent advancements have made it possible to calibrate even massive datasets within a manageable timeframe (Li and Fotheringham, 2020). Such advancements have paved the way for wider interdisciplinary applications of MGWR in various fields such as epidemiology and tourism (Mansour et al., 2021; Shabrina et al., 2021), but its application in the field of LP&P remains limited. Notably, the MGWR modelling is grounded on the premise of linearity and additivity, implicitly suggesting that more complex non-additive higher-order interactions are negligible or absent. However, the validity of this assumption is often unknown, which poses a potential limitation. Notwithstanding, even without explicit underlying theories, MGWR serves as an exploratory tool to grasp the spatial context of the underlying heterogeneity of perspectives from experts and laypeople, thereby facilitating collaborative planning.

6.2 Comparing scenic evaluations between experts and non-experts

6.2.1 Answering research question 2:

To what extent do people's photographic ratings for landscape scenic beauty correspond with expert-led, character-based evaluations of scenic quality? (Chapter 4)

Average of public scenic perceptions within character areas

Chapter 4 presents a comparative study to determine the extent to which public perceptual ratings for landscape scenery, derived from the Scenic-Or-Not campaign, correspond with expert-led, character-based evaluations of scenic quality, sourced from the LANDMAP visual and sensory dataset. The underlying central tendency and variability of public perceptions were initially obtained through calculating the mean and entropy of crowdsourced scenicness underneath each aspect area. The process of data aggregation is further considered and discussed in Section 6.4.2. These collective summary statistics for every Level-2 landform type were visualised through boxplots, as illustrated in Figure 4.2, to examine the assumption of homogeneous qualities made in this character-based evaluation. It comes no surprise that the most aesthetically appreciated landforms are water-related, including coastal waters, coastal, and inland water, which are contrasted to those landforms associated with development. This observation provides an evidence base to underpin much of the previous work on human landscape perception and preference (Real et al., 2000; Arriaza, J. F. Cañas-Ortega, et al., 2004). Among the four Level-1 classifications, water classes received the highest average scenic scores, followed by upland, lowland, and development classes, respectively.

Variability of public scenic perceptions within character areas

Landforms associated with upland regions, particularly 'exposed upland or plateau,' exhibited distributions sharply peaked around a higher average entropy of scenicness, indicating considerable variability. Conversely, those associated with water bodies and developed areas tend to demonstrate less variability. This discrepancy could be due to the increased diversity and complexity implied by elevated terrains where topographical and meteorological effects may be amplified (e.g., mountain grandeur or harsh weather conditions), resulting in a broad range of visual perceptions for landscape images. This observation points to a limitation in the broad classification of continuous elevation and land cover patterns into different landform categories. It is recognised that as the landform characteristics increase in dimension, for instance, an increase in relative elevation (e.g., from flatlands through hills to mountains), the land cover pattern becomes less important as an element of visual quality. Conversely, as landform dimensions decrease, the variety of ground cover becomes increasingly important for maintaining high scenic value. Nevertheless, the complex interrelation between elevation and land cover, which significantly impacts the landscape pattern in human visual perception, is not adequately accounted for by the current characterisation practices.

Comparing expert and non-expert perspectives

The crowdsourced average scenicness were regressed onto the dummy-coded covariates of the landform types, using an OLS model. The resulting coefficients were found to be statistically significant at the 0.001 level and could be ordered rationally, as detailed in Table 4.5. This result affirms the relevance of this typology for evaluating the aesthetic and perceptual value of landscapes. On the other hand, the effect of these landform types on each level of scenic quality, as evaluated by experts, were estimated by constructing four logistic regressions using the same reference category (built land). These coefficients were then used to rank the categories after being transformed into the odds of being rated at each of the four levels of scenic quality. Although these landform types are not influential to the evaluations of scenic quality, as indicated by their poor goodness-of-fits, most of the coefficients fell below the significance level of 0.05. It is important to note that the ranking order is subject to the uncertainty inherent in these estimates. Only the model for outstanding scenic quality yielded more reliable and stable coefficients than the other models, as depicted by their corresponding 95% error bars (see Figure 4.3). Despite this uncertainty, the pairwise comparisons of the rank-order between expert and public perspectives revealed a degree of rationale and similarity (see Figure 4.4). For example, both groups attributed the highest valuation to landforms associated with water among the four first-level classes.

6.2.2 Landscape appraisal: shared and divergent perspectives

While much of the literature concentrates on the development of analytical and expert techniques for landscape appraisal, a counterargument posits a lack of consensus on what constitutes aesthetic value (Daniel, 2001). Despite this contention, there is indisputable statistical significance in people’s preferences for landscapes (Real et al., 2000; Arriaza, J. F. Cañas-Ortega, et al., 2004). The findings presented in Chapter 4 further reinforce these common preferences, though differences in perspectives between the two groups do exist. Notably, a clear consensus between experts and the public is noticeable at the upper end of the scenic spectrum, where significant changes in rankings are minimal. This congruence in landscape appraisal also supports evolutionary principles and provides a foundation for formulating new hypotheses about the potential of leveraging crowdsourced scenicness. These could be used to predict the evaluative levels of scenic quality in line with LANDMAP’s qualitative scale, as demonstrated in Chapter 5. In contrast, less consensus is evident in the mid-range of the scenic spectrum, where some variation in landscape preferences between these two groups is observed. Public assessments play an integral role, complementing expert-based assessments of physical landscapes in guiding landscape management and planning. However, there is a challenge in reconciling the multi-dimensional landscape assessment used by experts in defining the character areas with the mono-dimensional response obtained from the public responding to the photographs, where the visual component clearly dominates the assessment. These discrepancies in landscape appraisal could guide future refinement of the approach. Therefore, this comparison serves as a starting point for further inquiry.

6.2.3 Methodological enhancement in landscape characterisation

LANDMAP provides a valuable hierarchical classification system for the visual and sensory characterisation of landscapes and the evaluation of their quality, utilising consistent and objective criteria. This output is represented as GIS polygons with distinct and sharp boundaries. One methodological limitation, however, is the focus of the landscape characterisation process on a bird-eye's view, rather than a ground-level perspective. Furthermore, human visual perceptions of landscapes are generally not unique, well-bounded spatial entities. Instead, their shape and extent depend on the observer's spatial position, height, and elevation. With advancements in fast intervisibility analyses, it is timely to stand back and reappraise the methodology and its outputs. There are promising possibilities in characterising visual landscapes via advanced GIS-based viewshed analyses. These techniques include vertical voxel viewsheds (Carver and Washtell, 2012), used in constructing specific wildness components, and experion viewsheds (Brabyn, 2015), which consider the viewer's localised contexts and the screening effect of nearby topographical features. These viewshed analyses offer a more nuanced and realistic understanding of how landscapes are perceived from a human perspective, thereby enhancing the relevance to human perception. This direction provides a promising avenue for future research and development in the field of LP&P.

6.3 Integrating objective and subjective landscape assessments

6.3.1 Answering research question 3

Can the integration of the subjective perceptions, objective assessments, and character-based evaluations mentioned above be used to effectively map landscape scenic quality? (Chapter 5)

The evidence from Chapter 3 and Chapter 4 somewhat provides an a priori expectation and justification for potential correlations in datasets derived from the Scenic-Or-Not, wildness mapping, and LANDMAP initiatives. Each of these initiatives captures different aspects of landscape quality using various measurement scales, which are closely associated with landscape aesthetics. The fundamental premise of this study is that these measures may exhibit a certain level of agreement or correlation, forming the basis for modelling these relationships in Chapter 5. The ultimate research aim is to integrate both expert-based, objective assessments and nonexpert-based, subjective assessments, namely wildness and scenicness, into a comprehensive nationwide landscape assessment. The expert-based judgements served as the standard criteria to determine the optimally combined configuration (weights) of wildness and scenicness assessment scales, which is then used to standardise these scales into a consistent format. Unlike the purpose of modelling in Chapter 3, spatial dependency was not considered in this modelling. To maximise predictability, the state-of-the-art machine learning (ML) algorithm, XGBoost, was employed which allowed for a thorough examination of potential nonlinear relationships and higher-order interactions among the variables. Two separate models were trained based on the assessments conducted by two different expert groups in the northern and southern parts of Wales, respectively. Both models achieved an overall classification accuracy exceeding 67%, providing empirical support for the underlying motive and evolutionary theories (refer to Section 2.2.1).

These models were then applied to predict the levels of scenic quality across Great Britain (GB), particularly benefiting England and Scotland, which lack such landscape baselines. One key caveat is that the models, built using aggregate measures, were used to predict grid units of 1 km, potentially presenting an unresolved issue further discussed in Section 6.4.2. Nonetheless, the spatial outputs from this extrapolation essentially reflect the ensemble of the predictor features, which can be regarded as a rescaled fusion of wildness and scenicness measures. The spatial 'resolution' of these outputs depends on the granularity of available scenic ratings. Despite this limitation, they offer adequate detail for national-scale management and planning, potentially supplementing the existing national/regional landscape character framework (refer to Section 2.5.2). These maps could be integrated into a single bivariate map, using a colormap to highlight regions of common outstanding and low quality for landscape conservation and enhancement, as shown in Figure 5.5 This map is particularly useful in identifying special areas based on scenery and aesthetics. The outstanding areas align well with some of the existing protected landscapes, including the Cairngorms, the Lake District, and the Snowdonia National Parks. Importantly, these findings could guide discussions about potential locations for Scotland's new national park, a topic on the Scottish government's landscape policy agenda. Following public consultation, seven potential sites have been identified. These findings offer authorities key insights into each site's scenic quality, which is a crucial consideration in designation decision-making.

6.3.2 Automating characterisation process through multiple data integration

The LANDMAP project crucially establishes a systematic and detailed methodology for assessing the visual and sensory characteristics of the landscape in Wales. However, the landscape characterisation and evaluation processes can be heavily manual and costly tasks for a country, let alone the whole of the UK. It is important to acknowledge the potential for incompatibility in such assessments, as being largely relied on expert interpretations and valuations. However, the consistency of assessments cannot be assured though within the highly institutionalised framework, as different experts may arrive at different conclusions to produce a typology and hierarchy of spatial units that may not be comparable. Further, the following field valuation for the prior-defined spatial typology of landscape may vary even among experts. Moreover, defining 'quality' landscape through such methodological approach may not adequately represent public opinions (Butler, 2016). Altogether, these considerations might explain why landscape planners are sometimes reluctant to tackle the visual and perceptual aspects of landscapes, despite the prevalence of character-based approaches in the UK's landscape assessments.

Previous research has demonstrated the effective use of CGI data from social media platforms (i.e., Panoramio, Flickr, and Instagram) in modelling landscape values (Boris T. van Zanten et al., 2016). However, the incorporation of perceptual ratings from people, like the crowdsourced scenicness highlighted in this thesis, is still uncommon in contemporary landscape assessment practices (Medeiros et al., 2021). The Scenic-Or-Not campaign offers an otherwise bottom-up approach to crowdsourcing public perceptions of scenic beauty with a basic level of engagement. However, this approach is not without potential biases, which will be further discussed in Section 6.4.1. A significant challenge with using CGI is data scarcity, often resulting from issues of accessibility. Certain remote areas may be difficult to access, resulting in limited or no data availability. Efficient indicators, such as the wildness index used herein, are particularly needed to fill this gap, and they should rely on open and regularly updated data sources. The methodology proposed in Chapter 5 highlights the analytical and inferential benefits of integrating ground-level views with bird's-eye views of the landscape for empirical modelling. This indicates a potential opportunity for automating and standardising the land characterisation process of scenic quality. However, further research is needed in this direction due to the lack of a shared theoretical basis, a common operational framework, and a harmonised data collection approach.

6.3.3 Pros and cons of machine learning

To integrate both objective and subjective assessments from expert and the public respectively, Chapter 5 presents the advantages of employing a machine learning (ML) predictive model, XGBoost. Unlike traditional statistical modelling, ML models often do not require certain assumptions such as normality, independence, or linearity. XGBoost is highly flexible and efficient in capturing complex non-linear relationships among variables while reducing the need for extensive data pre-processing tasks such as imputation and normalisation. This is particularly useful when integrating CGI data with missing values into the model parameterisation process. As shown in Chapter 5, each trained model provides a prediction accuracy comparable to sophisticated statistical methods used in previous landscape perception research (Boris T. van Zanten et al., 2016).

However, XGBoost models tend to overfit the training data if the model complexity is too high or if the hyperparameters are not properly tuned. In this case, model complexity is not an issue as the XGBoost models involves two predictor features. Bayesian optimization provides a computationally efficient and effective solution for the latter issue, as detailed in Section 5.3.3. It is also worth noting that ML/AI techniques cannot infer statistical significance, which limits their applicability for hypothesis testing. Furthermore, their predictions have often been characterised as lack of transparency and interpretability. Therefore, understanding how these 'black box' models work is vital for their practical implementation in decision-making processes including land use planning. The gain metric, illustrated in Figure 5.3, offers an overall estimate of each feature's contribution to the XGBoost model's predictions, but fails to capture the specific influence each predictor has on a given prediction. Moreover, unlike the geographically weighted (GW) modelling, spatial autocorrelation cannot be directly embedded within ML/AI models and their predictions. Recent studies have used coordinates and their derivatives as proxies to partially account for spatial dependency effects (Li, 2019; Li, 2022).

Future endeavours can, thus, focus on integrating locational covariates to account for spatial dependency and continuity, as well as on analysing the importance of each feature in every prediction using local interpretative frameworks such as shapley additive explanations (SHAP) (Lundberg, Erion, Chen, DeGrave, Prutkin, Nair, Katz, Himmelfarb, Bansal and S.I. Lee, 2020). There remain opportunities to leverage XGBoost and some local interpretative frameworks to facilitate more nuanced spatial-context investigations.

6.4 Data limitations and methodological issues

6.4.1 Data quality issues

Despite the large volume and bottom-up nature of CGI can supplement authoritative information, potential issues related to its representativeness and accuracy should not be ignored. Considering that the images used in Scenic-Or-Not originate from Geograph, it is necessary to first discuss any potential biases inherent in Geograph. Moreover, although using photographs as a proxy is a widely accepted practice in landscape perception studies, potential quality issues that might arise in Scenic-Or-Not are also worthy of discussion.

Potential biases in Geograph

The Geograph project principally encourage public contributions of geographically representative photographs for each grid square, rather than focusing on 'artistic' images. This project calls for the actual visitation of each square to document prominent geographical features. Yet, the essence in photographing landscapes often involves a process where photographers direct their 'aesthetic, socio-cultural and historical lens' towards the landscapes, expressing their complex experiences. This entails a process of interpretation mediated by perceived meanings and perhaps emotional responses to the landscape. Further, the current distribution of contributed photographs is influenced by a combination of biophysical and socio-ecological factors, such as accessibility and population densities. While any submission of inappropriate images would be rejected by a panel of moderators, it remains inevitable that contributors can introduce a degree of personal biases through subjective selection and framing of landscape scenes.

Moreover, the current guidances stipulate that a genuine 'Geograph' image should be taken within the same grid square as the subject, providing a sufficient context of the surroundings. However, when in mountainous areas, the availability of vantage points can alter the scale and quantity of the landscape visible in both horizontal and vertical directions, thus changing people's perception of the landscape (Bell, 2019). Coastal regions also offer wider viewing distance with a unique character arising from the conjunction of land and water, thereby enhancing perceived aesthetic qualities. Photographs taken from these areas likely present 'cross-grid' views, capturing distant subjects or features beyond the one square kilometre extent. Viewing distance can greatly impact the perceived visual scale (Bell, 2019). However, such far-reaching photos are generally discouraged as the rules encourage visiting all possible land squares. This could lead to an underrepresentation of samples with viewsheds larger than 1 km, potentially affecting the scenic appreciation associated with these grid squares.

Potential biases in Scenic-Or-Not

Scenic-Or-Not participants, unlike Geograph contributors who generally visit landscapes in person, simply rate photographs presented randomly, without knowing their locations. Essentially, these online scenic ratings are surface-level reactions to visual stimuli, with the primary focus on immediate visual perception and no involvement of any conceptual judgement. Furthermore, the sequence in which these photographs are presented may direct attention and influence scenic evaluations (Bruce Hull and Revell, 1989). This basic cognitive process does not take into account historical and ecological ideologies, values, or perspectives. Consequently, images that lack of an effective pictorial composition, or that do not evoke excitement or amusement, are often perceived as less scenic. The reliability and validity of these photographic surrogates for 'real-world' experiences heavily relies on the representativeness of selected photographs (Bubalo et al., 2019), as each photo samples a fixed, limited view rather than the entire landscape of a given location. Additionally, general concerns persist that photographs would not be able to simulate the multi-sensory stimuli of an in-situ landscape experience (James F. Palmer and Hoffman, 2001).

In the Scenic-Or-Not campaign, most squares include only a single Geograph photograph, with unknown selection criteria. Notably, a case of non-random selection was observed in the Lake District, where nearly 850 out of the 2,300 chosen photographs were taken by the same Geograph user (Chesnokova, Nowak and Ross S. Purves, 2017). This could potentially introduce personal biases towards landscapes, thereby influencing the rating outcomes. Considering that a 1 km by 1 km area seldom consist of just one type of landscape, it might be more representative to use multiple pictures to capture the variety of the landscape. Hence, questions about the representativeness of the data raise when applied at larger scales. Despite this, the granularity of these perceptual ratings appears sufficient for national level applications, suggesting that the scope of application might influence the effectiveness of these ratings as suitable input for landscape assessment.

Notably, the feature in Scenic-Or-Not that allows rating without login or registration has both potential strength and weakness: strength because it may help in increasing motivation to contribute; weakness because it obscures the demographic information of participants, hindering further investigation of the differences between social groups. Consequently, it becomes a challenge to determine if the crowdsourced scenic ratings truly reflect the wider public's opinions, or just represent those of a small, highly engaged subset of the population. Moreover, there may be potential source of uncertainty in the ratings caused by the misunderstanding of the Likert. For instance, participants might reverse the scoring system, attributing a score of 10 to the least scenic beauty and 1 to the most attractive scene. In addition, the issue of vandalism, which is pervasive among CGI data, might contribute to such uncertainty (Degrossi et al., 2018).

6.4.2 Change of support problem (COSP) and modifiable area unit problem (MAUP)

This thesis aimed to integrate three spatial datasets, collected at various types and scales: Scenic-Or-Not (point data), Wildness (raster data), and LANDMAP (areal data). The analyses conducted in this thesis inevitably led to concern about common issues discussed in geostatistics, specifically the change-of-support problem (COSP). The term "support" refers to the geometric properties – size, shape, and spatial orientation – of the regions from which the measurements are taken (Olea, 1991). The COSPs often arise from spatial misalignment, which occurs when observations are collected at one scale, but inferences are required at another one. Changing the support of a variable, usually through averaging or aggregation, yields a new variable with distinct spatial and statistical properties while maintaining a relationship with the original. Concurrently, the modifiable areal unit problem (MAUP) represents a relevant issue highlighted in spatial analysis and geographical literature (Openshaw and Taylor, 1979). Essentially, the MAUP is an area-to-area COSP, which is concerned with two interrelated problems: the scale/grouping effect or zoning/aggregation effect. Therefore, the results of spatial analyses are scale-dependent and sensitive to the defined spatial units (Gotway and Young, 2002). Hence, attention must be paid to both the aggregation and zoning aspects of the MAUP when working with spatial data.

In Chapter 3, the results from GW modelling were influenced by two intertwined scale effects: the scale of sampling block and the kernel bandwidth used for weight determination (S. Su et al., 2012). The analyses were conducted at a higher level of abstraction through aggregating variables, using a series of 5 km hexagons. This aggregation process led to noticeable improvements in both the correlations and the coefficients of determinations, compared to those obtained from point-level measurements as a reduction in heterogeneity at the individual level due to aggregation. As such, it becomes critical to contextualise these results in conjunction with the scale of the blocks used. For example, the optimally calibrated bandwidth of 5.7 km for the intercept term from MGWR closely aligns with the size of the hexagons. This implies a potential variability of the local intercept almost on a per-hexagon basis, with a slight degree of smoothing, which might prevent overfitting during the analysis. Notably, the local coefficients of MGWR for the intercept, as depicted in Figure 3.6, show close resemblance to the dependent variable (see Figure 3.1). This may suggest an over-reliance on the intercept term for accounting for spatial variability. Therefore, the MGWR outcomes for the other predictors is primarily to evaluate the local deviance rather than the magnitude of the dependent variable. While the individual bandwidths of MGWR do provide valuable insights into the scales of relationships between response and predictor variables, it remains uncertain whether the interpretation of these bandwidths should be associated with human-scale, particularly visual perception distances. In contrast to the global process of remoteness discussed earlier, the other three covariates related to visual landscape characteristics operate at scales ranging from local to regional. Specifically, the local parameters for naturalness vary at a broader spatial scale (118.6 km) than the other two local parameters. The optimal bandwidth for the relationship between scenicness and absence (32.9 km) is smaller than that between scenicness and ruggedness (48.7 km), indicating a more localised process for the former. However, the meaning of these bandwidths is difficult or impossible to ascertain. A systematic understanding of the impact of sampling block scale on the parameter-specific bandwidth derived from MGWR is yet to be established. This highlights the need for more comprehensive investigations and detailed discussions on scale-related issues in future research. Thus, further studies could aim to elucidate the intertwined scale effects within the MGWR framework.

In Chapter 4, different spatial supports and measurement are noted in the assessments conducted by non-experts and experts. The crowdsourced scenic ratings were collected through point-based observations. In contrast, the authoritative assessments were primarily characterised as polygonal areas with qualitative evaluations. Before comparing these two perspectives, resolving this mismatch in spatial support becomes essential, which may give rise to a point-to-area COSP. When analysing the central tendency and variability in public opinions, the applied metrics include the mean and entropy. These were calculated from aggregating responses across all images within a given aspect area. Following the logic of character-based evaluation, this aggregation procedure assumed a homogeneity of landscape quality within each aspect area. However, this is a speculative and unrealistic assumption as this aggregation failed to distinguish between individual disagreement and heterogeneity of landscape quality across different photos. Furthermore, larger aspect areas would naturally encompass more photos, given the spatially even distribution of photo sampling from Scenic-Or-Not. The current calculation for the average and entropy of scenicness do not consider the different number of photos for each aspect area, nor the varying number of votes received per photo. Both of these factors could influence the areal estimates of central tendency and variability in scenicness. While most CGI data is point-referenced observations, contemporary LCA methodologies are specific to area-based characterisation and evaluation. The challenge of relating these two distinct types of spatial data in a way that allows for valid inference remains an unsolved problem.

In Chapter 5, measures of scenicness and wildness were overlaid with the visual and sensory character areas. Summary statistics, particularly measures of centre and variation, were considered as predictor features for each of these character areas. These discrete areas did not align well with the spatial structure of continuous wildness indices, resulting in quite skewed distribution. Therefore, for the wildness covariates, the resistant measures, specifically median and interquartile range, could provide more representative statistics. Following the data aggregation procedure introduced in Chapter 4, the mean and entropy continue to serve as the preferred measures to summarise the centre and variability of crowdsourced scenicness. However, as these three measurements inherently differ in their spatial scales and supports, the aggregation processes inevitably introduce a point-to-area COSP. These central tendency and variability measures were subsequently used as input features for predictive modelling. The models were trained at the character-based level, and then applied over the smaller grid level (downscaling). It is important to note that using such models of scenic quality was based on the assumption that the relationship existing between the covariates and the response variable at the training level are consistent with that at the prediction level. This assumption may not always hold true, potentially making the results susceptible to area-to-area COSP or MAUP. Therefore, the limitations and confidence of such results should be informed, as much as their advantages. However, end-users of these models should keep in mind the quote from Box et al., 2005: "The most that can be expected from any model is that it can supply a useful approximation to reality: All models are wrong; some models are useful."

6.5 Possible implications

6.5.1 Implications for landscape planning and management

The implications for landscape management and planning are threefold. First, geographically weighted spatial statistics has the potential for applications in landscape management practices. The outcome of GW analyses is a set of mappable local statistics, particularly useful for strategic planning and management purposes. This allows landscape practitioners to make better informed decisions for the priority allocation of resources. For example, based on the results of the MGWR analysis, it was found that the public's perception of scenic beauty in Scotland is influenced by the degree of perceived naturalness, more so than in the other two countries within Great Britain. Indeed, Scotland is universally acclaimed for its remarkable natural beauty. Accordingly, landscape authorities or institutes could lay out more targeted and practical guidance for the national conservation interest. A further implication for landscape assessments is the potential application of wildness mapping. This methodology, reflecting aspects of the natural state, has not yet been widely used in LCA practices. In particularly, the visibility-based approaches that characterise visual landscape structure (i.e., the visual absence of man-made structures, perceived naturalness of land cover, and terrain ruggedness) are closely related to human perceptual dimension. These could provide valuable references for landscape visual character, which is foundational to any assessment and management action.

The second broad implication is that the integration of objective and subjective landscape assessments also has the potential to spawn new areas of investigation. The further understanding of the difference in perspectives between experts and non-experts can also push the current assessment practices forward, likely yielding more inclusive evaluations. While the intent of this thesis is not to provide a comprehensive assessment, it seeks to inspire interdisciplinary conversations and encourage dual communication between experts and non-experts.

Third, the individual maps of predicted scenic quality could be amalgamated into a bivariate map. This carries significant implications for the ongoing discussion concerning the establishment of a new National Park in Scotland. From the perspective of scenery and natural beauty, the southern uplands of Scotland including landscapes of Dumfries and Galloway, merit further consideration for conservation. Galloway is currently home to several distinct landscape designations, including the Galloway Forest Park, the Galloway International Dark Sky Park, and a UNESCO Biosphere Reserve. Given the "very real concerns" about the costs associated with establishing any new national parks, the Scottish government is eager to maximise the potential of existing designations and associated economic development opportunities. However, the existing designation status of Galloway is considered insufficient for ensuring its adequate protection. In light of rapid land use changes—reflected in large-scale afforestation, hydro development, shifts in agriculture, and renewable energy development—Galloway's exceptional natural beauty warrants dedicated efforts and resources for its preservation.

6.5.2 Implications for public participation in landscape perception

Level of public participation

Article 5 of the ELC stipulates the establishment of "procedures for public participation". The ELC's Explanatory Report recommends evaluating landscapes based on objective criteria first, and then comparing the findings with the views of various public groups (Council of Europe, 2000). The recommended procedures for fostering public participation include providing the public with information, consulting all representative bodies, using the media, and conducting awareness-raising campaigns at all levels. However, these procedures are somewhat top-down in nature, which is hardly compatible with the bottom-up approach advocated elsewhere and with the conception of landscape as 'an area … as perceived by people' (Council of Europe, 2000). From a practical standpoint, the investments required for these top-down procedures can be time-consuming and cost-intensive, which may have a deterrent effect on planners. While public participation is a leitmotiv of the ELC (Prieur and Durousseau, 2006), the achievable level of involvement or engagement in practice remains unclear. To this point, instances of proactive, large-scale public involvement in landscape perception have been rarely seen, largely due to concerns that it could overwhelm the public, thereby potentially impeding the overall planning process. Ideally, viable participation methods should allow individuals to unconsciously contribute to spatial planning processes without feeling overstrained. The crowdsourcing projects utilised in this thesis may yield insights into implementing the participatory stance, as detailed follow.

Gamification strategy

Gamification is a popular strategy used to enhance participation in crowdsourcing initiatives, as exemplified by platforms such as Geograph and Scenic-Or-Not. These initiatives leverage gamified data collection methods to mobilise public engagement, encouraging people to document landscapes with varying levels of cognitive participation, as detailed in Section 2.6.2. For instance, Geograph invites contributors to provide representative images and associated information for every square kilometre, a task that necessitates an interpretative level of cognitive processing. This repository has so far contained abundant landscape photographs and related descriptive texts. These resources offer a wealth of information not only for traditional visual-centric studies, but also for explorations into other sensory modalities like aural landscape perception, thus enabling a more holistic understanding of people's experiences and perceptions of landscapes (Chesnokova and Purves, 2018; Chesnokova et al., 2019). Conversely, many participants in Scenic-Or-Not are likely drawn by the pursuit of simple pleasure and entertainment, as the campaign demands relatively basic cognitive processing compared to Geograph. Hence, this photo-based survey can effectively capture the broad public perception of scenic beauty. This form of participation involves a simple appraisal of the landscape appearance and does not require deep though or extensive interpretation.

Landscape sampling and presentation strategy

A landscape can be viewed from an infinite number of positions and perspectives. When assessing landscape visual quality, it is critical to systematically sample this infinitude, though little attention has been paid to landscape sampling methods. To date, a universally optimal approach for landscape sampling has not been found. However, the Geograph project offers a systematic bottom-up data collection process, with simple guidance that seemingly helps to bridge this gap. It can assist in identifying landscape scenes used by persons who visit the landscape. Persons conducting the sample will be located in the landscape, familiar with it and hence sensitive to its nuances and meanings - the sample will reflect this. Similar practices involving public participation in landscape sampling could present an opportunity for a democratic and inclusive governance in accordance with the ELC, in which rights and responsibilities in relationship to landscape are shared.

On the other hand, the Scenic-Or-Not campaign has adopted a spatially even sampling procedure to ensure broad spatial coverage and avoid clustering of samples. However, the underlying criteria for selecting representative photographs remain undisclosed and the photos looked randomly chosen within the grid squares. Methodologically, the sampling grid size can be refined further to achieve a finer resolution than a 1 km square, allowing for more nuanced evaluation of spatial variations in public perceptions and preferences of landscapes. Far more experimental work is required to determine the effectiveness of the sampling granularity for the scope of landscape assessment. In addition, involving both insiders (local experts) and outsiders (landscape planners) is proffered in obtaining the best representative photographs. However, this approach is time-consuming and typically needs large-scale efforts to be implemented, compared to the use of scientific sampling. Moreover, future practices can also consider the use of advanced panoramic representations at any given location to truly reflect a human's normal field of vision to address the limitation of photographic surrogates discussed in Section 6.4.1. Technologies employing panoramic street-level imagery or photosphere, together with an interactive, rotatable viewer that presents the image in a natural projection, have evolved to deliver a virtual, immersive experience of visual landscapes.

6.6 Conclusions and prospects

In response to the ELC's call for public participation in the management and planning of all landscapes, this thesis proposes a pathway towards more inclusive and integrative landscape assessment methodologies. These methodologies aim to maintain a balance between expert-based, objectivist evaluations of landscapes and nonexpert-based subjectivist ones. To facilitate this integration, the thesis introduces innovative tools, demonstrating their applications in landscape perception and preference (LP&P) studies. The primary objective of this thesis is to answer three key research questions.

First, the spatial relationships between crowdsourced perceptions of scenic beauty and expert-based evaluations of landscape wildness quality were explored, using MLR, GWR, and MGWR models. The results highlighted the limitations of GWR and revealed the promise of MGWR. The latter, by allowing for variable-specific bandwidths, yielded more spatially nuanced and statistically significant results. Furthermore, a distinct difference in the spatial the spatial scales of processes between the non-visual factor of remoteness and the other visual landscape characteristics—absence, naturalness, and ruggedness—was identified by the MGWR findings. This study reinforced the advocate that MGWR should function as the default model for geographically weighted analyses.

Second, a comparison of landscape scenic evaluations from experts and non-experts was undertaken. Essentially, these evaluations are respectively based on two different viewpoints: bird's-eye and ground-level perspectives. The results indicated that water-related landforms received the highest aesthetic ratings, while those associated with development scored the lowest. Greater variability in the perceptual ratings was evident in upland-related landforms. Further statistical exploration confirmed the relevance of the LANDMAP visual and sensory typology in assessing landscape aesthetics. Moreover, when the scenic rankings in the landform types were compared, a commonality in scenic evaluations between experts and the public was found, particularly at the higher end of the spectrum.

Third, an integrative landscape assessment methodology has been proposed that combines subjective perceptions, objective assessments, and character-based evaluations for the empirical modelling of scenic quality. This methodology has resulted in models with a classification accuracy of over 67%. Techniques such as crowdsourcing and machine learning, specifically XGBoost, offer promising avenues for automating landscape characterisation, even though potential biases and data scarcity issues may arise. However, these models often suffer from overfitting and lack of interpretability. Future research could focus on enhancing model transparency by applying local interpretative frameworks. This might be coupled with the incorporation of locational information to address the lack of consideration for spatial dependency and continuity in current ML modelling practices.

This thesis highlights the potential of CGI data in landscape assessment, advocating for the further development of analytical strategies to leverage these valuable, albeit imperfect, data sources. Recent studies have delved into the diverse contents of crowdsourced data, which include geographic, textual, and imagery dimensions. These sources provide a wealth of perceptual details and contextual information about landscapes (Callau et al., 2019; Koblet and Purves, 2020). Such studies offer potential solutions to some of the challenges in current LCA practices. It is expected that future LP&P studies will increasingly investigate how individuals perceive, conceptualise, and value landscapes via these data sources. These studies will acknowledge not just the visual aspects but also other sensory dimensions. The application of interdisciplinary techniques, such as text mining and computer vision, will likely be instrumental. Such inputs and approaches, when intertwined with existing LCA, can enhance both the validity and integrity of landscape assessments.

Reference

Arriaza, M., Cañas-Ortega, J.F., Cañas-Madueño, J.A. and Ruiz-Aviles, P. 2004. Assessing the visual quality of rural landscapes. *Landscape and Urban Planning*. **69**(1).

Bell, S. 2019. *Elements of Visual Design in the Landscape*.

Box, G.E.P., Hunter, J.S. and Hunter, W.G. 2005. Statistics of Experimenters 2nd Edition. *John Wiley & sons.* **1**.

Brabyn, L. 2015. Modelling landscape experience using 'experions'. *Applied Geography*. **62**.

Brodsky, I. 2018. H3: Uber's hexagonal hierarchical spatial index. *Available from Uber Engineering website: https:///www.uber.com/blog/h3/*. [Online], p.30. [Accessed 28 January 2023]. Available from: https://www.uber.com/blog/h3/.

Bruce Hull, R.I. and Revell, G.R. 1989. *Issues in Sampling Landscapes for Visual Quality Assessments*.

Bubalo, M., van Zanten, B.T. and Verburg, P.H. 2019. Crowdsourcing geo-information on landscape perceptions and preferences: A review. *Landscape and Urban Planning*. **184**, pp.101–111.

Butler, A. 2016. Dynamics of integrating landscape values in landscape character assessment: the hidden dominance of the objective outsider. *Landscape Research*. **41**(2).

Callau, A.À., Albert, M.Y.P., Rota, J.J. and Giné, D.S. 2019. Landscape characterization using photographs from crowdsourced platforms: Content analysis of social media photographs. *Open Geosciences*. **11**(1).

Carver, S. and Washtell, J. 2012. Real-time visibility analysis and rapid viewshed calculation using a voxel-based modelling approach. *GISRUK 2012 Conference, Lancaster, UK, Apr*.

Chesnokova, O., Nowak, M. and Purves, R.S. 2017. A crowdsourced model of landscape preference *In*: *Leibniz International Proceedings in Informatics, LIPIcs*.

Chesnokova, O. and Purves, R.S. 2018. From image descriptions to perceived sounds and sources in landscape: Analyzing aural experience through text. *Applied Geography*. **93**, pp.103–111.

Chesnokova, O., Taylor, J.E., Gregory, I.N. and Purves, R.S. 2019. Hearing the silence: finding the middle ground in the spatial humanities? Extracting and comparing perceived silence and tranquillity in the English Lake District. *International Journal of Geographical Information Science*. **33**(12).

Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., Lü, Y., Murakami, D., Nakaya, T., Wang, Y. and Harris, P. 2023. A Route Map for Successful Applications of Geographically Weighted Regression. *Geographical Analysis*. **55**(1), pp.155–178.

Council of Europe 2000. *European Landscape Convention* [Online]. Florence. Available from: https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatynum=176.

Daniel, T.C. 2001. Whither scenic beauty? Visual landscape quality assessment in the 21st century. *Landscape and Urban Planning*. **54**(1–4), pp.267–281.

Degrossi, L.C., Porto de Albuquerque, J., Santos Rocha, R. dos and Zipf, A. 2018. A taxonomy of quality assessment methods for volunteered and crowdsourced geographic information. *Transactions in GIS*. **22**(2).

Fotheringham, A.S., Yang, W. and Kang, W. 2017. Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers*. **107**(6).

Gotway, C.A. and Young, L.J. 2002. Combining Incompatible Spatial Data. *Journal of the American Statistical Association*. **97**(458), pp.632–648.

Koblet, O. and Purves, R.S. 2020. Extracting Perceived Landscape Properties from Text Sources.

Li, L. 2019. Geographically weighted machine learning and downscaling for high-resolution spatiotemporal estimations of wind speed. *Remote Sensing*. **11**(11).

Li, Z. 2022. Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost. *Computers, Environment and Urban Systems*. **96**, p.101845.

Li, Z. and Fotheringham, A.S. 2020. Computational improvements to multi-scale geographically weighted regression. *International Journal of Geographical Information Science*. **34**(7), pp.1378–1397.

Li, Z., Fotheringham, A.S., Oshan, T.M. and Wolf, L.J. 2020. Measuring Bandwidth Uncertainty in Multiscale Geographically Weighted Regression Using Akaike Weights. *Annals of the American Association of Geographers*. **110**(5).

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N. and Lee, S.I. 2020. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*. **2**(1).

Mansour, S., Al Kindi, A., Al-Said, Alkhattab, Al-Said, Adham and Atkinson, P. 2021. Sociodemographic determinants of COVID-19 incidence rates in Oman: Geospatial modelling using multiscale geographically weighted regression (MGWR). *Sustainable Cities and Society*. **65**.

Medeiros, A., Fernandes, C., Gonçalves, J.F. and Farinha-Marques, P. 2021. Research trends on integrative landscape assessment using indicators – A systematic review. *Ecological Indicators*. **129**.

Olea, R.A. 1991. *Geostatistical Glossary and Multilingual Dictionary* [Online]. Oxford University Press. Available from: https://books.google.com.tw/books?id=SaM-800mhYAC.

Openshaw, S. and Taylor, P.J. 1979. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. *Statistical Applications in the Spatial Sciences*., pp.127–144.

Palmer, J.F. 2004. Using spatial metrics to predict scenic perception in a changing landscape: Dennis, Massachusetts. *Landscape and Urban Planning*. **69**(2–3), pp.201–218.

Palmer, J.F. and Hoffman, R.E. 2001. Rating reliability and representation validity in scenic landscape assessments. *Landscape and Urban Planning*. **54**(1–4).

Prieur, M. and Durousseau, S. 2006. Landscape and public participation *In*: *Landscape and sustainable development: challenges of the European Landscape Convention*. Council of Europe Publishing Strasbourg, pp.165–207.

Real, E., Arce, C. and Sabucedo, J.M. 2000. Classification of landscapes using quantitative and categorical data, and prediction of their scenic beauty in north-western Spain. *Journal of Environmental Psychology*. **20**(4).

Shabrina, Z., Buyuklieva, B. and Ng, M.K.M. 2021. Short-Term Rental Platform in the Urban Tourism Context: A Geographically Weighted Regression (GWR) and a Multiscale GWR (MGWR) Approaches. *Geographical Analysis*. **53**(4).

Su, S., Xiao, R. and Zhang, Y. 2012. Multi-scale analysis of spatially varying relationships between agricultural landscape patterns and urbanization using geographically weighted regression. *Applied Geography*. **32**(2), pp.360–375.

van Zanten, B.T., Van Berkel, D.B., Meentemeyer, R.K., Smith, J.W., Tieskens, K.F. and Verburg, P.H. 2016. Continental-scale quantification of landscape values using social media data. *Proceedings of the National Academy of Sciences*. **113**(46), pp.12974–12979.

Appendix A – Supplementary notes for Paper I

A.1 Interpretation of intercept variability

In this appendix, the importance of considering the variability in MGWR coefficient estimates for the intercept term is emphasised. Such variability holds critical implications for understanding the estimates for the other covariates. The optimally calibrated bandwidth for the intercept term is found to be 5.7 km, which closely aligns with the size of the hexagonal units of analysis. This suggests that there might be a variability in the local intercept on a nearly hexagon-by-hexagon basis. However, a degree of smoothing is introduced by the bandwidth being slightly larger than one hexagon. This is likely to prevent overfitting during the analysis.

Turning to Figure 3.6, it is observed that the local MGWR coefficients for the intercept bear a striking similarity to the dependent variable, as shown in Figure 3.1. This resemblance may point to a potential over-reliance on the intercept term when it comes to accounting for the spatial variability of the dependent variable. As such, the interpretation of the MGWR outcomes for the other predictors is to evaluate local deviance, rather than to explaining the spatial variation in the dependent variable.

Appendix B – Supplementary notes for Paper II

B.1 The limitation of the used aggregation method

The current approach to centre and entropy calculation, as outlined in Chapter 4, aggregates all responses across images within each aspect area. This approach has its limitation as it fails to discern between disagreements among individuals and variations in landscape quality across different photos within the same aspect area. Consequently, this approach operates based on the implicit assumption of homogeneity in landscape quality within each aspect area—an assumption that may be speculative and potentially misleading.

Furthermore, another influential variable in this context is the number of votes each photo receives as well as the number of photos each aspect area includes. Although each photo in the Scenic-Or-Not dataset has at least 3 ratings, photos with more votes might offer more reliable measures of publicly perceived scenic beauty, being less prone to outlier influence. Furthermore, larger aspect areas would naturally contain more photos due to the spatially even distribution of photo sampling from Scenic-Or-Not. These factors could impact the estimated central tendency and variability for every aspect area, yet were not considered in the current approach.

To address these limitations, a more nuanced approach would be to incorporate the use of Bayesian statistics when estimating average ratings, thereby taking into account the number of ratings each photo receives. Such a Bayesian framework would be more sensitive to the number of ratings each photo receives as well as the quantity of photos each aspect area contain. This enabling a more robust and accurate reflection of public perceptions. Future research could also focus on developing and testing more sophisticated metrics capable of capturing the heterogeneity of public opinions in this context. This will help in overcoming the limitations identified in the current study and provide more nuanced insights into the landscape quality assessment process.

B.2 Uncertainty and sensitivity in parameter ranking

The limited goodness-of-fit of the four logistic regression models is recognized, suggesting constraints in the ability of landform types to fully explain scenic quality evaluations. Despite this limitation, the derived coefficients' statistical significance at the 0.001 level and their rational ordering are noteworthy. Specific attention is drawn to the sensitivity of parameter rankings, particularly in the 'moderate' and 'high' categories – where R^2 values are indeed low (0.082 and 0.144).

It is admitted that the inherent uncertainty in some coefficient estimates, depicted by their corresponding error bars, necessitates caution in using these estimates for ranking purposes. In light of this and given that bootstrapping was not used in this study, it's recognized that concluding firm differences in the aesthetic valuation of landscapes between experts and non-experts may be challenging.

Despite the uncertainty, the existence of a degree of coherence between expert and non-expert perspectives is encouraging. Future research can potentially benefit from techniques such as bootstrapping to better quantify uncertainty linked to each predictor's ranking. Such an approach could contribute to enhancing the robustness of the methodology, and in turn, the validity of the conclusions drawn.

B.3 Exploring public perceptions in designated areas

Considering the potential implications for landscape policy and management, particularly in Wales, future research could further explore the realm of public perceptions of scenic beauty within designated areas such as National Parks (NP) and Area of Outstanding Natural Beauty (AONB). This research could employ a methodology analogous to that used in Chapter 4, by aggregating public perception data from the Scenic-Or-Not dataset within these designated area boundaries. Such a study would allow an evaluation of the influence of designation status, coded as dummy variables, on public perceptions of landscape aesthetics. While this line of inquiry extends beyond the central focus of the current study—a comparative analysis of expert and non-expert perspectives on scenic quality—it undoubtedly holds potential. This proposed research could shed light on whether the legal protection and status of landscapes, such as those offered by National Park designation, can influence public evaluations of scenic quality.

B.4 Detailed landscape evaluation criteria

The table below is an expanded version of Table 4.2 specifying the criteria for evaluating landscape quality across four distinct levels.

Evaluation Score	Definition of Importance	Description
Outstanding	International or national	A landscape offering many scenes of a picturesque quality throughout the area, which are aesthetically pleasing in composition. The area is iconic for these nationally and internationally.
High	Regional and county	A landscape with some scenes of a picturesque quality, which are aesthetically pleasing in composition. The area is notable for these regionally.
Moderate	Local	Landscapes with a few scenes of a picturesque quality, which are aesthetically pleasing in composition. These areas are notable locally for these.
Low	Little or no importance	Landscapes with very limited aesthetically pleasing scenes.

B.5 Geograph photos of each landscape typology

A random selection of three Geograph photos for each LANDMAP visual and sensory Level-2 typologies, as utilised in Chapter 4, is now provided as below. These images demonstrate the distinct characteristics and diversity inherent within each typology, giving a more intuitive understanding of each landform type.

Coastal waters



Coastal



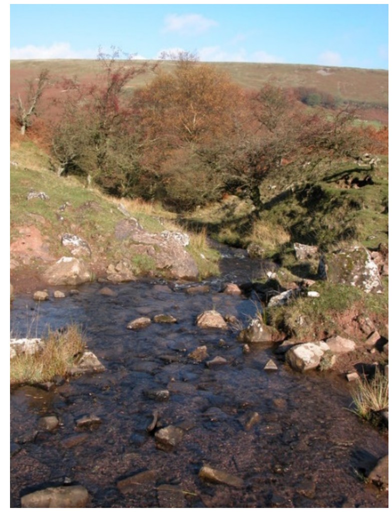
Inland water



Exposed upland or plateau

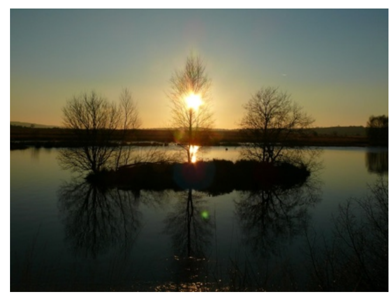


Upland valleys



Lowland valleys

Flat lowland or
levels



Hills, lower plateau,
and scarp slopes

Rolling lowland



Developed unbuilt land



Built land (reference)



Photographers of scenic images from left to right, top to bottom: Nigel Homer, Mick Lobb, Jonathan Wilkins, Humphrey Bolton, Colin Park, Robin Drayton, Shaun Butler, Nigel Brown, Row17, Bill Rowley, Alan Bowring, Philip Halling, Nigel Davies, Graham Cole, Philip Halling, Andrew Jones, Dot Potter, Dylan Moore, David Stowell, Colin Bell, Ian Medcalf, Graham Horn, John Lord, Eirian Evans, Penri Williams, Dylan Moore, Dot Potter, Philip Halling, Paul Roberts, Robin Drayton, Natasha Ceridwen de Chroustchoff, Natasha Ceridwen de Chroustchoff, Robert Cuthill.