Acta Universitatis Agriculturae Sueciae

Doctoral Thesis No. 2024:30
Faculty of Forest Sciences

The nature of human habitats

Revealing outdoor recreation preferences through landscape utilization

Carl lehto

The nature of human habitats

Revealing outdoor recreation preferences through landscape utilization

Carl Lehto
Faculty of Forest Sciences
Department of Ecology
Uppsala

SWEDISH UNIVERSITY
OF AGRICULTURAL
SCIENCES

DOCTORAL THESIS
Uppsala 2024

Acta Universitatis Agriculturae Sueciae
2024:30

Cover: A photo of a typical Swedish waterfall, with typical Swedish recreationists. (Photographer: Carl Lehto)

ISSN 1652-6880
ISBN (print version) 978-91-8046-324-9
ISBN (electronic version) 978-91-8046-325-6
https://doi.org/10.54612/a.5nfb4mggh5

Swedish University of Agricultural Sciences, Department of Ecology, Uppsala, Sweden

Print: SLU Grafisk service, Uppsala 2024

The nature of human habitats: revealing outdoor recreation preferences through landscape utilization

Abstract

A reduction of accessible green space has deteriorated peoples' opportunities for recreation. To be able to assess landscapes' potential for recreation, indices and frameworks have been developed. These have mostly relied on expert knowledge rather than analyses of empirical data. Empirical data could be achieved by studying actual landscape usage by recreationists via Public Participatory GIS (PPGIS), where surveys are employed to gain spatial data of peoples' recreational habits. Analysing such data is challenging, as other aspects than preference, for instance accessibility, also affect where recreation occurs.

This thesis investigates what landscape characteristics are important for recreationists, how an index of recreational potential can be created, and how PPGIS methodology can be improved to better understand recreation. It also evaluates the Perceived Sensory Dimensions framework, a proposed design tool based on how humans perceive environments. These aims are achieved through a literature review of which forest characteristics are preferable, combined with two PPGIS studies employing novel methodology to analyse the choice of location for recreation in Sweden.

The literature review resulted in a proposal for a recreation potential index for forests in Sweden, where large trees, proximity to water, and the absence of traces of forestry were identified as the most important elements. The PPGIS studies showed that the improved methodology, including the use of machine learning models and viewshed analysis, yielded accurate models. The models indicated several characteristics of particular importance for recreationists, such as proximity to water, recreational infrastructure and lack of urban noise. Finally, the evaluation of the PSD framework revealed it to have good internal validity, aligning with theoretical expectations. However, it also concluded that it is unsuitable as a tool for mapping landscapes based on their characteristics.

Keywords: PPGIS, Outdoor recreation, Landscape preference, Recreation potential, Machine learning

Friluftslivets natur: människors rörelsemönster avslöjar landskapspreferenser

Abstract

En minskning av grönområden har försämrats människors möjligheter för friluftsliv. För att kunna bedöma ett landskaps potential för friluftsliv har olika index och ramverk utvecklats. Dessa har främst förlitat sig på expertkunskap snarare än analyser av empiriska data. Empiriska data skulle kunna uppnås genom att studera hur friluftslivsutövare rör sig i landskapet, till exempel genom Public Participatory GIS (PPGIS), där enkäter används för att samla in rumsliga data om människors rörelsemönster. Att analysera sådana data är dock utmanande, då andra aspekter än preferens, till exempel tillgänglighet, också påverkar var friluftsliv äger rum.

Denna avhandling undersöker vilka landskapsegenskaper som är viktiga för människor, hur ett index på landskapspotential för friluftsliv kan skapas, och hur analysmetoder för PPGIS kan förbättras. Den utvärderar också Perceived Sensory Dimensions-ramverket, ett designverktyg baserat på hur människor uppfattar miljöer. Dessa mål uppnås genom en litteraturöversikt över vilka skogsegenskaper som människor föredrar, samt två PPGIS-studier som använder nyskapande metodik för att analysera svenskars val av område för friluftsliv.

Litteraturöversikten resulterade i ett förslag på ett index för landskapspotential för skogar i Sverige, där stora träd, närhet till vatten och frånvaron av spår från skogsbruk identifierades som de viktigaste elementen. PPGIS-studierna visade att den förbättrade metodiken, inklusive användningen av maskininlärningsmodeller och siktfältsanalys, gav starka modeller. Modellerna indikerade flera egenskaper av särskild betydelse, såsom närhet till vatten, anläggningar och anordningar för friluftsliv, och avsaknad av buller. Slutligen visade utvärderingen av PSD-ramverket att det inte är lämpligt som ett verktyg för att kartlägga landskap baserat på deras egenskaper, men gav stöd för att ramverkets har en god intern validitet.

Nyckelord: PPGIS, Friluftsliv, Landskapspreferenser, Index, Maskininlärning

Dedication

To Edith.

Contents

List of publications 9

1. Introduction, aims & outline 11

2. Background 15
   - 2.1 Recreation potential indices 15
     - 2.1.1 Indicators and weighting 16
     - 2.1.2 Validation 16
     - 2.1.3 Spatial extent and resolution 18
   - 2.2 The Perceived Sensory Dimensions framework 18
   - 2.3 Recreation preference research and PPGIS 20
   - 2.4 Research gaps 24

3. Methods 27
   - 3.1 Literature review 27
   - 3.2 PPGIS 28
     - 3.2.1 Survey 28
     - 3.2.2 Use-availability framework 31
     - 3.2.3 Defining and sampling the experienced landscapes 32
   - 3.3 Evaluating the Perceived Sensory Dimensions framework 34
   - 3.4 Statistical modelling: Boosted regression trees 35

4. Results and discussion 39
   - 4.1 Aim 1 – An indicator framework for Swedish forests 39
     - 4.1.1 The conceptual framework 39
     - 4.1.2 Literature review of forest qualities 40
   - 4.2 Aim 2 – Landscape characteristics and personal attributes........ 43
     - 4.2.1 Preference for landscape characteristics......................... 43
     - 4.2.2 Preference heterogeneity 47

4.3 Aim 3 – A Novel methodology for PPGIS studies 48
4.3.1 Boosted regression trees 48
4.3.2 Controlling for accessibility 50
4.3.3 Viewshed analysis 50
4.4 Aim 4 – Evaluating the Perceived Sensory Dimensions framework 51
4.4.1 Internal validity 52
4.4.2 Linking PSDs to landscape characteristics 53
4.4.3 Applicability of the PSD framework 54

5. Conclusions and future perspective 57

References 61

Popular science summary 73

Populärvetenskaplig sammanfattning 75

Acknowledgements 77

List of publications

This thesis is based on the work contained in the following articles, referred to by Roman numerals in the text:

I. Lehto C, Sirén A, Hedblom M, Lindhagen A, Fredman P. A framework for indicators of outdoor recreation in forests. (Submitted manuscript).

II. Lehto C, Hedblom M, Öckinger E, Ranius T (2022). Landscape usage by recreationists is shaped by availability: Insights from a national PPGIS survey in Sweden. Landscape and urban planning, vol 227. https://doi.org/10.1016/j.landurbplan.2022.104519

III. Lehto C, Hedblom M, Filyushkina A, Ranius T. Seeing through their eyes: Revealing recreationists' landscape preferences through viewshed analysis and machine learning. (Accepted for publication in *Landscape and Urban Planning*).

IV. Stoltz J, Lehto C, Hedblom M (2023). Favourite places for outdoor recreation: Weak correlations between perceived qualities and structural landscape characteristics in Swedish PPGIS study. People and Nature, 6 (1), pp 269-285. https://doi.org/10.1002/pan3.10574

Papers **II and IV** are published as open access articles under the Creative Commons CC-BY NC ND 4.0 license. Paper **III** will be published under a CC-BY 4.0 license.

The contribution of Carl Lehto to the papers included in this thesis was as follows:

I. Main author. **Paper I** is an adaptation of a report for the Swedish Forestry Agency that CL co-authored. CL had the main responsibility in this adaptation, which involved translating and reworking the Swedish report and producing additional figures. All co-authors contributed to the final version of the manuscript.

II. Main author. Data was collected by MH. CL conceived the methodology and analysed the data. CL interpreted the results together with MH, EÖ and TR. CL wrote the manuscript supported by all co-authors.

III. Main author. CL designed the survey supported by MH, AF and TR. CL conceived the methodology and analysed the data. CL interpreted the results together with MH, AF and TR. CL wrote the manuscript supported by all co-authors.

IV. Second author. Same data source as **Paper III** collected by CL. CL, JS & MH conceived the methodology. CL analysed the data, CL, JS & MH interpreted the results. CL contributed to the writing of the paper.

1. Introduction, aims & outline

Nature experiences are crucial for human well-being (Bratman et al., 2019; Hartig et al., 2014). However, outdoor recreation has been in decline in recent decades, leading to a suggested disconnect from nature: an ‘extinction of experience’ (Soga & Gaston, 2016). One reason for this decline could be the reduction and fragmentation of urban green spaces, which are the primary venues for outdoor recreation (Richards & Belcher, 2020). Although the importance of recreation as an ecosystem service is broadly recognized, its intangible nature makes its value hard to quantify monetarily, leading to its frequent oversight in urban development (Colding et al., 2020). Consequently, proposals have emerged for indicators of accessible green space, such as the proportion of people having a maximum distance of 300 m to a green space of at least 0.5 ha (WHO, 2015). However, the complexity of recreation—encompassing varied activities, demands, and preferences—suggests that merely increasing green space may not suffice (Kaplan & Kaplan, 1989). An area simply having vegetation is not enough to satisfy all recreationists; quality also plays a role (Kajosaari et al., 2024). Spending time in higher quality greenspace has been found to improve health outcomes, such as increased psychological well-being (Nguyen et al., 2021). Therefore, it is important to know what determines the quality of green spaces, and how quality can be quantified.

Several approaches have been applied to measure or classify landscapes' potential for recreation. On the smaller scale, such as for municipal planning, frameworks like the Recreation Opportunity Spectrum (Manning, 2022) or Perceived Sensory Dimensions (PSD; Stoltz & Grahn, 2021) have been employed. These frameworks mostly rely on expert knowledge to classify areas, which is resource-intensive and difficult to apply on larger scales. The need to estimate recreational potential on a larger scale has been recognized by several governmental bodies (Swedish Environmental Protection Agency, 2023; Nordic Council of Ministers, 2013; Forest Europe, 2020). To meet this need, researchers have developed spatial indices where landscape characteristics are used as indicators of high recreational values, in order to calculate indices of recreation potential (Casado-Arzuaga et al., 2014; de Vries et al., 2007; Komossa et al., 2018; Paracchini et al., 2014; Peña et al., 2015; Walz & Stein, 2018; Weyland & Laterra, 2014). The indicators used in these indices vary, but have often adopted simplified measures, such as assuming that landscapes that are more natural have higher quality. One reason for this is the lack of available mapping material for many landscape characteristics, but also a lack of knowledge about which characteristics contribute to high recreational values. Furthermore, these indices of recreation potential have seldom been validated by comparing the predictions with the landscape quality experienced by recreationists. The indices have also been developed to estimate recreation potential for an average person, an approach that has been challenged on the grounds that recreation is a multifaceted phenomenon, where different people have different preferences and needs (Komossa et al., 2019). As such, there may be a need to develop different indices for different groups of recreationists (Komossa et al., 2018). All these points raise questions about how well previously developed indices align with reality, especially when applying them across cultural borders (Gosal et al., 2021).

Indices need to have a basis in knowledge about landscape preferences, which can be achieved by various methodologies. Traditional approaches have largely focused on stated preference, where individuals express their likes or dislikes, often by evaluating photographs (Kaplan & Kaplan, 1989) or through choosing between hypothetical recreation sites with different characteristics (Pröbstl-Haider et al., 2020). However, such methods fall short in capturing the full sensory experience of being in a landscape—a picture does not convey the smell of flowers, nor birdsong and the rustling wind. Recently, the focus has shifted towards revealed preference, where actual landscape usage by recreationists is studied. This shift is to a large degree due to technological advancements, especially the widespread adoption of smartphones. This has enabled researchers to investigate landscape usage by, for example, analysing social media posts (Karasov et al., 2022; Tieskens et al., 2018; Yoshimura & Hiura, 2017), collecting GPS logs from mobile devices (Beeco et al., 2014; Byczek et al., 2018; Korpilo et al., 2017), or using cell tower triangulation (Bergroth, 2019). A methodology that has become increasingly popular is Public Participatory Geographic Information Systems (PPGIS), where citizens are asked to share their geographic information through surveys (Brown & Fagerholm, 2015).

Revealed preference studies, especially with PPGIS, have vastly improved access to spatial data on outdoor recreation. This is both a blessing and a curse: large amounts of spatial data can be daunting to handle, and the field still grapples with developing best practices for analysis, making comparisons between studies challenging (Brown & Fagerholm 2015; Hermes et al., 2018). One obstacle lies in the fact that it is not solely the characteristics of the landscape that influences why recreationists choose certain areas; accessibility also plays a crucial role (Koppen et al., 2014) and in analyses this needs to be controlled for (de Valck & Rolfe 2018). Moreover, similar to other ecosystem services, recreation is a result of complex relationships with many expected non-linear effects and high-dimensional interactions, which is difficult to handle with traditional statistical methods. Machine learning methods have been proposed as a suitable approach to handle this, especially since they are able to handle large amounts of data (Scowen et al., 2021). Lastly, a challenge for these studies is how to define the landscape that each recreationist experienced from the spatial data that was provided. The spatial data is most commonly point data, requiring extrapolation to estimate what landscape the recreationist experienced.

This thesis revolves around the question of how we can estimate which landscapes have high recreational values. It explores how landscape characteristics can be used as indicators of recreational potential, and how knowledge of landscape preferences can be gained by improving current PPGIS methodology. It also evaluates the Perceived Sensory Dimensions framework, which has been suggested as a possible tool for mapping recreational suitability.

More specifically, the thesis aims are:

I. to develop a framework of indicators of recreational values of forests, using Sweden as a case study (**Paper I**)

II. to investigate how characteristics of a landscape (e.g. land cover composition, heterogeneity, topography, recreational infrastructure, forest characteristics) influence the likelihood of use by recreationists. Furthermore, investigate to what degree this depends on attributes of the recreationist (e.g. age, gender, level of education), or the type of recreational activity (**Paper II-III**)

III. to further the field of PPGIS analysis by developing and implementing a novel analysis methodology, based on analysing choice of location for recreation on an individual level (**Paper II-III**)

IV. to evaluate the Perceived Sensory Dimensions framework, especially in regards to how applicable it is for large-scale mapping using landscape characteristics (**Paper IV**)

The geographical context is Fennoscandia, and particularly Sweden. Sweden has a long tradition of outdoor recreation being an integral part of daily life, with 50 percent of the adult population spending leisure time outdoors daily (Swedish Environmental Protection Agency, 2019b). This is facilitated by the right to public access, which permits outdoor recreation on almost all land, both publicly and privately owned. Outdoor recreation is also reflected in policy and legislation, with the Swedish Parliament in 2012 adopting ten national goals for outdoor recreation (Skr. 2012/13:51). However, progress follow-ups show that overall, the conditions for outdoor recreation have worsened, and that the goals relating to the amount of areas suitable for recreation are difficult to evaluate, mainly due to a lack of robust methodology to define such areas (Swedish Environmental Protection Agency 2019b, 2023).

This thesis is divided into 6 sections, with section 2 giving a detailed background on mapping recreational values and PPGIS studies, section 3 describing the methods I have used, section 4 discussing my results, and section 5 a summary and outlook towards future research.

2. Background

Research on the recreational attractiveness of landscapes began in the 1960's and 70's, spurred on by legislation in both the U.S. and Great Britain aimed at protecting 'scenic resources' (Zube et al., 1982). The research gained significant momentum with the advent of the Millennium Ecosystem Assessment in the early 2000s, which brought the ecosystem service concept to the forefront (Vihervaara et al., 2010). At this point, recreation was recognized as one of several cultural ecosystem services offered by landscapes, which since then has yielded an increasing volume of research (Hermes et al., 2018; Morse et al., 2022). A general goal for ecosystem service research has been to quantify how a landscape's provision of services depends on its characteristics, which is useful e.g. to predict how land use policy and management will affect its provision. When it comes to recreation, this has led to the development of several different indices that estimate recreational potential of landscapes, using various characteristics as indicators of recreational attractiveness. There have also been efforts to create frameworks more applicable on a local level, e.g. for municipal planning. In this section I review efforts to produce indices of recreational potential, and how PPGIS studies can be used to develop such indices. I also describe the Perceived Sensory Dimensions framework.

2.1 Recreation potential indices

Table 1 compares eight indices of recreational potential developed by researchers that were identified in the literature review of **Paper I** (see section 3.1). Their implementation varies on a number of parameters: the spatial extent (study area) and spatial resolution; which landscape characteristics are used as indicators; how the indicators are weighted against each other; and whether and how the indicators were validated. All these aspects are important in the applicability of the index and are discussed in the sections below.

2.1.1 Indicators and weighting

An essential step to develop an index of recreation potential is the decision of which indicators to include, and how they should be weighted relative to each other, i.e. their relative importance in influencing the output. Most studies have relied on expert knowledge or literature reviews for selecting indicators (e.g. Walz & Stein, 2018; Paracchini et al., 2014; Komossa et al., 2018), often highlighting the limited availability of map data as a constraint. Common indicators identified across studies include the concept of naturalness or the degree of anthropogenic influence (hemeroby), with higher naturalness or lower hemeroby associated with greater recreational potential. Water bodies also frequently emerge as an indicator of increased recreational potential (Table 1).

In most cases, included indicators were either assigned equal weight or grouped with similar indicators into composite measures, which were then equally weighted against other indicators (e.g. Casado-Arzuaga et al., 2014). However, a few studies have adopted data-driven approaches to weighting: de Vries et al. (2007) used survey data on preferences to parameterize a model of recreation potential in the Dutch countryside, while Weyland & Laterra (2014) employed the density of campsites as a response variable in regression models, operating under the assumption that a higher density of campsites reflect higher recreational quality.

2.1.2 Validation

Only half of the studies in Table 1 attempted to validate the accuracy of the resulting index. Two employed a PPGIS analysis, examining visitor numbers (Casado-Arzuaga et al., 2015) or visitor satisfaction (de Vries et al., 2007). Komossa et al. (2018) used map data on the presence of recreational facilities to validate the index, assuming that a higher density of recreational facilities indicated more attractive landscapes.

Table 1. A summary of studies developing indices of recreation potential found during the literature review of **Paper I** (see section 3.1).

Study	Resolution	Indicators	Study area	Weighting	Validation
Walz & Stein (2018)	5×5 km	Naturalness; topographic diversity; open space; ecotone density; riparian areas; coastlines; unfragmented open spaces; skyline disturbance	Germany	Equal weight of indicators	None
Peña et al. (2015)	Viewsheds of varying size	Naturalness; protected areas; water bodies; sites of geological interest; relief; mountains; landscape heterogeneity; landmarks	Basque Country, Spain	Equal weight of indicators	Photo-survey (n=629)
Paracchini et al. (2014)	Not explicitly stated	Naturalness; water quality and proximity to coast; protected areas	EU	Equal weight of naturalness, water, protected areas	None
Nahuelhual et al. (2013)	Not explicitly stated	Large trees; scenic beauty; roads; natural attractions; tourism services; land cover type	Ancud, Chile	Weighted by ‘experts’ and ‘ecotourists’	None
Weyland & Laterra (2014)	32×32 km	Mean temperature; thermal amplitude; ruggedness; NDVI; tree cover	Argentina	Regresses campsites against indicators	None
Komossa et al. (2018)	1×1 km	Naturalness; water; elevation; land cover; air quality; light pollution; noise; livestock; edible plants; trails; rare species; protected areas; sacred species	EU	Weighted by experts	Recreation facilities
Casado-Arzuaga et al. (2014)	2×2 m	Naturalness; land cover; protected areas; coast; recreational areas; climbing sites; cycling paths; routes of geological interest; mountain summits	Bilbao, Spain	Equal weight of infrastructure, naturalness, coast	PPGIS; Photo-survey (n=500; 64)
De Vries et al. (2007)	250×250 m	Naturalness; relief; monuments; urbanization; skyline disturbance; noise	Netherlands	Weighted by survey data	PPGIS (n=5226)

2.1.3 Spatial extent and resolution

The study area or population to which the index is applied to varied from a single municipality to the entirety of the EU (Table 1). Designing a set of indicators that remain relevant across diverse natural landscapes presents a complex challenge, further compounded by the variability due to cultural context on recreational preferences (Gosal et al., 2021). Even within a single country, there is evidence of preference heterogeneity, as seen in Weyland & Laterra (2014) who identified significant regional differences in recreational preferences across Argentina. This variability raises questions about the applicability of broad-scale approaches, like the index implemented across the entire EU by Paracchini et al. (2014). The authors acknowledge this short-coming, suggesting that for more nuanced analyses at smaller scales indicators should be weighted according to local conditions. Additionally, the availability of landscapes for recreational activities can vary significantly between countries, influenced by factors such as public access legislation.

The spatial resolution of the indices (i.e. the scale at which the indices were calculated) also varied by several orders of magnitude: from 2×2 m up to 32×32 km. The choice of spatial resolution was sometimes motivated, such as in Weyland & Laterra (2014) where the authors indicate that 32×32 km is an accessible landscape within a day's trip, or Walz & Stein (2018) arguing that 5×5 km captured a ‘municipal scale’. Most of the studies did not discuss the choice of spatial resolution at length, with e.g. Casado-Arzuaga et al. (2014) simply stating that the resolution of all data was 2×2 m, or de Vries et al. (2007) writing that 250×250 m was a ‘convenient size’. Paracchini et al. (2014) and Nahuelhual et al. (2013) did not explicitly state the spatial resolution employed. Many indicators used in the indices are scale-dependent, such as estimates of landscape heterogeneity or any indicator calculated as an average, e.g. naturalness or ruggedness.

2.2 The Perceived Sensory Dimensions framework

Recreational research has developed several frameworks mainly applied for recreational management and physical planning on smaller scales. Such localised frameworks benefit from enhanced availability of map data, detailed knowledge of resident demographics, and specialised expertise that are not available at larger scales. One such is the Perceived Sensory Dimensions (PSD) framework, which defines eight core qualities that humans value in outdoor settings: *Natural*, *Cultural*, *Cohesive*, *Diverse*, *Sheltered*, *Open*, *Serene*, and *Social* (Figure 1). These attributes aim to address diverse recreational needs, encompassing both relaxation and activation. Originally conceived in the 1990s by Berggren-Bärring & Grahn (1995) as Park Character Analysis, these qualities were derived from factor analysis of surveys on green space experiences across various Swedish population segments (Stoltz & Grahn, 2021).

Presently, PSD remains a vibrant area of research with over 60 studies worldwide employing this framework in diverse ways. It is also used in practice, predominantly as a design tool for green space by landscape architects (Stoltz & Grahn, 2021). It has been employed as a mapping tool by several Swedish municipalities, where recreational areas have been classified by experts according to which PSDs can be experienced there (Skärbäck, 2007). Several studies have assumed that PSDs are associated with certain landscape characteristics, such as land cover type, and attempted to predict PSD occurrence using these assumed associations (Björk et al., 2008; Annerstedt van den Bosch et al., 2015). However, none of the studies have validated this approach, and PSDs connection to landscape characteristics is thus a current research gap.

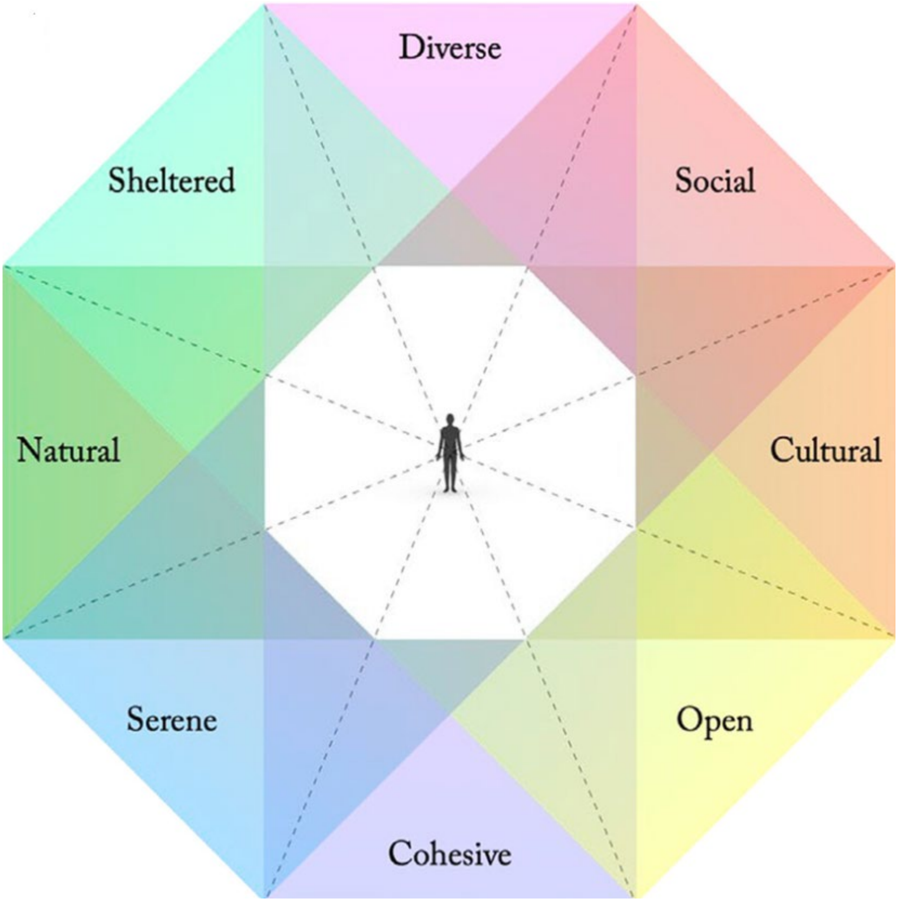


Figure 1. Schematic relations between the eight perceived sensory dimensions (PSDs). The figure suggests which perceived qualities are expected to correlate, with positive correlations for adjacent PSDs and negative between opposites (Stoltz, Lehto & Hedblom 2024).

2.3 Recreation preference research and PPGIS

To develop accurate recreation indices and frameworks, correctly identifying the underlying drivers of recreation is paramount. Research on recreational preferences can broadly be classified as either employing a stated or revealed preference approach. Stated preference studies capture respondents' preferences through for example asking them to rate photographs (Gundersen & Frivold, 2008); choose between hypothetical scenarios (Pröbstl-Haider et al., 2020); or through qualitative interviews (Scott et al., 2009). While such studies have offered valuable insights into recreationists' preference and behavior, there is always the question of how well knowledge gained in the lab translates to the real world. Revealed preference studies, where actual landscape usage is observed, have shown both agreement (Silvennoinen et al., 2022; Kienast et al., 2012) and disagreement (Boll et al., 2014; Cordingly et al., 2015) with findings from stated preference studies. This suggests that for a comprehensive understanding of preferred landscapes for recreation, both approaches are needed.

PPGIS studies of recreation are methodologically diverse, but common approaches have been to ask people to identify either their favourite place in the landscape (Scholte et al., 2018), typical places they visit (Gerstenberg et al., 2020) or the most recent places they have visited (Agimass et al., 2018). Most often point data is collected (Kienast et al., 2012; Scholte et al., 2018; Ridding et al., 2018; Baumeister et al., 2020; de Valck et al., 2016), with some studies collecting route data (Korpilo et al., 2018; Gerstenberg et al., 2020). To analyse the data, hotspot analysis is commonly employed, where areas containing large amounts of points or routes are assumed to have a higher recreational value, with landscape characteristics then being analysed within these hot spots (Scholte et al., 2018; Ridding et al., 2018; Baumeister et al., 2020; de Valck et al., 2016).

Revealed preference studies must make an assumption about which area the recreationist has experienced in order to analyse the landscape characteristics within it. In hotspot analysis, this is usually done through kernel density estimation (KDE), where points or routes are smoothed into a continuous surface. To implement KDE requires setting the bandwidth parameter, which determines how far each point or route is smoothed, and thus, the portion of the landscape assumed to have been experienced by the recreationist (Figure 2). The bandwidth in PPGIS studies of recreation has varied widely, from 50 metres (Gerstenberg et al., 2020) to 15 kilometres (Scholte et al., 2018). What bandwidth is reasonable depends on the scale that the study is operating on: Scholte et al. (2018) analysed the choice of recreation location on a much larger scale (all of the Netherlands) than Gerstenberg et al. (2020), who focused on recreation in selected forest areas. However, the impact of bandwidth choice on the analysis is rarely discussed, and ultimately subjectively chosen.

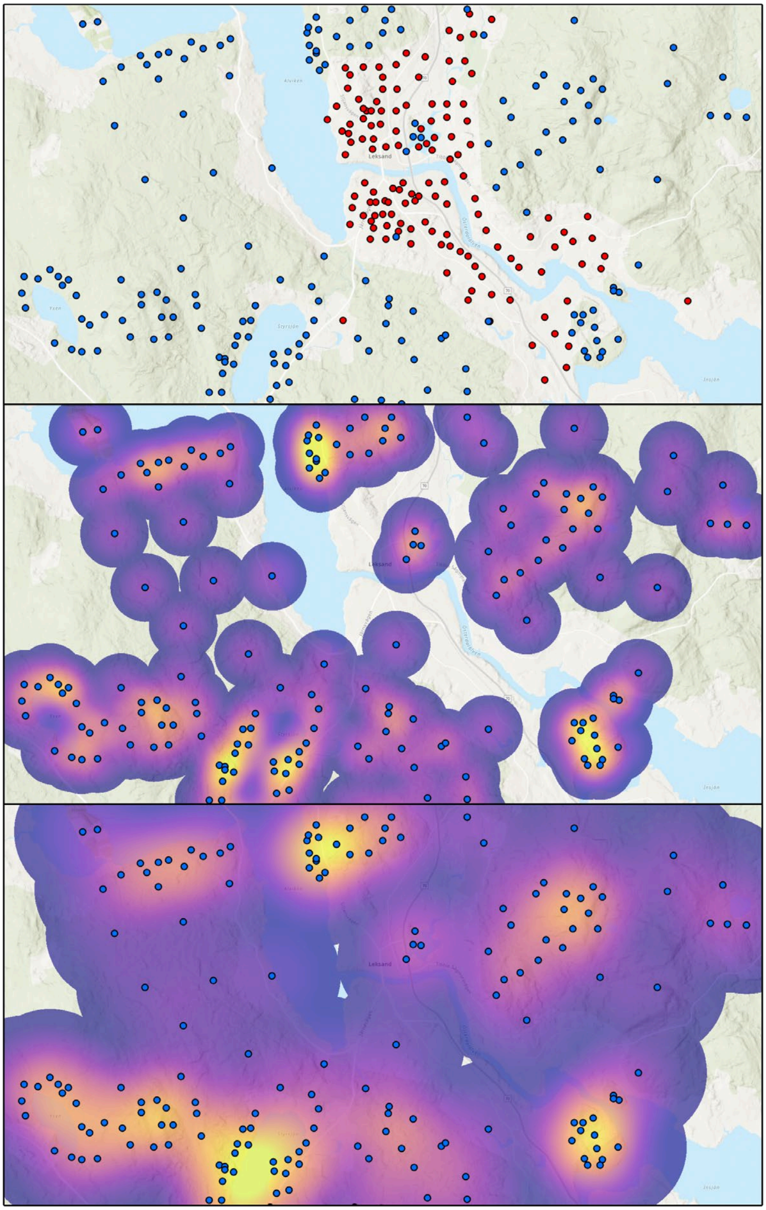


Figure 2. A hypothetical example of a hot spot analysis showing how the bandwidth parameter in a kernel density estimate analysis affects the results. The top map shows individuals (red dots) who have been surveyed for their favourite spots for recreation (blue dots). The middle and bottom image shows kernel density estimates of the visits using two different bandwidths, with a larger bandwidth (bottom image) yielding a more smoothed surface.

Analyses that do not use hotspot analysis have instead mostly sampled landscape properties in a buffer around the points (Figure 3). The question then arises about what a relevant distance is: Baumeister et al. (2020) used 50 metres, arguing that it is an estimate of the 'perceptible area', while Ridding et al. (2018) used two buffers, 500 metres to represent the local scale and 5 kilometres to represent the landscape scale.

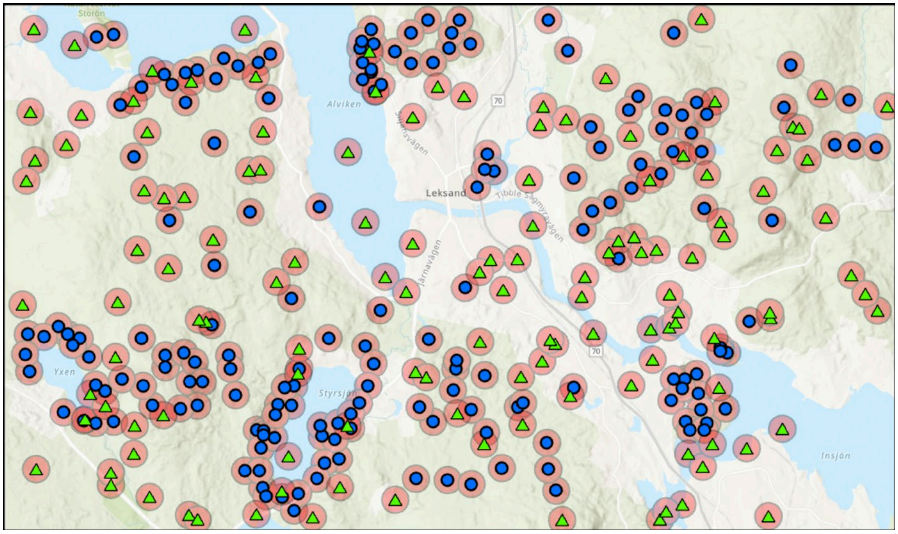


Figure 3. An alternative approach to a hot-spot analysis, using the same hypothetical data as in figure 2. Here, random points (green triangles) have been placed within the same extent as the favourite spots (blue dots). Landscape characteristics are sampled in buffers around each point, and logistic models are created that compare the characteristics of the chosen points to the random points.

When employing hot spot analysis, the response variable is the intensity of use of the landscape. This makes it more challenging to examine preference heterogeneity, e.g. to draw conclusions about how individual characteristics like gender or age influence preferred landscape qualities. In order to still draw such conclusions, a solution has been to subdivide the sample and create different models, e.g. dividing by the types of activity engaged in (Korpilo et al., 2018, de Valck et al., 2016; Gerstenberg et al., 2020), by origin of the recreationist (Scholte et al., 2018), or by age groups (Kienast et al., 2012). This approach is viable, but lowers the sample size, making it untenable to test interactions between several personal attributes at the same time. A study design focusing on the individual's choice process, where the choice of location for recreation is compared to the landscape that was accessible to the recreationist makes such analysis easier, as exemplified in Agimass et al. (2018). In the study, Danish recreationists were asked to point out their recent forest visits, followed by an analysis comparing the forests they chose with other forests they could have chosen instead based on their home location. This approach has been less common, possibly because collecting data on location of peoples' homes can be sensitive information, and also the difficulty in defining what landscape is accessible to an individual.

The point of accessibility is crucial also for many of the studies employing hot spot analysis. There is strong evidence for a landscape's attractiveness for recreation not only depending on its physical characteristics, but also depending on where it is located (de Valck & Rolfe, 2018). For recreation, accessibility has proven to be a critical factor, with people significantly more prone to using areas that are in proximity (Neuvonen et al., 2007; Hörnsten & Fredman 2000; Koppen et al., 2014). To draw accurate conclusions about people's preferences based on revealed preference studies, this factor needs to be controlled for. While many authors are aware of this, how it has been handled methodologically varies. Some studies have not considered it at all (Baumeister et al., 2020; Gerstenberg et al., 2020), while others (Kienast et al., 2012; Scholte et al., 2018; Ridding et al., 2018) include distance to settlements as a proxy for accessibility. Often, this predictor yields strong effects, providing an estimate of the magnitude of the effect of accessibility. However, it does not control for it, suggesting the recreationist's 'true' preference remains masked. To properly account for this effect, an individual level analysis is necessary, where the accessible landscape for each recreationist is estimated.

2.4 Research gaps

There are several deficiencies making the implementation of indices for estimating recreational potential difficult. One of these is the spatial resolution employed: the difference between assessing a landscape's recreation appeal on a 5×5 km versus a 2×2 m scale is substantial, yet this aspect is rarely discussed, or in some cases not even explicitly defined. The selection of indicators has frequently been arbitrary, driven more by map material availability than informed decision-making. This is further highlighted by the common practice of not weighting indicators, opting instead for equal weighting. The most glaring issue, however, is the lack of validation; many indices remain unvalidated or rely on rudimentary measures, as in Komossa et al. (2018). While de Vries (2007) presents a more ambitious validation approach using a large dataset, it also falls short as the dataset was not specifically collected for the purpose of index validation, instead repurposed from previous data. Selection of indicators and validation can be achieved through the implementation of PPGIS, but care needs to be taken with study design. Accessibility needs to be controlled for on an individual level to draw correct inferences of which landscape characteristics matter for recreationists. Performing individual level analysis also improves the ability to draw conclusions on preference heterogeneity between individuals. Finally, how to estimate what landscape the recreationist experienced from the spatial data they provided is an underdeveloped aspect, which needs further consideration. The goal of this thesis is to fill in some of these research gaps, as specified in the aims in section 1.

3. Methods

This section summarises the methods I have used in my four papers to reach the aims stated in section 2.

3.1 Literature review

In **Paper I**, we performed a literature review to determine which forest characteristics could be used as indicators to develop an index of recreation potential of Swedish forests. The review was conducted as a scoping study, where relevant literature was successively identified through a snowball methodology (Arksey & O'Malley, 2005). Initially, we surveyed scientific literature concerning recreational preference of forests that we were already acquainted with, supplemented by searches across Scopus, Google Scholar, and Web of Science databases. Our search terms encompassed combinations such as 'Recreation', 'Forest, 'Boreal', 'Temperate', 'Indicator', 'Preference' etc. We systematically expanded our search by examining literature cited by identified publications, particularly review articles, and by scrutinising databases for more recent references citing these sources. While our primary focus was on research in the Fennoscandian region, we also incorporated relevant studies from other regions as deemed appropriate. We included studies on both stated and revealed preferences.

We assessed the strength of each forest characteristic in affecting recreation potential by individually rating the relative significance from the evidence found in the literature, and then discussing our choices to arrive at a final ranking. We also assessed whether there was evidence for heterogeneity in preferences related to the characteristic, and to what degree there currently exist map data or methods to deploy them as indicators on a national level in Sweden.

3.2 PPGIS

To evaluate which landscape characteristics (combined with characteristics of the person and the activity) affect the choice of area for recreation, and to further the field of PPGIS methodology, we analysed two separate survey datasets on recreational usage of landscapes in Sweden.

3.2.1 Survey

In **Paper II**, we analysed a large spatial dataset of Swedish recreationists' latest visit to nature collected by the Swedish Environmental Protection Agency in 2014 (Figure 4). This data was collected as a part of their recurring survey on Swedes' recreational habits (Swedish Environmental Protection Agency, 2015). In the survey the respondents were asked to mark their latest visit to nature on a map and supply additional details such as the type of activity, time spent on the location and how far they travelled to get to the location.

For **Paper III-IV**, we designed and deployed an online survey for residents in the city of Umeå (Figure 5). Similarly to the previous survey, the respondents were asked to map their outdoor recreation activities. However, instead of marking their latest recreational visit with a point, the survey was divided into two parts: the first tasked the respondents with drawing their typical recreational routes within Umeå municipality, along with answering follow-up questions regarding each route, such as type of activity, mode of transportation, visit frequency, duration, season etc. The second part asked respondents to mark their favourite recreational places, which were defined as places 'holding any specific importance, such as a place of beauty or somewhere you often stop and spend time in'. For each favourite place they marked, they were asked questions related to the Perceived Sensory Dimensions framework, which was used for the analysis in **Paper IV** (section 3.3). The respondents were also asked to mark their home location, provide socio-demographic background information, and assess to what degree they saw themselves as an urban-oriented and nature-oriented person respectively. This final question was included as it has been shown to affect perception of greenspace in previous studies (Ode-Sang et al., 2016; Gunnarsson et al., 2017).

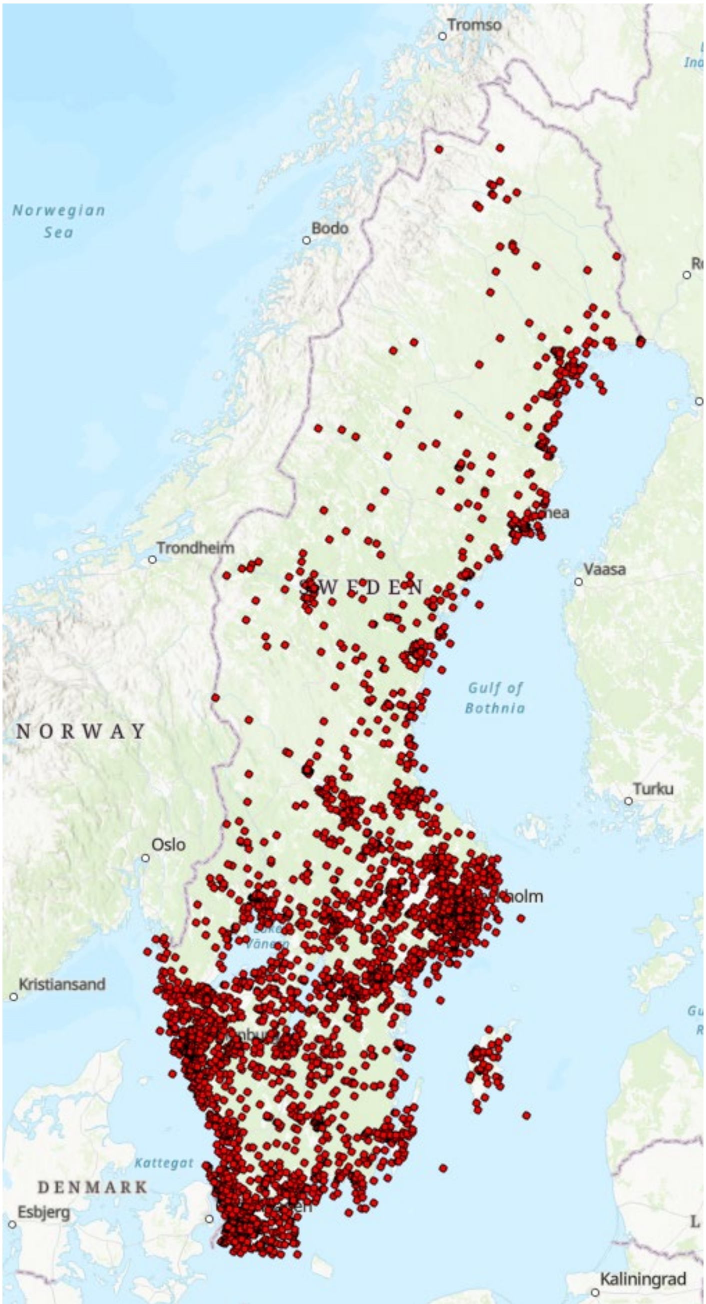


Figure 4. A visualization of the dataset used for **Paper II**. Each point represents a recreationist's latest visit to nature (n=3853).

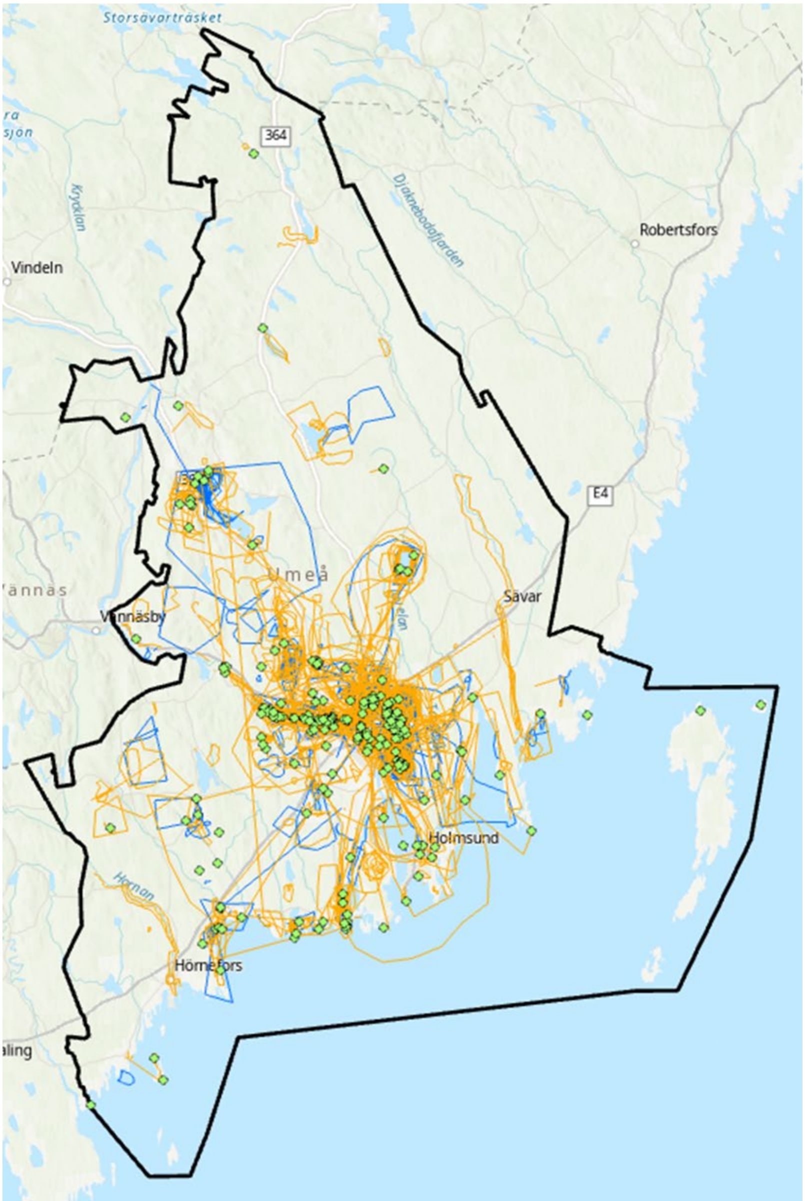


Figure 5. The dataset collected for **Paper III-IV**. Green dots are favourite places (n=275), orange lines are routes performed in summer (n=947), blue lines are in winter (n=442). The dataset originated from 358 individuals. The Umeå municipality border is outlined in black.

3.2.2 Use-availability framework

To draw conclusions on which landscape characteristics matter for recreationists, a *use-availability* framework was employed. This is a methodology commonly used in studies of animal habitat selection, where landscape characteristics at locations where the study animal has been (the use sample) are compared to randomly selected locations from the surrounding landscape (the availability sample; Northrup et al., 2013). The rationale behind employing this approach was to focus on the individual's choice process, both to properly control for the effect of accessibility (only comparing the chosen locations to what the recreationist actually had accessible to them) and to be able to draw conclusions on how individual characteristics affect choice of location. How the availability sample was created differed between **Paper II** and **Paper III**.

In **Paper II**, the use sample consisted of the points marked by the survey respondents, and the availability sample was generated by randomly placing a point within twice the distance the respondent had travelled to reach their location. Drawing the availability sample in this way was due to the lack of exact information regarding where the respondent had travelled from: had this been known the availability sample would instead have been drawn from an equidistant point from the home location.

In **Paper III**, the availability sample was created in a more intricate way. For the routes, first a spatial network analysis was performed. This analysis used path and road map data to determine which areas could have been reached in the same time it took to reach the beginning of the route from the home of the respondent, using the same mode of transportation (on foot, by bike or by car/public transportation). A random point was then placed at a point that was equally accessible as the performed route, and a copy of the performed route was placed to serve as the availability sample (Figure 6). To construct the availability sample for the favourite places, points were randomly placed around each home location within a distance equal to the distance to the favourite place. Network analysis was not applied here since information on the mode of transportation was not requested for the favourite places. Multiple random points were placed for each favourite place to test model sensitivity to the size of the availability sample.

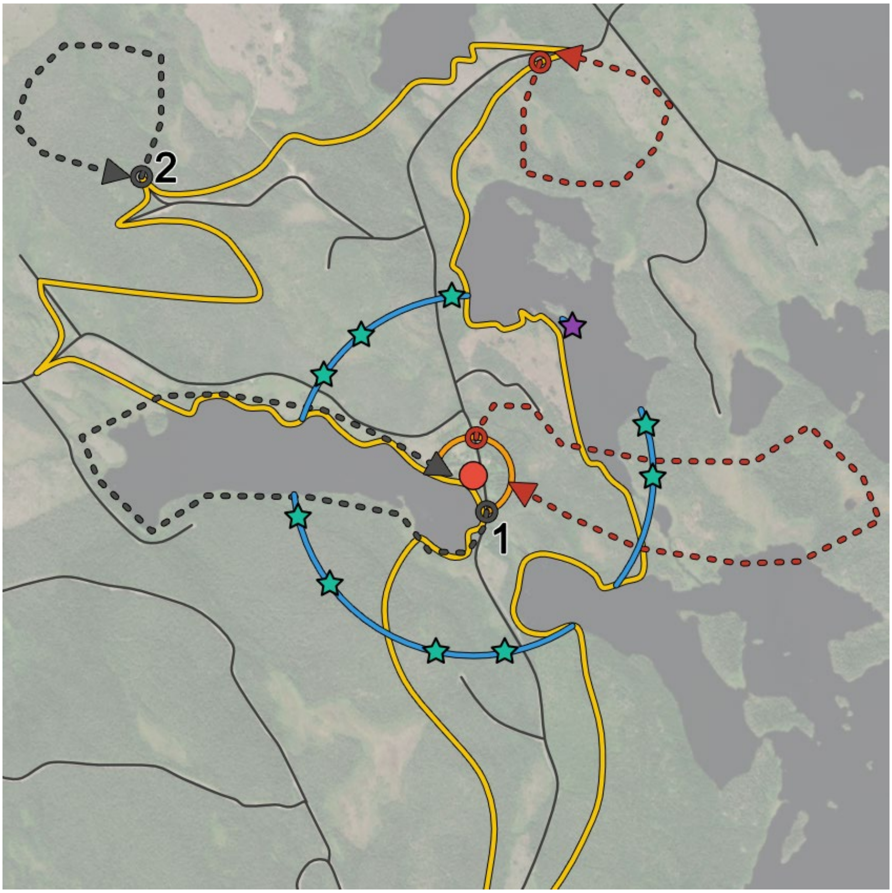


Figure 6. An example illustrating the creation of an availability sample for one respondent. The respondent identified their home location (red dot), indicated two recreational routes (gray dashed lines), and designated a favourite place (purple star). Availability samples for the routes were generated by mirroring each route's shape (red dashed lines) and placing it randomly within terrestrial areas accessible in the same travel time, considering the mode of transportation. The orange circle indicates points as accessible as the start of route 1, reached on foot, while the yellow line represents areas as accessible as the start of route 2, accessible by car. Nine random locations (teal stars) were placed at an equal distance from the home location as the distance to the favorite place (blue circle).

3.2.3 Defining and sampling the experienced landscapes

As described in section 2.3, a challenge in studies of revealed landscape preference is how to define what landscape each recreationist experienced based on the spatial data they provided.

In **Paper II,** we only had a single point marking the location of each recreationist's last outdoor recreation. These activities ranged widely, from brief dog walks to extended two-hour bike rides, as such the size of the experienced landscape was expected to vary. To accommodate this variation, we employed circular buffers, where we increased the buffer size with duration of the visit: longer visits were associated with larger buffers. We conducted a sensitivity analysis on the function that determined buffer size from visit duration by creating models with varying buffer sizes and assessing their accuracy.

In **Paper III,** we attempted to increase the realism of the analysis through a two-pronged approach. First, a 50 m buffer was created along the routes and at the favourite places, which represented a minimum estimate of what landscape the recreationist had experienced. In addition to these buffers, we also employed viewshed analysis, which is a technique to determine what areas are visible from a given vantage point, considering terrain elevation and obstructions like trees and buildings. It is a tool commonly used in physical planning e.g. to assess visual impacts of new developments. It has been previously applied in recreation research, primarily focusing on modelling the aesthetic value of landscapes or vistas using crowdsourced photographs (Karasov et al., 2020; Tenerelli et al., 2017; Yoshimura & Hiura, 2017) but also to analyse PPGIS data on recreation (Ridding et al., 2018). Here, we constructed viewsheds at the favourite places and along the routes using LiDAR data (Lantmäteriet, 2023), which provide high-resolution heightmaps of both the ground terrain and any obstacles that block vision, such as trees and buildings (Figure 7).

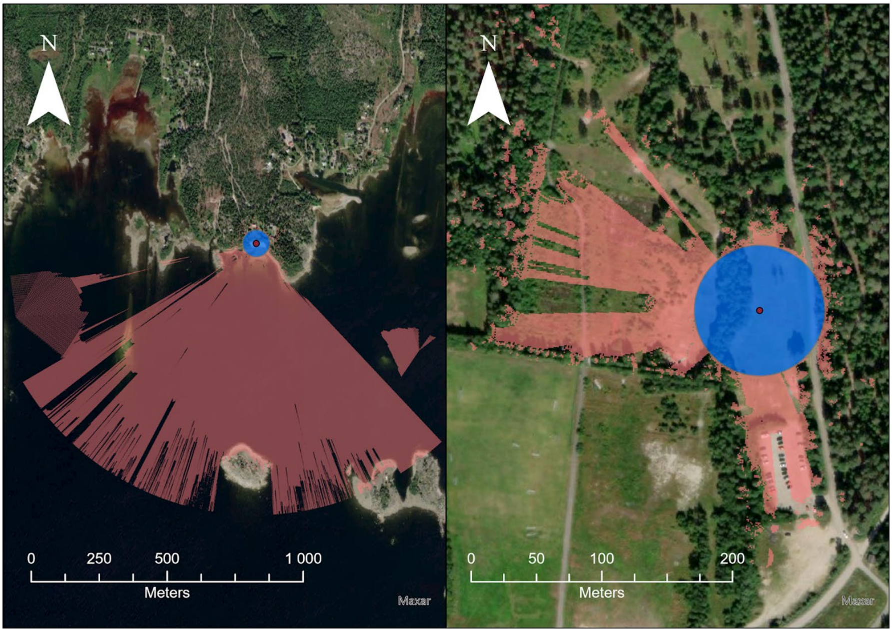


Figure 7. Example of the sampled landscape around two favourite places. The red point is the favourite place provided by the survey respondent, the blue circle is the 50 m buffer, and the red area is the estimated visible landscape when standing at the point (viewshed). Viewsheds were also estimated along the routes in the study.

Map data of various landscape characteristics that were expected to affect recreational attractiveness were then sampled in the buffers and viewsheds: land cover composition, land cover heterogeneity, presence of roads and paths, protected areas, tree height, tree species composition, and elevation. For **Paper III-IV** also noise and distance to recreational infrastructure (toilets, fire pits, and shelters) were included. All these characteristics, together with the personal attributes of the recreationists and of the recreational visit, were subsequently used as predictors to train machine learning models (section 3.4).

3.3 Evaluating the Perceived Sensory Dimensions framework

The aim of **Paper IV** was to evaluate the Perceived Sensory Dimensions framework (Stoltz & Grahn, 2021) and its connection to landscape characteristics. This was done by analysing data collected in the second

PPGIS study. When respondents were marking their favourite places, they were also asked to answer eight statements, each corresponding to one PSD (Table 2). These were based on the definitions of the PSDs described by Stoltz & Grahn (2021), and were phrased as simple one sentence statements, intended to capture the essence of each PSD. For each statement, the respondents were presented with a slider that ranged from 0 to 100, where 0 corresponded to 'Not at all' and 100 to 'Fully'.

Table 2. Statements for each PSD experienced at recreationists' favourite places in the survey, answered using sliders from 0 ('Not at all') to 100 ('Fully').

PSD	The place evokes a sense of…
Natural	… wild and untouched nature
Cultural	… being shaped by humans
Open	… openness, with opportunities for vistas
Social	… a cohesive whole, of being a world in itself
Diverse	… diversity and variation
Sheltered	… shelter
Serene	… serenity

To test whether the PSDs experienced at the favourite places were in agreement with the framework's underlying theory, correlations between all PSDs were calculated. A cluster analysis was performed using k-means clustering, to see if the favourite places could be organised into discernible clusters according to which PSDs were experienced. To test the connection between PSDs and landscape characteristics, machine learning models (see section 3.4) were fitted using each PSD as a response variable, with the predictors being the same sampled landscape characteristics and attributes related to the person as in **Paper III**.

3.4 Statistical modelling: Boosted regression trees

Statistical analyses in **Paper II-IV** were performed using Boosted Regression Trees (BRT), also known as Gradient Boosting Machine or Generalised Boosting Model (GBM). BRT is a machine learning method based on decision trees (Friedman, 2001). In this approach, the algorithm seeks to find homogeneous clusters in a dataset by using cut-off values of predictors to perform binary splits (Figure 8).

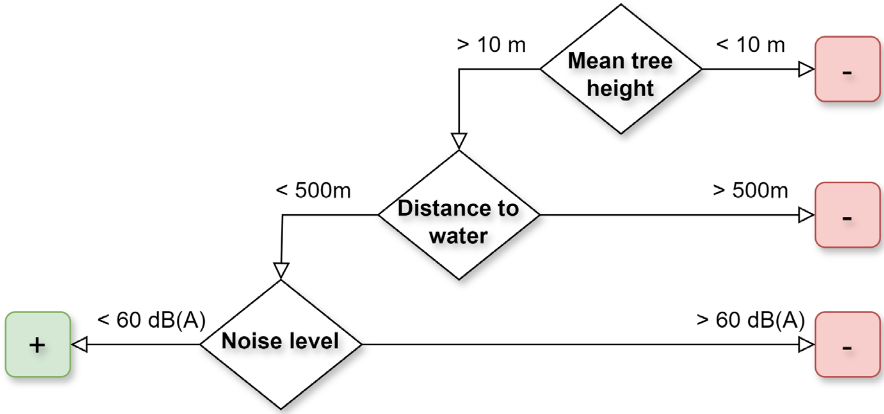


Figure 8. A visual representation of a decision tree model. At each node, the algorithm identifies the predictor that most effectively separates the data into similar groups. This iterative process continues until a stopping criterion, such as a predefined maximum tree depth (illustrated here as three levels of splits). In this hypothetical example, the decision tree begins by partitioning the dataset based on mean tree height, with areas with trees less than 10 metres high being less favoured for recreation. It further divides the data based on distance to water, identifying areas further from water sources as less preferable. Finally, the algorithm partitions the remaining data based on ambient noise levels, favouring quieter areas.

BRT employs an iterative process, building a series of decision trees in a stepwise manner, with each subsequent tree focusing on the instances that were previously misclassified or had large residuals. The final model is the sum of all of these decision trees, which often number in the hundreds or thousands. The iterative process allows BRT to gradually refine its predictions, capturing complex relationships and non-linear patterns that may exist within the data. It can be used for both classification and regression, and handles both numerical and categorical predictors, making it versatile for a wide range of datasets and problems. BRT is robust to overfitting, due to the boosting mechanism, which penalises complex models.

There is of course no free lunch: BRT model fitting can be time and resource-intensive, as several hyperparameters that govern the algorithm need to be tuned to find the best model. This is typically done by fitting models with combinations of a large number of parameter values, and evaluating model accuracy using cross-validation. Another drawback is its potential to act as a 'black box' model, making the inner workings hard to interpret, such as the relationship between the response and specific predictors. This is especially true when higher-order interaction effects are present. For a more in-depth explanation of boosted regression trees, see Elith et al. (2008).

4. Results and discussion

In this section, I summarise the main results of my articles as they relate to the overarching aims of the thesis and discuss their broader implications.

4.1 Aim 1 – An indicator framework for Swedish forests

4.1.1 The conceptual framework

In **Paper I** we presented a recreational indicator framework. The framework posits that individuals' decisions to visit forests for recreation are influenced by two main factors: *accessibility* and the forest's *qualities.* Accessibility refers to the ease with which a visitor can reach a forest (Koppen et al., 2014), while qualities are features of the forest and its environment that affect its appeal for recreation. These were further subdivided into *intrinsic*, *extrinsic*, and *facilitation* qualities: Intrinsic qualities are physical, tangible aspects of the forest, such as its structure and tree species composition. Extrinsic qualities relate to features of the forest's surroundings, such as ambient noise or proximity to water bodies. Facilitation qualities refer to recreational infrastructure such as paths, picnic tables, restrooms, fire pits, informational signage, and other amenities that facilitate the recreational experience).

The framework further defines *realized recreation* as the 'true' recreational value of a specific forest or of a specific recreationist. This can be estimated as the quantity of visits and/or the experienced satisfaction. Conversely, *recreation potential* is a forest's theoretical attractiveness for recreation, regardless to what degree it is currently being used for recreation or not. This aspect can be estimated by using accessibility and the qualities as indicators (Figure 9).

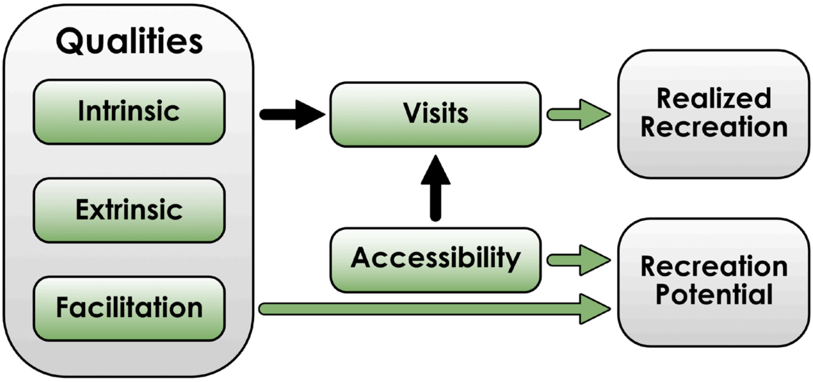


Figure 9. A conceptual model of how recreation visits to a forest are driven by its qualities and accessibility (black arrows). Realized recreational values can be estimated by measuring the number and quality of forest visits by users, while the recreation potential of a forest can be assessed by using the forest's qualities and accessibility as indicators (green arrows).

4.1.2 Literature review of forest qualities

In the review of literature on what forest qualities recreationists prefer we identified six intrinsic qualities (related to forest structure) and four extrinsic qualities (related to the forest's location and surroundings) that play a significant role for the recreational value of forests (Table 3). We found there was evidence for facilitation qualities (e.g. related to recreational infrastructure such as paths, benches, signs etc.) to be important for recreational use, but that there was not sufficient knowledge to split these into separate qualities. For each quality we assessed its current feasibility for inclusion in calculating a recreation potential index for Swedish forests: how strong does the effect on recreational value seem to be, and is it a general effect true for most recreationists? How could each quality be defined? Are the qualities currently possibly to map on a national level in Sweden? For some qualities these questions were easily answered, such as proximity to water: it seems to be a generally preferred quality which strongly increases the recreational value of a forest, of which map data is readily available. For other qualities, such as biodiversity, the feasibility was lower: the evidence for effects on recreational values are mixed, possibly dependent on the specific attitudes of the recreationist, and there is no adequate map data available. In some cases, the issue was not the lack of knowledge on whether it affects preferences or the availability of map data, but more methodological in nature, such as for scenic views: there is good evidence for scenic views being a positive quality, but it is challenging to quantify it in a spatial indicator.

Table 3. Forest qualities identified as important for recreational value in our literature review. 'Importance' represents how strong we assessed the connection between the quality and recreational preferences to be on a scale from 0-3. 'Current feasibility' shows how applicable we assess the quality to be as an indicator, weighing the combined significance with current data availability and methods to estimate the quality.

Intrinsic qualities	**Importance**	**Current feasibility**
Tree size/age	+++	+++
Stand density/visibility	++	+
Traces of forestry operations	+++	
Stand heterogeneity	++	+
Tree species composition	+	++
Biodiversity	+	
Extrinsic qualities		
Proximity to water	+++	+++
Noise	++	++
Topography and views	++	+
Landscape heterogeneity	++	+
Facilitation qualities		
Recreational infrastructure	+++	++

A difficulty in implementing an index of recreational potential is preference heterogeneity—not every recreationist wants the same forest. There is evidence from studies of stated preference that people have different landscape preferences (Eriksson et al., 2012; Elbakidze et al., 2022; Juutinen et al., 2017; Ode Sang et al., 2016; Scott et al., 2009). These differences have been found to be weakly correlated with socio-demographic factors such as age and gender (Giergiczny et al., 2015; Ode-Sang et al., 2016), but more strongly to environmental attitudes, nature-relatedness, and ideological stances (Eriksson et al., 2012; Juutinen et al., 2017; Ode Sang et al., 2016; Scott et al., 2009). Type of activity engaged in has also shown to affect landscape preferences, especially for landscapes' recreational infrastructure (Kienast et al., 2012; de Valck et al., 2016, 2017; Korpilo et al., 2017).

The general magnitude of preference heterogeneity is difficult to estimate, but has in some studies been strong (Giergiczny et al., 2015; Abildtrup et al., 2013), casting some doubt on the approach of calculating a single index for an 'average person'. A solution to this has been suggested in the form of user typologies, where recreationists are grouped into archetypes according to a combination of landscape preferences, recreational behaviour and socio-demographic characteristics (Komossa et al., 2019). A similar line of thinking is found in recreation studies using the wilderness purism scale (Gundersen et al., 2015) and in planning frameworks such as the Recreation Opportunity Spectrum (Manning, 2022), where recreationists are defined on a continuum between 'urban oriented' recreationists (urbanists), who prefer easily accessible and comfortable natural settings, to 'wilderness oriented' recreationists (purists), who seek solitude and more challenging environments. We proposed weighting indicators for different recreationist groups as a reasonable way to address preference heterogeneity, a strategy requiring studies like Komossa et al. (2019), which explore user typologies within the target population of the index.

In **Paper I**, we suggested which indicators may be suitable for creating an index of recreation potential for Swedish forests, but we did not reach a full implementation. Parts remain to be solved, such as how the indicators should be weighted relative to each other, and the spatial resolution to calculate the index on. To achieve this, multiple approaches can be considered. While we believe our review was thorough, it lacked systematic rigour. A systematic review with meta-analysis could have offered a way to parameterize and weight indicators. However, the diverse methodologies in the studies we identified, ranging from observed site choices to expressed preferences via photographs or choice experiments, complicate this approach. Instead, we propose weighting indicators based on realized recreation—studying where people actually go and what they experience. PPGIS methodology would be suitable for capturing both movement patterns and data on the recreationist. To calibrate and validate a nationwide index, large-scale implementation would be needed, to capture a diversity of landscapes and preferences. This could also be performed as part of a systematic monitoring program, tracking how preferences and activities change over time. An embryo for such a program can be seen in the PPGIS data analysed in **Paper II**, which was collected as part of the Swedish Environmental Protection Agency's recurring survey on Swedish people's recreational habits.

4.2 Aim 2 – Landscape characteristics and personal attributes

Paper II and **III** aimed to explore which landscape characteristics, combined with personal attributes, affect where people choose to engage in outdoor recreation. The studies showed partly divergent outcomes, with overall much stronger effects of the tested variables in **Paper III**.

4.2.1 Preference for landscape characteristics

In **Paper II**, the BRT models showed low accuracy, and were only slightly better than chance at distinguishing between the use and the availability samples. This indicated that the included predictors (land cover, heterogeneity, topology, path and road density, forest characteristics and protected areas) had limited impact on site choice (Figure 10).

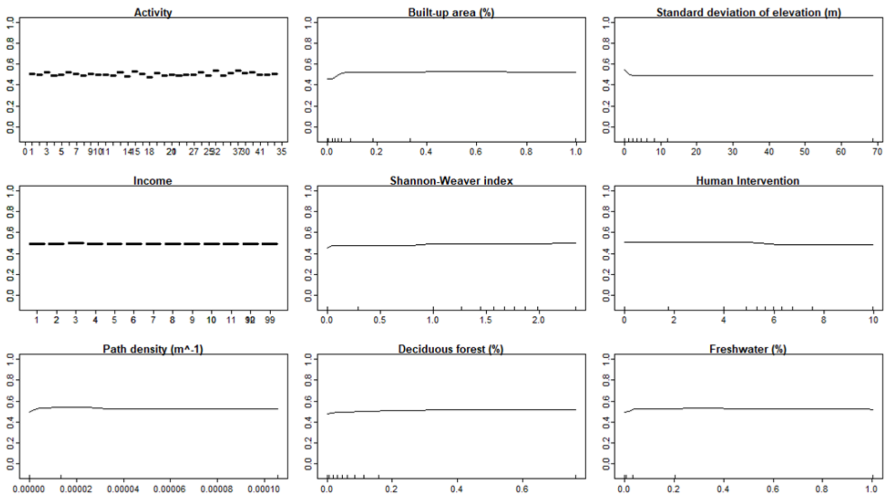


Figure 10. Partial dependence plots for the most influential predictors of one model from **Paper II**, showing how varying each predictor affects the chance an area was chosen for recreation. The flat shapes, suggesting none of the predictors had a large effect, was similar across all influential predictors of the models.

In contrast, the models of **Paper III** were much more accurate, being able to discern between the use sample and availability sample in ~80 % of cases. Analysing the most influential predictors of the models showed that for the route model (Figure 11), proximity to recreational amenities (shelters, fireplaces, and toilets) had the strongest positive effect, where shorter distances increased the probability that a route was used. Increased amounts of built-up area in the viewshed had a strongly negative effect, while path density, deciduous forest, tree height, and freshwater were all positively related with route use. Noise and clearcuts showed negative correlations with route use.

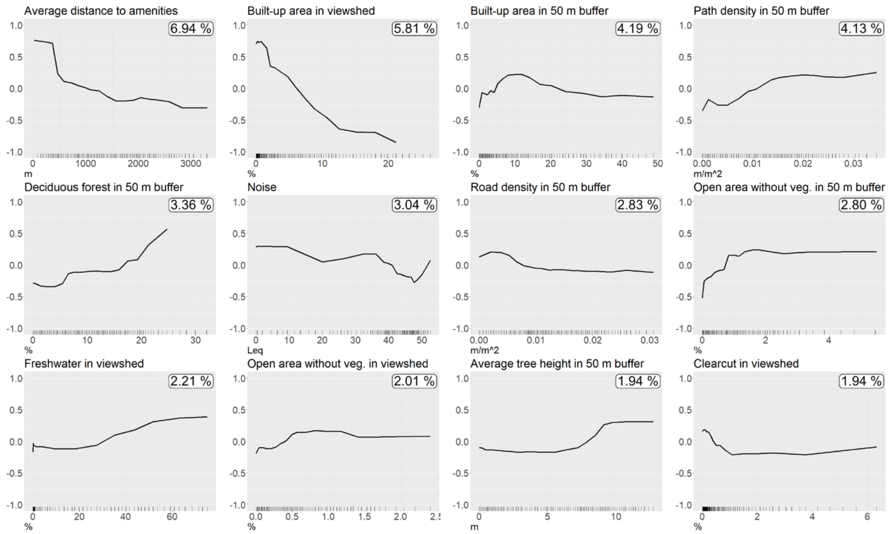


Figure 11. **Paper III**'s comparison of landscape characteristics between used and random routes, showing accumulated local effects for the 12 most influential predictors. A higher y-axis value indicates a greater likelihood of being a used route. The boxes in the upper right corner depict each predictor's influence on model accuracy. The x-axis displays a rug plot, showing data distribution with each notch representing one percentile. Outliers are removed by cutting off the x-axis at 95% of each variable's range.

In the model predicting favourite places (Figure 12), two predictors were considerably more influential than the others and strongly positively correlated with being a favourite place: the amount of freshwater in the viewshed, and proximity to the nearest amenity. Freshwater within the 50 m buffer was also positive, along with the standard deviation of elevation. In the viewshed, the fraction of sea and median elevation were positive, while the fraction built-up area and pine forest was negative. Moreover, viewshed size was positively correlated with being a favourite place.

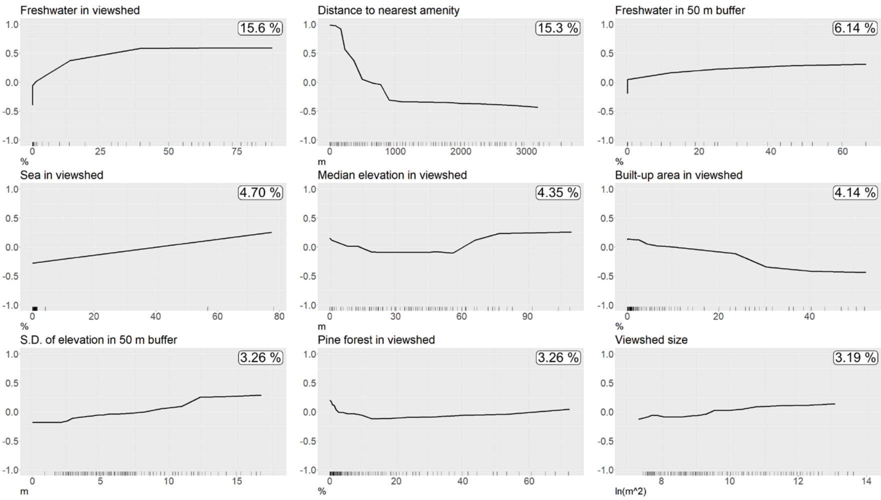


Figure 12. **Paper III**'s comparison of landscape characteristics between favourite places and random locations, showing accumulated local effects for the 9 most influential predictors. A higher y-axis value indicates a greater likelihood of being a favourite place. The boxes in the upper right corner depict each predictor's influence on model accuracy. The x-axis displays a rug plot, showing data distribution with each notch representing one percentile. Outliers are removed by cutting off the x-axis at 95% of each variable's range.

The effects of landscape characteristics on choice of area for recreation in **Paper III** were mostly in line with what was expected from literature. The magnitude of certain effects was surprising, however, especially the strong positive effect of recreational infrastructure. This aspect has been shown to affect where people recreate previously (de Valck et al., 2016; Giergiczny et al., 2015; Kienast et al., 2012), but not to the extent shown here: recreational infrastructure was more influential on model accuracy than all forest-related predictors combined in the favourite places model. This shows the importance that these kinds of facilitations can have on the recreational potential of landscapes, implying that they are important to include as an indicator of recreation potential as well. As described in 4.1, this is however complicated by preference heterogeneity linked to recreational infrastructure, and a lack of available map data.

Water environments were important for the choice of recreation sites, which also has been shown previously in studies of both stated (de Valck et al., 2017; Kaplan & Kaplan, 1989; White et al., 2010) and revealed preferences (Baumeister et al., 2020; Kienast et al., 2012; Laatikainen et al., 2017). Our results highlight the importance of preserving water environments for recreational use and protecting them from exploitation.

The observed preference for taller trees and avoidance of clearcuts is consistent with previous studies on stated preferences (Gundersen et al., 2019). The presence of deciduous trees were positive, which could be explained either from a preference for deciduous trees themselves, or that the presence of deciduous trees yields a higher perceived diversity, which has been shown to be preferable (Filyushkina et al., 2017). These results support the claims of higher recreational values when applying alternative forestry methods such as continuous cover forestry rather than even-aged forestry with large clearcuts, which currently is the prevalent method in Fennoscandia (Pukkala et al., 2012).

Both models consistently showed that recreationists tend to avoid urban areas, with the route model highlighting a negative impact of noise. Anthropogenic noise, particularly from sources like traffic, has been proven to adversely affect perceptions of natural settings in research studies (Benfield et al., 2020; Li et al., 2018). Tranquility has emerged as a significant motivator for recreationists, with the concept denoting not just the absence of noise but also the restorative experience certain landscape elements (such as water) can imbue (Wartmann et al., 2019). This aligns with increasing research on the concept of soundscapes, which explores the spatial perception of sounds within landscapes (Ratcliffe, 2021), and which have been integrated into PPGIS research (Korpilo et al., 2023). Despite its importance, noise's impact on landscape use remains a relatively understudied area; we could only find one study revealing changes in actual behaviour linked to noise (Krog et al., 2010).

The models showed that topography was important: people preferred a landscape of varying altitude that also yielded a large view, as long as those views were not of urban areas or clearcuts. We found that both a high elevation and a low elevation were positive, which we interpret as representing both a preference for height and for close-to-sea areas, both of which yield large views with long sight lines. Earlier studies have found a general preference for views (Gundersen et al., 2019; Kaplan & Kaplan, 1989; Kienast et al., 2012). Our results confirm such findings, and highlight that this is an important aspect to consider in landscape planning for recreation, with viewshed analysis being a possible route to identify this characteristic in the landscape.

The most likely explanation of the increase in model accuracy between the two studies is the improved methodology: **Paper II** controlled for accessibility in a less accurate way due to the lack of information on home location, and had a less realistic manner of sampling what landscape the recreationist experience, only employing circular buffers and not viewsheds. The difference in outcomes could, however, possibly also be due to other reasons: The scopes were different, with the restricted study area of **Paper III** (Umeå municipality) enabling the usage of slightly more detailed map data, including e.g. recreational infrastructure. The smaller scale also probably captured a more uniform subset of recreationists than the national sample in **Paper II**.

4.2.2 Preference heterogeneity

Both PPGIS surveys included predictors that previously have been shown to modulate the effect of landscape preferences, such as recreationists' gender and age (Gunnarsson et al., 2017; Kienast et al., 2012), the time of year (Gatti et al., 2022) and type of activity performed (de Valck et al., 2017; Gerstenberg et al., 2020). As noted above, the models in **Paper II** were weak, with the most influential effect stemming from the type of activity engaged in. This predictor's explanatory power derived from many interactions with other landscape predictors. These interactions were so weak as to be uninterpretable by themselves but suggest that type of activity to some extent changes which landscape recreationists choose, which is in line with previous research both on stated and revealed preferences (de Valck et al., 2016, 2017; Gerstenberg et al., 2020).

In **Paper III**, the models were much more robust, but here we instead saw no evidence of preference heterogeneity due to socio-demographic variables, characteristics of the visit (such as type of activity, visit frequency etc.) nor the season. To note here is that for **Paper III** the number of activities that the respondents could choose were much reduced, and the dataset was dominated by walking, which combined with the smaller sample size of **Paper III** could be one reason for not finding an effect of activity type. Regarding socio-demographic factors the lack of an effect was unsurprising, as these earlier studies also have found them to be weak (Giergiczny et al., 2015). Contrary to previous findings with Swedish recreationists (Ode-Sang et al., 2016; Gunnarsson et al., 2017), our questions about the respondent's identification as an 'urban person' or 'nature person' did not yield significant effects. We see these questions as connected to the broader concept of natural relatedness (Nisbet et al., 2009), which has been shown to affect preferences and recreational behaviours (Flowers et al., 2016; Elbakidze et al., 2022).

Preference heterogeneity has mostly been shown in studies using stated preference methods, and only a few relying on revealed preference (Kienast et al., 2012; Ode-Sang et al., 2016; de Valck et al., 2016; Agimass et al., 2018). Among these, only Agimass et al. (2018) analyse individual-level data, whereas the others focus on population-level analysis, segmenting samples by activity (de Valck et al., 2016; Gerstenberg et al., 2020), or age (Kienast et al., 2012) for comparison. Our findings suggest that preference heterogeneity observed in stated preference studies may not convert into actual differences in landscape usage. However, our studies did not account for potential influences like environmental attitudes (Eriksson et al., 2012; Juutinen et al., 2017) and knowledge (van der Wal et al., 2014). Given our models' limitations, some preference heterogeneity may remain unexplained.

4.3 Aim 3 – A Novel methodology for PPGIS studies

In my PPGIS studies I employed three methods (Boosted Regression Trees modelling, viewsheds, and controlling for accessibility with network analysis) that while not novel in themselves, had either not been used in PPGIS studies before, or not been combined in this way. In **Paper III**, this combined methodology yielded high accuracy models.

4.3.1 Boosted regression trees

Machine learning is currently in vogue, receiving increased attention from both researchers and the public (Holzinger et al., 2018). It is a broad term, covering many different algorithms and techniques. One suggested definition has been ‘a field of statistical research for training computational algorithms that split, sort and transform a set of data to maximise the ability to classify, predict, cluster or discover patterns in a target dataset’ (Reichstein et al., 2019). With machine learning techniques, data are empirically modelled with few or no prior assumptions about the system, enabling data inferences without relying on causal theory. Handling non-linear predictors and high-dimensional interactions, it has been suggested to be particularly useful in research on ecosystem services, as these often are complex systems (Scowen et al., 2021).

The type of machine learning I have employed (BRT) is not novel (Friedman, 2001), but as with most innovations in statistical methods, adoption by researchers is slow (Sharpe, 2013). Within recreation research I have not been able to find other examples of it being used, but Random Forests, a similar algorithm based on ensembles of decision trees, has been deployed a few times (Baumeister et al., 2020, Nyelele et al., 2023, Manley & Egoh 2022). Of these, the study by Baumeister et al. (2020) is the most similar to my approach, also analysing PPGIS data with landscape covariates. A common use case for machine learning within recreation research has been in the analysis of social media data, which can produce large amounts of data. Although I did not set out to explicitly compare the benefits of BRT modelling to other techniques, I found working with them highly suitable for the task. Non-linear relationships were anticipated for many predictors (e.g. land cover composition), combined with the presence of high-order interactions between predictors, while we had few prior hypotheses regarding their exact nature. By avoiding the need for model selection or the pre-specification of interactions, I was able to incorporate all relevant map data available for analysis.

As described in section 3.4, machine learning modelling can be computationally demanding and yield less interpretable models. While hardware upgrades can address the computational issue, model interpretability remains a challenge. However, techniques like the visualizations in the Interpretable Machine Learning package in R (Molnar et al., 2018) can aid in this. Alternatively, methods like Local Interpretable Model-agnostic Explanations (LIME) offer interpretable surrogate models for specific predictions (Ribeiro et al., 2016). Implementing machine learning has become more accessible with numerous guides tailored to specific research areas, such as biology and ecology (Elith et al., 2008; Greener et al., 2022), survey data analysis (Kern et al., 2019), and managing big data in outdoor recreation studies (Dagan & Wilkins 2023).

4.3.2 Controlling for accessibility

Many PPGIS studies on recreational preference have not controlled for accessibility, or only done so to a limited extent (e.g. on a population level), possibly yielding inaccurate results. Our solution, employing network analysis in a use/available framework (section 3.2.2) makes it possible to control for this factor for each individual recreationist. Network analysis has been employed previously in studies of recreation, but mainly as a tool to analyse the accessibility of various features, such as attractive water environments (Laatikainen et al., 2017). It is not difficult to employ but necessitates detailed road/path network data. Network analysis has been suggested to be an improvement over using straight line distances for estimating accessibility to green space (Wang et al., 2021; Wolffe, 2021).

Our methodology has room for improvement. Here, we created an availability sample by duplicating the performed route at an equally accessible distance from home. While this approach aims to compare chosen routes with potential alternatives, replicating exact routes may be unrealistic. Recreationists do not typically follow specific route shapes; instead, the landscape itself influences their movement through barriers and paths. An alternative could be agent-based modelling: by using recreationists' actual routes to train a model (Morelle et al., 2019), synthetic routes that better mimic real-world landscape usage can be generated, serving as a more representative availability sample.

4.3.3 Viewshed analysis

In **Paper III-IV**, we included viewsheds along visited routes and at favourite places to increase the realism of the estimate of what landscape each recreationist experienced. Viewsheds have previously been used in recreation research, for instance for modelling the aesthetic value of landscapes using crowdsourced photographs (Karasov et al., 2020; Tenerelli et al., 2017; Yoshimura & Hiura, 2017). We could however only find one example of it being applied to PPGIS data: Ridding et al. (2018) also calculated viewsheds using point data of peoples ‘important outdoor places’. Similarly to our analysis, they sampled landscape characteristics in viewsheds and circular buffers around each point and compared them to random points in the vicinity.

To employ viewshed analysis a digital surface model (DSM) with high spatial accuracy is needed (Lagner et al., 2018), and it is relatively demanding on computing power. There are several avenues of possible development of viewshed methodology that could be useful for recreation research. One is the issue of how to handle vegetation: vegetation here was modelled as being entirely transparent up until 50 m from each point, after which it was assumed to be entirely opaque. This was a compromise, as only using the digital terrain model (DTM) for calculation without taking vegetation into account would lead to unrealistically large viewsheds, while the opposite would lead to unrealistically small viewsheds. There have been attempts to solve this issue by taking the permeability of different kinds of vegetation into account (Ruzickova et al., 2021), or even modelling individual trees using LiDAR (Budei et al., 2018). Such methods are however not currently incorporated in available GIS software. In a similar vein, our viewsheds were not adjusted for the difference between summer and winter, where the loss of leaves from deciduous trees improves visibility.

A further development would be to move beyond treating the viewshed as a single entity. Certain characteristics of the landscape probably affect recreational attractiveness more when in the near view than when far away, and vice versa. A possible route of analysis to see this is to divide the viewshed into distinct zones and analyse these separately, which could provide a more nuanced understanding (Schirpke et al., 2013). Additionally, there have been suggestions of employing more advanced methodologies to assess the perceived visual impact of various landscape elements (Chamberlain & Meitner, 2013; Nutsford et al., 2015). Such approaches could enable a more comprehensive evaluation of visual aesthetics and landscape quality.

4.4 Aim 4 – Evaluating the Perceived Sensory Dimensions framework

The aim of **Paper IV** was to evaluate the PSD framework by studying the experienced PSDs reported from the favourite places collected in the Umeå PPGIS survey. This was done both in terms of whether the PSDs experienced at the favourite places for recreation aligned with the underlying theory of the framework, but also to what degree the PSDs were linked to landscape characteristics.

4.4.1 Internal validity

The distributions of PSDs experienced at favourite places skewed towards the higher end of the scale, with all median values at 50 or above (Figure 13). Sheltered, Open and Serene were overall the qualities most strongly perceived. Sheltered and Serene are both qualities that have been strongly associated with restoration of high stress levels and cognitive fatigue (Pálsdóttir et al., 2018), which is a common motivation for outdoor recreation. Open is associated with long, unbroken sightlines, and plenty of space to roam freely without physical obstacles, which lines up with the results of **Paper III** where large viewsheds were shown to be important.

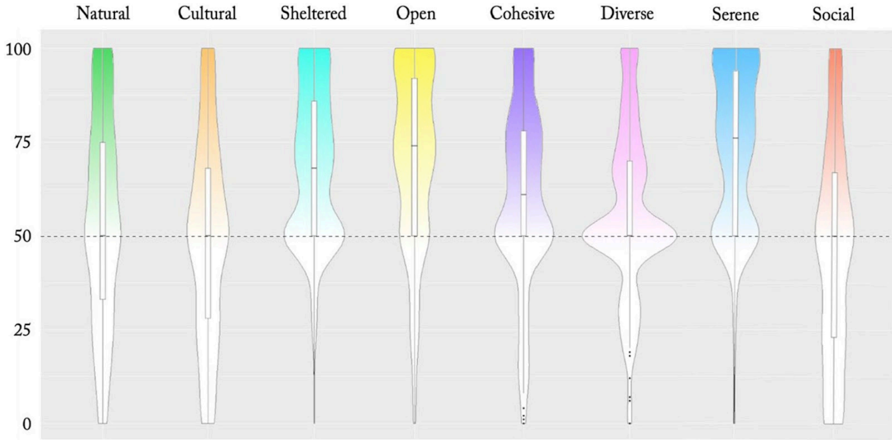


Figure 13. Violin diagrams showing rated perceived sensory dimensions at people's favourite places. The width of each violin is a smoothed density plot, corresponding to the amount of data for each value along the y-axis. Within each violin is a box plot showing quantiles.

The theoretical basis of the PSD framework is an oblique rotation factor analysis, where each quality is meant to assess a distinct aspect of the perceived environment. However, some correlation between qualities is expected, with those closer to each other in the PSD wheel (section 2.2, Figure 1) being positively correlated and PSDs opposite each other negatively correlated. The correlations between the different PSDs experienced at the favourite places in **Paper IV** revealed this to be mostly true, with no correlation coefficient exceeding 0.5, and the signs and magnitudes of the correlations in line with expectations (Table 4).

Table 4. Correlations between PSDs experienced at recreationists' favourite places. Positive correlations in green and negative in red, with more saturated colours indicating stronger correlations.

	Cohesive	Serene	Natural	Sheltered	Diverse	Social	Cultural
Open	0.17	0.1	0.16	0.1	0.03	0.24	0.14
Cul.	0.00	-0.17	-0.28	0.09	0.17	0.43	
Soc.	0.02	-0.11	-0.07	0.14	0.29		
Div.	0.42	0.15	0.27	0.44			
She.	0.42	0.48	0.20				
Nat.	0.48	0.35					
Ser.	0.38						

To see which PSDs often occurred together, a clustering analysis was performed. This showed that the favourite places were split in two main groups, with Group 1 expressing a more outward directed or activity-oriented recreational experience (related to the PSDs Cultural and Social), whereas Group 2 expressing a more rest-oriented recreation (related to PSDs Natural and Serene; Figure 14).

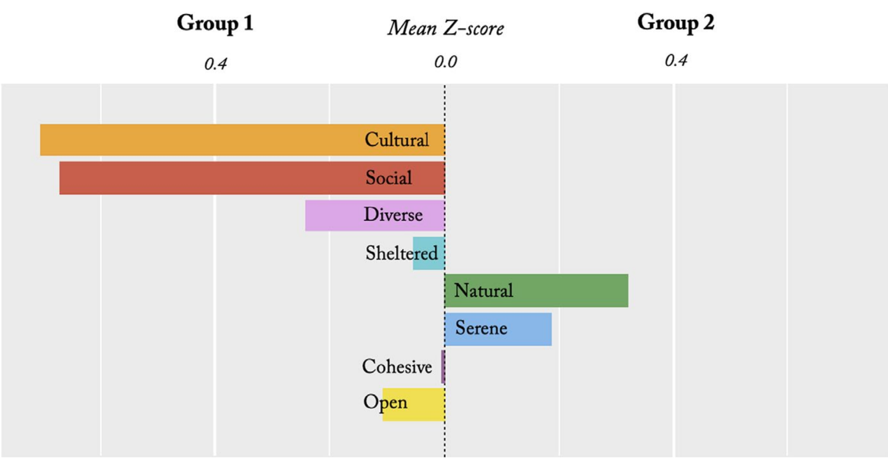


Figure 14. A clustering analysis of the dataset showed a main split between a Cultural-social axis and a Natural-Serene axis.

4.4.2 Linking PSDs to landscape characteristics

The BRT analysis revealed that landscape characteristics could not explain the PSDs experienced at favourite places. Of the eight BRT models, most had very low explanatory power, with the strongest models being the ones related to the PSDs Natural ($R^2 = 0.27$), Social ($R^2 = 0.19$) and Cultural ($R^2 = 0.14$). The Natural model had several expected effects, such as being negatively correlated with built-up areas and noise, however the strongest effect was the degree to which the respondent self-identified as a nature-oriented person, with a positive correlation (Figure 15). This predictor, along with its counterpart (to what degree the respondent self-identified as an urban person) were also influential in the Social and Cultural models.

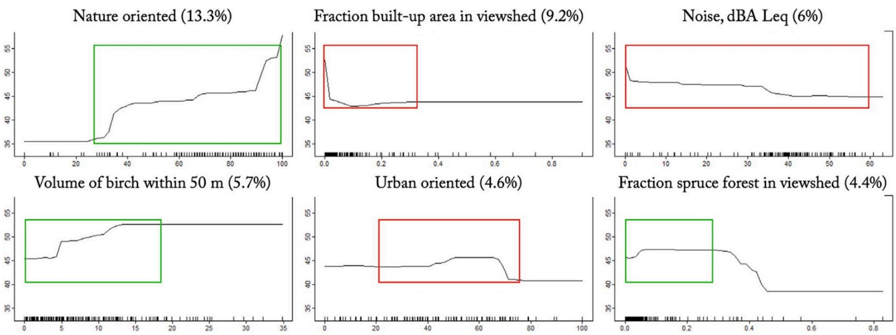


Figure 15. Partial dependence plots for the six most influential predictors of the Natural PSD model, showing how the PSD value (Y-axis) is predicted to change as the value of each predictor changes. Relative influence of each predictor within parentheses (%). Above the x-axis is a rug plot, showing the distribution of values within the data set, with each notch representing 1% of the data set. The graphs show the entire range of values for each predictor within the data set, but as the algorithm fits few trees where there is little data, interpretation should focus on sections with higher data densities, approximately highlighted with rectangles in green (positive) and red (negative).

Overall, the results suggest that the link between PSDs and landscape characteristics are weak, and that they cannot reliably be estimated by landscape characteristics alone, instead emanating from the individual experience. This finding is contrary to assumptions in previous research, which have used landscape characteristics as proxies for the PSDs (Björk et al., 2008; Annerstedt van den Bosch et al., 2015). It also contradicts results showing that there is consistency between individuals in how PSDs are experienced (Qiu & Nielsen, 2015).

4.4.3 Applicability of the PSD framework

This study was an example of how PPGIS can be used to validate frameworks and indices. Considering these results, we argue that the PSD framework is not robust enough to be applied in the same manner as, for instance, recreational potential indices discussed in **Paper I**. Personal attributes were a stronger predictor of PSDs than the landscape characteristics, which is not surprising as the framework has its basis in how individuals perceive their environment. However, our results confirm the internal validity of the framework, and we argue that it still has merit as a planning and design tool on a smaller scale.

5. Conclusions and future perspective

This thesis has explored various facets of landscape preferences of recreationists. **Paper I** reviewed current knowledge of which forest characteristics are preferred, assessing how applicable they are as indicators of recreational potential, and arranged them into a framework. **Papers II-IV** developed novel methodology for PPGIS research, which in **Paper III** yielded highly accurate models that gave actionable insights into landscape preferences of recreationists. Finally, **Paper IV** showcased how PPGIS can be used for validation, by evaluating the Perceived Sensory Dimensions framework's mapping suitability based on landscape characteristics, revealing it as unsuitable for this purpose.

Paper I serves as a starting point for an index of recreational potential in Swedish forests. This need is underscored in Sweden's review of national outdoor recreation goals, which identifies a lack of suitable methods for defining recreation areas as a barrier to assessing goal achievement (Swedish Environmental Protection Agency, 2023). To complete the work, collection of more data is required to weight indicators and validate the output. This could take the form systematic monitoring: deployment of large-scale PPGIS of Swedish recreational visits, where experiences of various landscapes are collected. Proposals of such programmes have been the subject of government commissions on outdoor life, protected nature and rural development previously on several occasions, but have never been implemented (Nordic Council of Ministers, 2013). The PPGIS data used in **Paper II** is an embryo of such a programme, which could be further developed. The results of **Paper I** are also relevant on a smaller scale, as the literature review of which forest characteristics matter for recreation can inform e.g. municipal forest management. The index of recreational potential could be adjusted to address the more detailed scale of local planning, where a need for useful tools has been identified (Petersson-Forsberg, 2014).

One of the goals of the methodology developed in **Paper II-III** was to control for the effect of accessibility. It is a strong factor influencing where recreation occurs, and if the goal of a revealed preference study is to determine which landscape characteristics are attractive, this factor must be addressed. Accessibility is also relevant when it comes to measuring one of the ultimate goals of recreation research: that residents have access to suitable areas for recreation. Previously suggested policy goals, e.g. that people should have a certain maximum distance to green space, might be too simplistic, especially if straight line distances are used rather than network analysis. There's also an argument to be made that only focusing on minimising this distance might be too reductive, as people do not only use their closest area, instead having a 'portfolio of natural places' they visit on various spatial scales (Bijker & Sijtsma, 2017). As such, estimating the cumulative recreational area at various distances could instead be a suitable indicator of accessibility to recreation.

A further methodological advancement was the inclusion of viewsheds in **Paper III-IV**. With this approach we increased the realism of spatial studies by trying to see through the eyes of the recreationist to estimate the landscape they experienced. We also attempted to hear through their ears by including estimated noise maps as a covariate. While previous research has demonstrated the adverse effects of noise on individuals, **Paper III** revealed its influence on recreationists' movement patterns, showing that individuals actively avoid noisy areas. The field of soundscapes has gained prominence recently, emphasising the significance of both positive elements like birdsong or running water, and negative aspects like noise (Ratcliffe, 2021). Elements of the sound environment have also been integrated into PPGIS analyses, where residents are asked to map areas where they perceive positive and negative sounds (Korpilo et al., 2023). Looking ahead, future PPGIS studies might consider incorporating additional sensory dimensions; for instance, the smell of nature has been found to significantly impact environmental experiences (Hedblom et al., 2019).

The results of **Paper III** have several policy implications. Avoiding the visibility and noise of the urban environment seem to be an essential aspect of recreation. Noise is a pervasive problem: a national Swedish survey found that 50 % of recreationists regularly experience noise (Swedish

Environmental Protection Agency, 2019a). In our models we employed maps of estimated noise levels, which demonstrate the utility of such noise mapping, and underscores efforts to map and protect 'quiet areas' where noise is minimised (Cerwén & Mossberg, 2019). Such quiet areas are currently rare, suggesting a need to set aside larger areas for recreation, or minimise noise pollution, to provide the sense of escape and tranquillity recreationists seek (Jönköping Administrative County Board, 2015). Proximity to water and views of water were highly preferred, highlighting the need to protect such areas from exploitation. Shoreline protection legislation is a recurring topic in Swedish politics, with the latest proposition for a weakening of shoreline legislation proposed in 2022 (Proposition 2021/22:168). Finally, recreational infrastructure showed strong positive effects, suggesting increased prioritisation of resource allocation into facilitation of areas as a way to improve recreational opportunities.

The findings from **Paper III**, coupled with the insights from the literature review in **Paper I**, indicate that current silvicultural practices are often ill-suited for forests intended for recreation. Key concerns include minimising visible forestry impacts, as these significantly diminish recreational value. Conversely, unmanaged forests may become too dense, requiring selective logging to enhance visibility. Management strategies should prioritise expanding views of natural features like water bodies, while limiting sightlines towards urban areas. Additionally, efforts should be made to promote the growth of large trees, while still ensuring a diversity of tree sizes and species, especially deciduous species.

While outdoor recreation today is acknowledged as a fundamental human need, it is often neglected in policy and legislation (Mann et al., 2010). This is partly reflected in ecosystem service research, where despite a large amount of effort and many comprehensive sets of indicators being developed, recreation remains elusive (Tiemann & Ring, 2022). These shortcomings are not surprising given the complexity of recreation—a subjective experience crafted in the interplay between individuals and landscapes (Lothian 1999). Attempting to quantify it is a challenge, and a flawless index of recreation will never be developed; aspects are always lost in translation, such as place attachment (Lewicka, 2011). But as the old aphorism goes: 'All models are wrong, but some are useful'. My sincere hope is that this thesis has contributed to the quest of achieving such useful models.

References

Abildtrup, J., Garcia, S., Olsen, S. B., & Stenger, A. (2013). Spatial preference heterogeneity in forest recreation. *Ecological Economics*, *92*, 67–77. doi - 10.1016/j.ecolecon.2013.01.001

Agimass, F., Lundhede, T., Panduro, T. E., & Jacobsen, J. B. (2018). The choice of forest site for recreation: A revealed preference analysis using spatial data. *Ecosystem Services*, *31*, 445–454. doi - 10.1016/j.ecoser.2017.11.016

Annerstedt van den Bosch, M., Östergren, P-O., Grahn, P., Skärbäck, E., & Währborg, P. (2015). Moving to Serene Nature May Prevent Poor Mental Health—Results from a Swedish Longitudinal Cohort Study. *International Journal of Environmental Research and Public Health*, *12*(7), 7974–7989. doi - 10.3390/ijerph120707974

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. doi - 10.1080/1364557032000119616

Baumeister, C. F., Gerstenberg, T., Plieninger, T., & Schraml, U. (2020). Exploring cultural ecosystem service hotspots: Linking multiple urban forest features with public participation mapping data. *Urban Forestry & Urban Greening*, *48*, 126561. doi - 10.1016/j.ufug.2019.126561

Beeco, J. A., Hallo, J. C., & Brownlee, M. T. J. (2014). GPS Visitor Tracking and Recreation Suitability Mapping: Tools for understanding and managing visitor use. *Landscape and Urban Planning*, *127*, 136–145. doi - 10.1016/j.landurbplan.2014.04.002

Benfield, J. A., Rainbolt, G. A. N., Troup, L. J., & Bell, P. A. (2020). Anthropogenic noise source and intensity effects on mood and relaxation in simulated park environments. *Frontiers in Psychology*, *11*. Retrieved from https://www.frontiersin.org/articles/10.3389/fpsyg.2020.570694

Bergren-Bärring, A.-M., & Grahn, P. (1995). *Grönstrukturens betydelse för användningen.* (No. 95(3)). Institutionen för landskapsplanering.

Bergroth, C. (2019). *Uncovering population dynamics using mobile phone data: The case of Helsinki Metropolitan area*. Master thesis, University of Helsinki.

Bijker, R. A., & Sijtsma, F. J. (2017). A portfolio of natural places: Using a participatory GIS tool to compare the appreciation and use of green spaces inside and outside urban areas by urban residents. *Landscape and Urban Planning*, *158*, 155–165. doi - 10.1016/j.landurbplan.2016.10.004

Björk, J., Albin, M., Grahn, P., Jacobsson, H., Ardö, J., Wadbro, J., … Skärbäck, E. (2008). Recreational values of the natural environment in relation to neighbourhood satisfaction, physical activity, obesity and wellbeing. *Journal of Epidemiology & Community Health*, *62*(4), e2–e2. doi - 10.1136/jech.2007.062414

Boll, T., von Haaren, C., & von Ruschkowski, E. (2014). The Preference and Actual Use of Different Types of Rural Recreation Areas by Urban Dwellers—The Hamburg Case Study. *PLoS ONE*, *9*(10), e108638. doi - 10.1371/journal.pone.0108638

Bratman, G. N., Anderson, C. B., Berman, M. G., Cochran, B., de Vries, S., Flanders, J., … Daily, G. C. (2019). Nature and mental health: An ecosystem service perspective. *Science Advances*, *5*(7), eaax0903. doi - 10.1126/sciadv.aax0903

Brown, G., & Fagerholm, N. (2015). Empirical PPGIS/PGIS mapping of ecosystem services: A review and evaluation. *Ecosystem Services*, *13*, 119–133. doi - 10.1016/j.ecoser.2014.10.007

Budei, B. C., St-Onge, B., Hopkinson, C., & Audet, F.-A. (2018). Identifying the genus or species of individual trees using a three-wavelength airborne lidar system. *Remote Sensing of Environment*, *204*, 632–647. doi - 10.1016/j.rse.2017.09.037

Byczek, C., Longaretti, P.-Y., Renaud, J., & Lavorel, S. (2018). Benefits of crowd-sourced GPS information for modelling the recreation ecosystem service. *PLOS ONE*, *13*(10), e0202645. doi - 10.1371/journal.pone.0202645

Casado-Arzuaga, I., Onaindia, M., Madariaga, I., & Verburg, P. H. (2014). Mapping recreation and aesthetic value of ecosystems in the Bilbao Metropolitan Greenbelt (northern Spain) to support landscape planning. *Landscape Ecology*, *29*(8), 1393–1405. doi - 10.1007/s10980-013-9945-2

Cerwén, G., & Mossberg, F. (2019). Implementation of Quiet Areas in Sweden. *International Journal of Environmental Research and Public Health*, *16*(1), 134. doi - 10.3390/ijerph16010134

Chamberlain, B. C., & Meitner, M. J. (2013). A route-based visibility analysis for landscape management. *Landscape and Urban Planning*, *111*, 13–24. doi - 10.1016/j.landurbplan.2012.12.004

Colding, J., Gren, Å., & Barthel, S. (2020). The Incremental Demise of Urban Green Spaces. *Land*, *9*(5), 162. doi - 10.3390/land9050162

Cordingley, J. E., Newton, A. C., Rose, R. J., Clarke, R. T., & Bullock, J. M. (2015). Habitat Fragmentation Intensifies Trade-Offs between Biodiversity and Ecosystem Services in a Heathland Ecosystem in Southern England. *PLOS ONE*, *10*(6), e0130004. doi - 10.1371/journal.pone.0130004

Dagan, D. T., & Wilkins, E. J. (2023). What is “big data” and how should we use it? The role of large datasets, secondary data, and associated analysis techniques in outdoor recreation research. *Journal of Outdoor Recreation and Tourism*, *44*, 100668. doi - 10.1016/j.jort.2023.100668

De Valck, J., Broekx, S., Liekens, I., De Nocker, L., Van Orshoven, J., & Vranken, L. (2016). Contrasting collective preferences for outdoor recreation and substitutability of nature areas using hot spot mapping. *Landscape and Urban Planning*, *151*, 64–78. doi - 10.1016/j.landurbplan.2016.03.008

De Valck, J., Landuyt, D., Broekx, S., Liekens, I., De Nocker, L., & Vranken, L. (2017). Outdoor recreation in various landscapes: Which site characteristics really matter? *Land Use Policy*, *65*, 186–197. doi - 10.1016/j.landusepol.2017.04.009

De Valck, J., & Rolfe, J. (2018). Spatial Heterogeneity in Stated Preference Valuation: Status, Challenges and Road Ahead. *International Review of Environmental and Resource Economics*, *11*(4), 355–422. doi - 10.1561/101.00000097

de Vries, S., Lankhorst, J. R.-K., & Buijs, A. E. (2007). Mapping the attractiveness of the Dutch countryside: a GIS-based landscape appreciation model. *For. Snow Landsc. Res.*

Elbakidze, M., Dawson, L., Milberg, P., Mikusiński, G., Hedblom, M., Kruhlov, I., … Grodzynskyi, M. (2022). Multiple factors shape the interaction of people with urban greenspace: Sweden as a case study. *Urban Forestry & Urban Greening*, *74*, 127672. doi - 10.1016/j.ufug.2022.127672

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*(4), 802–813. doi - 10.1111/j.1365-2656.2008.01390.x

Eriksson, L., Nordlund, A. M., Olsson, O., & Westin, K. (2012). Recreation in Different Forest Settings: A Scene Preference Study. *Forests*, *3*(4), 923–943. doi - 10.3390/f3040923

Filyushkina, A., Agimass, F., Lundhede, T., Strange, N., & Jacobsen, J. B. (2017). Preferences for variation in forest characteristics: Does diversity between stands matter? *Ecological Economics*, *140*, 22–29. doi - 10.1016/j.ecolecon.2017.04.010

Flowers, E. P., Freeman, P., & Gladwell, V. F. (2016). A cross-sectional study examining predictors of visit frequency to local green space and the impact this has on physical activity levels. *BMC Public Health*, *16*(1), 420. doi - 10.1186/s12889-016-3050-9

FOREST EUROPE. (2020). *State of Europe's Forests 2020*. Retrieved from https://foresteurope.org/state-europes-forests-2020/

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5). doi - 10.1214/aos/1013203451

Gatti, E. T. J., Brownlee, M. T. J., & Bricker, K. S. (2022). Winter recreationists' perspectives on seasonal differences in the outdoor recreation setting. *Journal of Outdoor Recreation and Tourism*, *37*, 100366. doi - 10.1016/j.jort.2021.100366

Gerstenberg, T., Baumeister, C. F., Schraml, U., & Plieninger, T. (2020). Hot routes in urban forests: The impact of multiple landscape features on recreational use intensity. *Landscape and Urban Planning*, *203*, 103888. doi - 10.1016/j.landurbplan.2020.103888

Giergiczny, M., Czajkowski, M., Żylicz, T., & Angelstam, P. (2015). Choice experiment assessment of public preferences for forest structural attributes. *Ecological Economics*, *119*, 8–23. doi - 10.1016/j.ecolecon.2015.07.032

Gosal, A. S., Giannichi, M. L., Beckmann, M., Comber, A., Massenberg, J. R., Palliwoda, J., … Ziv, G. (2021). Do drivers of nature visitation vary spatially? The importance of context for understanding visitation of nature areas in Europe and North America. *Science of The Total Environment*, *776*, 145190. doi - 10.1016/j.scitotenv.2021.145190

Greener, J. G., Kandathil, S. M., Moffat, L., & Jones, D. T. (2022). A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology*, *23*(1), 40–55. doi - 10.1038/s41580-021-00407-0

Gundersen, V., & Frivold, L. H. (2008). Public preferences for forest structures: A review of quantitative surveys from Finland, Norway and Sweden. *Urban Forestry & Urban Greening*, *7*(4), 241–258. doi - 10.1016/j.ufug.2008.05.001

Gundersen, Vegard, Tangeland, T., & Kaltenborn, B. P. (2015). Planning for recreation along the opportunity spectrum: The case of Oslo, Norway. *Urban Forestry & Urban Greening*, *14*(2), 210–217. doi - 10.1016/j.ufug.2015.01.006

Gunnarsson, B., Knez, I., Hedblom, M., & Sang, Å. O. (2017). Effects of biodiversity and environment-related attitude on perception of urban green space. *Urban Ecosystems*, *20*(1), 37–49. doi - 10.1007/s11252-016-0581-x

Hartig, T., Mitchell, R., de Vries, S., & Frumkin, H. (2014). Nature and Health. *Annual Review of Public Health*, *35*(1), 207–228. doi - 10.1146/annurev-publhealth-032013-182443

Hedblom, M., Gunnarsson, B., Iravani, B., Knez, I., Schaefer, M., Thorsson, P., & Lundström, J. N. (2019). Reduction of physiological stress by urban green space in a multisensory virtual experiment. *Scientific Reports*, *9*(1), 10113. doi - 10.1038/s41598-019-46099-7

Hermes, J., Van Berkel, D., Burkhard, B., Plieninger, T., Fagerholm, N., von Haaren, C., & Albert, C. (2018). Assessment and valuation of recreational ecosystem services of landscapes. *Ecosystem Services*, *31*, 289–295. doi - 10.1016/j.ecoser.2018.04.011

Holzinger, A., Kieseberg, P., Weippl, E., & Tjoa, A. M. (2018). Current advances, trends and challenges of machine learning and knowledge extraction: from machine learning to explainable AI. In A. Holzinger, P. Kieseberg, A. M. Tjoa, & E. Weippl (Eds.), *Machine Learning and Knowledge Extraction* (pp. 1–8). Cham: Springer International Publishing. doi - 10.1007/978-3-319-99740-7_1

Hörnsten, L., & Fredman, P. (2000). On the distance to recreational forests in Sweden. *Landscape and Urban Planning*, *51*(1), 1–10. doi - 10.1016/S0169-2046(00)00097-9

Jönköping Administrative County Board. (2015). *Ostörda områden - Var finns de? En GIS-modell för identifiering av bullerfria områden*. Retrieved from http://ext-dokument.lansstyrelsen.se/jonkoping/GDK_dokument/2015-01_Ostorda_omraden.pdf

Juutinen, A., Kosenius, A.-K., Ovaskainen, V., Tolvanen, A., & Tyrväinen, L. (2017). Heterogeneous preferences for recreation-oriented management in commercial forests: the role of citizens' socioeconomic characteristics and recreational profiles. *Journal of Environmental Planning and Management*, *60*(3), 399–418. doi - 10.1080/09640568.2016.1159546

Kajosaari, A., Hasanzadeh, K., Fagerholm, N., Nummi, P., Kuusisto-Hjort, P., & Kyttä, M. (2024). Predicting context-sensitive urban green space quality to support urban green infrastructure planning. *Landscape and Urban Planning*, *242*, 104952. doi - 10.1016/j.landurbplan.2023.104952

Kaplan, R., & Kaplan, S. (1989). *The Experience of Nature: A Psychological Perspective*. CUP Archive.

Karasov, O. (2020). Landscape coherence revisited: GIS-based mapping in relation to scenic values and preferences estimated with geolocated social media data. *Ecological Indicators*, 14.

Kern, C., Klausch, T., & Kreuter, F. (2019). Tree-based Machine Learning Methods for Survey Research. *Survey Research Methods*, *13*(1), 73–93.

Kienast, F., Degenhardt, B., Weilenmann, B., Wäger, Y., & Buchecker, M. (2012). GIS-assisted mapping of landscape suitability for nearby recreation. *Landscape and Urban Planning*, *105*(4), 385–399. doi - 10.1016/j.landurbplan.2012.01.015

Komossa, F., van der Zanden, E. H., Schulp, C. J. E., & Verburg, P. H. (2018). Mapping landscape potential for outdoor recreation using different archetypical recreation user groups in the European Union. *Ecological Indicators*, *85*, 105–116. doi - 10.1016/j.ecolind.2017.10.015

Komossa, F., van der Zanden, E. H., & Verburg, P. H. (2019). Characterizing outdoor recreation user groups: A typology of peri-urban recreationists in the Kromme Rijn area, the Netherlands. *Land Use Policy*, *80*, 246–258. doi - 10.1016/j.landusepol.2018.10.017

Koppen, G., Tveit, M. S., Sang, Å. O., & Dramstad, W. (2014). The challenge of enhancing accessibility to recreational landscapes. *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography*, *68*(3), 145–154. doi - 10.1080/00291951.2014.904399

Korpilo, S., Jalkanen, J., Virtanen, T., & Lehvävirta, S. (2018). Where are the hotspots and coldspots of landscape values, visitor use and biodiversity in an urban forest? *PLOS ONE*, *13*(9), e0203611. doi - 10.1371/journal.pone.0203611

Korpilo, S., Nyberg, E., Vierikko, K., Nieminen, H., Arciniegas, G., & Raymond, C. M. (2023). Developing a Multi-sensory Public Participation GIS (MSPPGIS) method for integrating landscape values and soundscapes of urban green infrastructure. *Landscape and Urban Planning*, *230*, 104617. doi - 10.1016/j.landurbplan.2022.104617

Korpilo, S., Virtanen, T., & Lehvävirta, S. (2017). Smartphone GPS tracking—Inexpensive and efficient data collection on recreational movement. *Landscape and Urban Planning*, *157*, 608–617. doi - 10.1016/j.landurbplan.2016.08.005

Krog, N. H., Engdahl, B., & Tambs, K. (2010). Effects of changed aircraft noise exposure on experiential qualities of outdoor recreational areas. *International Journal of Environmental Research and Public Health*, *7*(10), 3739–3759. doi - 10.3390/ijerph7103739

Laatikainen, T., Piiroinen, R., Lehtinen, E., & Kyttä, M. (2017). PPGIS approach for defining multimodal travel thresholds: Accessibility of popular recreation environments by the water. *Applied Geography*, *79*, 93–102. doi - 10.1016/j.apgeog.2016.12.006

Lagner, O., Klouček, T., & Šímová, P. (2018). Impact of input data (in)accuracy on overestimation of visible area in digital viewshed models. *PeerJ*, *6*, e4835. doi - 10.7717/peerj.4835

Lantmäteriet. (2023). *National Elevation Model*. Retrieved from https://www.lantmateriet.se/en/geodata/geodata-products/product-list/terrain-model-download-grid-1/

Lewicka, M. (2011). Place attachment: How far have we come in the last 40 years? *Journal of Environmental Psychology*, *31*(3), 207–230. doi - 10.1016/j.jenvp.2010.10.001

Li, J., Burroughs, K., Halim, M. F., Penbrooke, T. L., Seekamp, E., & Smith, J. W. (2018). Assessing soundscape preferences and the impact of specific sounds on outdoor recreation activities using qualitative data analysis and immersive virtual environment technology. *Journal of Outdoor Recreation and Tourism*, *24*, 66–73. doi - 10.1016/j.jort.2018.08.001

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

Manley, K., & Egoh, B. N. (2022). Mapping and modeling the impact of climate change on recreational ecosystem services using machine learning and big data. *Environmental Research Letters*, *17*(5), 054025. doi - 10.1088/1748-9326/ac65a3

Mann, C., Pouta, E., Gentin, S., & Jensen, F. S. (2010). Outdoor recreation in forest policy and legislation: A European comparison. *Urban Forestry & Urban Greening*, *9*(4), 303–312. doi - 10.1016/j.ufug.2010.06.004

Manning, R. E. (2022). *Studies in Outdoor Recreation: Search and Research for Satisfaction*. Corvallis: Oregon State University Press. Retrieved from https://muse.jhu.edu/pub/205/monograph/book/1583

Molnar, C., Bischl, B., & Casalicchio, G. (2018). *iml: An R package for Interpretable Machine Learning.*

Morelle, K., Buchecker, M., Kienast, F., & Tobias, S. (2019). Nearby outdoor recreation modelling: An agent-based approach. *Urban Forestry & Urban Greening*, *40*, 286–298. doi - 10.1016/j.ufug.2018.07.007

Morse, W. C., Stern, M., Blahna, D., & Stein, T. (2022). Recreation as a transformative experience: Synthesizing the literature on outdoor recreation and recreation ecosystem services into a systems framework. *Journal of Outdoor Recreation and Tourism*, *38*, 100492. doi - 10.1016/j.jort.2022.100492

Nahuelhual, L., Carmona, A., Lozada, P., Jaramillo, A., & Aguayo, M. (2013). Mapping recreation and ecotourism as a cultural ecosystem service: An application at the local level in Southern Chile. *Applied Geography*, *40*, 71–82. doi - 10.1016/j.apgeog.2012.12.004

Neuvonen, M., Sievänen, T., Tönnes, S., & Koskela, T. (2007). Access to green areas and the frequency of visits – A case study in Helsinki. *Urban Forestry & Urban Greening*, *6*(4), 235–247. doi - 10.1016/j.ufug.2007.05.003

Nguyen, P.-Y., Astell-Burt, T., Rahimi-Ardabili, H., & Feng, X. (2021). Green Space Quality and Health: A Systematic Review. *International Journal of Environmental Research and Public Health*, *18*(21), 11028. doi - 10.3390/ijerph182111028

Nisbet, E. K., Zelenski, J. M., & Murphy, S. A. (2009). The nature relatedness scale: Linking individuals' connection with nature to environmental concern and behavior. *Environment and Behavior*, *41*(5), 715–740. doi - 10.1177/0013916508318748

Nordic Council of Ministers. (2013). *Social Indicators in the Forest Sector in Northern Europe: A review focusing on Nature-based recreation and Tourism*. Nordic Council of Ministers. doi - 10.6027/TN2013-584

Northrup, J. M., Hooten, M. B., Anderson, C. R., & Wittemyer, G. (2013). Practical guidance on characterizing availability in resource selection functions under a use–availability design. *Ecology*, *94*(7), 1456–1463. doi - 10.1890/12-1688.1

Nutsford, D., Reitsma, F., Pearson, A. L., & Kingham, S. (2015). Personalising the viewshed: Visibility analysis from the human perspective. *Applied Geography*, *62*, 1–7. doi - 10.1016/j.apgeog.2015.04.004

Nyelele, C., Keske, C., Chung, M. G., Guo, H., & Egoh, B. N. (2023). Using social media data and machine learning to map recreational ecosystem services. *Ecological Indicators*, *154*, 110606. doi - 10.1016/j.ecolind.2023.110606

Ode Sang, Å., Knez, I., Gunnarsson, B., & Hedblom, M. (2016). The effects of naturalness, gender, and age on how urban green space is perceived and used. *Urban Forestry & Urban Greening*, *18*, 268–276. doi - 10.1016/j.ufug.2016.06.008

Pálsdóttir, A. M., Stigsdotter, U. K., Persson, D., Thorpert, P., & Grahn, P. (2018). The qualities of natural environments that support the rehabilitation process of individuals with stress-related mental disorder in nature-based rehabilitation. *Urban Forestry & Urban Greening*, *29*, 312–321. doi - 10.1016/j.ufug.2017.11.016

Paracchini, M. L., Zulian, G., Kopperoinen, L., Maes, J., Schägner, J. P., Termansen, M., … Bidoglio, G. (2014). Mapping cultural ecosystem services: A framework to assess the potential for outdoor recreation across the EU. *Ecological Indicators*, *45*, 371–385. doi - 10.1016/j.ecolind.2014.04.018

Peña, L., Casado-Arzuaga, I., & Onaindia, M. (2015). Mapping recreation supply and demand using an ecological and a social evaluation approach. *Ecosystem Services*, *13*, 108–118. doi - 10.1016/j.ecoser.2014.12.008

Petersson-Forsberg, L. (2014). Swedish spatial planning: A blunt instrument for the protection of outdoor recreation. *Journal of Outdoor Recreation and Tourism*, *5–6*, 37–47. doi - 10.1016/j.jort.2014.03.003

Pröbstl-Haider, U., Hunt, L. M., Rupf, R., & Haegeli, P. (2020). Choice experiments in outdoor recreation. *Journal of Outdoor Recreation and Tourism*, *32*, 100321. doi - 10.1016/j.jort.2020.100321

Pukkala, T., Lähde, E., & Laiho, O. (2012). Continuous cover forestry in Finland – Recent research results. In T. Pukkala & K. von Gadow (Eds.), *Continuous Cover Forestry* (pp. 85–128). Dordrecht: Springer Netherlands. doi - 10.1007/978-94-007-2202-6_3

Qiu, L., & Nielsen, A. B. (2015). Are Perceived Sensory Dimensions a Reliable Tool for Urban Green Space Assessment and Planning? *Landscape Research*, *40*(7), 834–854. doi - 10.1080/01426397.2015.1029445

Ratcliffe, E. (2021). Sound and Soundscape in Restorative Natural Environments: A Narrative Literature Review. *Frontiers in Psychology*, *12*. Retrieved from https://www.frontiersin.org/articles/10.3389/fpsyg.2021.570563

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. doi - 10.1038/s41586-019-0912-1

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data*

Mining, 1135–1144. New York, NY, USA: Association for Computing Machinery. doi - 10.1145/2939672.2939778

Richards, D. R., & Belcher, R. N. (2020). Global changes in urban vegetation cover. *Remote Sensing*, *12*(1), 23. doi - 10.3390/rs12010023

Ridding, L. E., Redhead, J. W., Oliver, T. H., Schmucki, R., McGinlay, J., Graves, A. R., … Bullock, J. M. (2018). The importance of landscape characteristics for the delivery of cultural ecosystem services. *Journal of Environmental Management*, *206*, 1145–1154. doi - 10.1016/j.jenvman.2017.11.066

Ruzickova, K., Ruzicka, J., & Bitta, J. (2021). A new GIS-compatible methodology for visibility analysis in digital surface models of earth sites. *Geoscience Frontiers*, *12*(4), 101109. doi - 10.1016/j.gsf.2020.11.006

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

Scholte, S. S. K., Daams, M., Farjon, H., Sijtsma, F. J., van Teeffelen, A. J. A., & Verburg, P. H. (2018). Mapping recreation as an ecosystem service: Considering scale, interregional differences and the influence of physical attributes. *Landscape and Urban Planning*, *175*, 149–160. doi - 10.1016/j.landurbplan.2018.03.011

Scott, A., Carter, C., Brown, K., & White, V. (2009). 'Seeing is not everything': Exploring the landscape experiences of different publics. *Landscape Research*, *34*(4), 397–424. doi - 10.1080/01426390903009289

Scowen, M., Athanasiadis, I. N., Bullock, J. M., Eigenbrod, F., & Willcock, S. (2021). The current and future uses of machine learning in ecosystem service research. *Science of The Total Environment*, *799*, 149263. doi - 10.1016/j.scitotenv.2021.149263

Sharpe, D. (2013). Why the resistance to statistical innovations? Bridging the communication gap. *Psychological Methods*, *18*(4), 572–582. doi - 10.1037/a0034177

Silvennoinen, H., Koivula, M., Tikkanen, J., & Tyrväinen, L. (2022). Consistency of photo and field surveys in measuring the attractiveness of forest stands. *Journal of Outdoor Recreation and Tourism*, *40*, 100586. doi - 10.1016/j.jort.2022.100586

Skärbäck, E. (2007). Planning for healthful landscape values. In Ü. Mander, H. Wiggering, & K. Helming (Eds.), *Multifunctional Land Use: Meeting Future Demands for Landscape Goods and Services* (pp. 305–326). Berlin, Heidelberg: Springer. doi - 10.1007/978-3-540-36763-5_19

Soga, M., & Gaston, K. J. (2016). Extinction of experience: the loss of human-nature interactions. *Frontiers in Ecology and the Environment*, *14*(2), 94–101. doi - 10.1002/fee.1225

Stoltz, J., & Grahn, P. (2021). Perceived sensory dimensions: An evidence-based approach to greenspace aesthetics. *Urban Forestry & Urban Greening*, *59*, 126989. doi - 10.1016/j.ufug.2021.126989

Swedish Environmental Protection Agency. (2015). *Friluftsliv 2014: nationell undersökning om svenska folkets friluftslivsvanor* (No. 6691). Stockholm: Swedish Environmental Protection Agency.

Swedish Environmental Protection Agency. (2019a). *Friluftsliv 2018: nationell undersökning av svenska folkets friluftslivsvanor* (No. 6887). Stockholm: Naturvårdsverket. Retrieved from Naturvårdsverket website: https://www.naturvardsverket.se/978-91-620-6887-5

Swedish Environmental Protection Agency. (2019b). *Uppföljning av målen för friluftslivspolitiken 2019* (No. 6904).

Swedish Environmental Protection Agency. (2023). *Uppföljning av målen för friluftslivspolitiken 2023* (No. 7123; p. 117). Retrieved from https://www.naturvardsverket.se/amnesomraden/friluftsliv/sveriges-friluftlivsmal/

Tenerelli, P., Püffel, C., & Luque, S. (2017). Spatial assessment of aesthetic services in a complex mountain region: combining visual landscape properties with crowdsourced geographic information. *Landscape Ecology*, *32*(5), 1097–1115. doi - 10.1007/s10980-017-0498-7

Tiemann, A., & Ring, I. (2022). Towards ecosystem service assessment: Developing biophysical indicators for forest ecosystem services. *Ecological Indicators*, *137*, 108704. doi - 10.1016/j.ecolind.2022.108704

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

van der Wal, R., Miller, D., Irvine, J., Fiorini, S., Amar, A., Yearley, S., … Dandy, N. (2014). The influence of information provision on people's landscape preferences: A case study on understorey vegetation of deer-browsed woodlands. *Landscape and Urban Planning*, *124*, 129–139. doi - 10.1016/j.landurbplan.2014.01.009

Vihervaara, P., Rönkä, M., & Walls, M. (2010). Trends in Ecosystem Service Research: Early Steps and Current Drivers. *AMBIO*, *39*(4), 314–324. doi - 10.1007/s13280-010-0048-x

Walz, U., & Stein, C. (2018). Indicator for a monitoring of Germany's landscape attractiveness. *Ecological Indicators*, *94*, 64–73. doi - 10.1016/j.ecolind.2017.06.052

Wang, S., Wang, M., & Liu, Y. (2021). Access to urban parks: Comparing spatial accessibility measures using three GIS-based approaches. *Computers, Environment and Urban Systems*, *90*, 101713. doi - 10.1016/j.compenvurbsys.2021.101713

Wartmann, F. M., Tieskens, K. F., van Zanten, B. T., & Verburg, P. H. (2019). Exploring tranquillity experienced in landscapes based on social media. *Applied Geography*, *113*, 102112. doi - 10.1016/j.apgeog.2019.102112

Weyland, F., & Laterra, P. (2014). Recreation potential assessment at large spatial scales: A method based in the ecosystem services approach and landscape metrics. *Ecological Indicators*, *39*, 34–43. doi - 10.1016/j.ecolind.2013.11.023

White, M., Smith, A., Humphryes, K., Pahl, S., Snelling, D., & Depledge, M. (2010). Blue space: The importance of water for preference, affect, and restorativeness ratings of natural and built scenes. *Journal of Environmental Psychology*, *30*(4), 482–493. doi - 10.1016/j.jenvp.2010.04.004

WHO. (2015). *A typology of urban green spaces, eco-system services provisioning services and demands* (No. D3, 1).

Wolff, M. (2021). Taking one step further – Advancing the measurement of green and blue area accessibility using spatial network analysis. *Ecological Indicators*, *126*, 107665. doi - 10.1016/j.ecolind.2021.107665

Yoshimura, N., & Hiura, T. (2017). Demand and supply of cultural ecosystem services: Use of geotagged photos to map the aesthetic value of landscapes in Hokkaido. *Ecosystem Services*, *24*, 68–78. doi - 10.1016/j.ecoser.2017.02.009

Zube, E. H., Sell, J. L., & Taylor, J. G. (1982). Landscape perception: Research, application and theory. *Landscape Planning*, *9*(1), 1–33. doi - 10.1016/0304-3924(82)90009-0

Popular science summary

Nature experiences are important for human well-being. Despite this, we see a decline in outdoor recreation, which some researchers argue risks leading people to lose their connection with nature. This could create a vicious cycle where decreased nature contact leads to less outdoor recreation, further reducing nature contact, and so on. The exact reasons for the decline are not fully understood, but one suggested cause is the decreased availability of natural areas suitable for outdoor recreation. To address this problem, we need methods to determine what constitutes a suitable outdoor recreation area, which requires an understanding of the characteristics that make areas attractive to people.

Such knowledge can be obtained in various ways: for example, by asking people to rate photographs of different landscapes. Another approach is to observe where people actually go for outdoor recreation and draw conclusions from that. Today, there are many ways to collect such information about people's movements. For instance, mobile phones can be tracked using cell tower triangulation, or photos from social media can be collected, which often contain geographical information. An increasingly popular method used by researchers is Public Participatory Geographic Information Systems (PPGIS), where surveys are used to gather spatial information from people. However, drawing conclusions about which characteristics make people choose one location over another based on movement patterns has its challenges. For instance, accessibility plays a significant role in outdoor recreation, with closer areas being visited more frequently simply because they are closer. Such effects can make it difficult to draw accurate conclusions: are people visiting a certain area because of its characteristics, or simply because it is close to where they live?

This thesis examines the characteristics that make certain areas more attractive for recreationists and how such characteristics can be used to assess which areas are suitable for outdoor recreation. This is done through two methods: first, a literature review compiling research on which forest characteristics recreationists prefer. Second, outdoor recreation patterns among Swedes are collected in two different PPGIS studies. The PPGIS studies develop a new methodology, where people's choices of outdoor recreation locations are compared with the landscape available to them using machine learning and viewshed analysis.

In the literature review of the characteristics that make forests attractive for outdoor recreation, we found that large trees, proximity to water, and the absence of signs of forestry are particularly important. The PPGIS studies also showed that people are strongly attracted to water environments, but also that outdoor recreation facilities such as shelters, fire pits, and trails greatly influence which locations are visited. Noise was also found to have an effect, with people avoiding areas of urban noise. The thesis then discusses how these characteristics can be combined into an index to estimate which areas are attractive for recreation.

Populärvetenskaplig sammanfattning

Naturupplevelser är oerhört viktiga för människors välmående. Trots detta ser vi att friluftsliv minskar, något som vissa forskare menar riskerar att leda till att folk får en försämrad kontakt med naturen. Detta skulle kunna leda till en ond cirkel, där sämre naturkontakt leder till mindre friluftsliv, som i sin tur leder till sämre naturkontakt, och så vidare. Det är inte helt klarlagt varför friluftslivet minskar, men en anledning som föreslagits är att människors tillgång till naturområden lämpliga för friluftsliv har minskat. För att lösa detta problem behöver vi ha metoder för att kunna avgöra hur ett område lämpligt för friluftsliv ser ut, något som kräver en förståelse för vilka egenskaper som gör områden attraktiva för människor.

Sådan kunskap kan inhämtas på olika vis. Ett sätt är att be folk betygsätta fotografier av olika typer av landskap. Ett annat sätt är att titta på var folk faktiskt beger sig för att utöva friluftsliv, och dra slutsatser utifrån vilka platser som är populära. Idag finns det många sätt att samla in sådan information om människors rörelser. Exempelvis kan man spåra mobiltelefoner med hjälp av triangulering, eller samla in foton från sociala medier, som ofta innehåller information om var de är tagna någonstans. En metod som forskare börjat använda sig allt mer av är det som kommit att kallas Public Participatory Geographic Information Systems (PPGIS), där man använder enkäter för att samla in rumslig information från människor. Att försöka dra slutsatser om vilka egenskaper som gör att folk väljer en viss plats över en annan utifrån rörelsemönster har dock sina svårigheter: t.ex. så har det visats att tillgänglighet spelar en stor roll för friluftsliv, där mer närbelägna områden besöks i mycket högre utsträckning, helt enkelt för att de ligger närmre. En sådan effekt kan göra det svårare att dra rätt slutsatser: är det på grund av hur ett visst område ser ut som människor är där, eller bara för att det ligger nära där de bor?

Den här avhandlingen tittar på vilka egenskaper som gör vissa områden mer attraktiva för friluftslivsutövare, och hur sådana egenskaper kan användas för att bedöma vilka områden som är lämpliga för friluftsliv. Detta görs genom två metoder: först genom en litteraturgenomgång, där forskning om vilka skogliga egenskaper som människor föredrar sammanställs. Sedan samlas rumsliga data om svenskars friluftslivsvanor in i två olika PPGIS-studier. Det insamlade data analyseras med en ny metodik, där människors val av plats för friluftsliv jämförs med landskapet de hade i sin närhet med hjälp av maskininlärning och siktfältsanalys.

I litteraturgenomgången av vilka egenskaper som gör skogar attraktiva för friluftsliv fastslår vi att stora träd, närheten till vattendrag och avsaknad av spår efter skogsbruk pekas ut som särskilt viktiga för friluftsliv. PPGIS-studierna visade också att människor dras starkt till vattenmiljöer, men också att friluftslivsanordningar som vindskydd, grillplatser och stigar gör mycket för att avgöra vilka platser som besöks. Buller visade sig också påverka, där människor sökte sig bort från stadens oljud. I avhandlingen diskuteras sedan hur dessa egenskaper tillsammans kan kombineras i formen av ett index för att försöka skatta vilka områden som är attraktiva för rekreation.

Acknowledgements

They said writing the acknowledgments was hard.
They weren't wrong.
It's been a wild couple of years. During the 4,5 years it took to excrete this lump of paper I've managed to grow a beard, develop a crippling Reddit addiction, and have a baby. What a ride! However, I did not achieve all this on my own.[1] As they say, it takes a village to write a thesis.

First, my supervisors: **Thomas**, I have many things to thank you for, but what I've really come to appreciate is your keen eye for cutting through my sloppy writing, and actually making me think about what I'm trying to express. To the rest of the supervisor team, **Anna**, **Marcus**, **Erik**: Thank you for all your support, in all the myriad ways. I've really enjoyed working with all of you.

Jag vill tacka min familj. **Mamma**, du har alltid varit en förespråkare av kunskap och det livslånga lärandet. Det blev ingen läkare av mig, men väl en doktor, och det tror jag du hade en stor roll i. **Pappa**, din flit och din stoicism är en inspiration. Du har lärt mig vikten av strävan.
Johan, **Erik, Arianna**—Det är en sådan trygghet att ha er. Tack för allt.
Bror & **Axel**—Att leka med er har varit det bästa sättet att släppa stressen med den här avhandlingen.
Edith—Du är den bästa bäbisen. Tack för att du höll dig till ett någorlunda rimligt sovschema under tiden som den här avhandlingen skrevs. Nu får du kaosa loss bäst du vill.
Sophia—Älskade, jag har så många smetiga ord till dig och saker att tacka dig för, men de tänker jag inte dela med allt löst folk som läser det här. De orden är endast för dig.

---

[1] Well, except for the beard I guess.

Anna—Tack för alla gånger du varit där när bägaren runnit över. You're the Greg to my Tom.
Staffan—Våra Puyo-skills må rosta, men vänskapen består.
Nick—Q: How did the man learn he was allergic to bees?[2]
Johnny—Thank you for your great care of my second baby, the Ecology Pub. May your hops be ever dank.
Kristina—Det har varit ett rent nöje att tvångsmata dig med förskingrade bullar över åren.
Dragos—We still need to play that game of Starcraft.
Lovisa—Tack för alla gånger du släpat ut mig i solen och regnet och slasket. Snart blir det mera vild dans!
Juliana—Kom tillbaka till ekologen, det är för långt till SoL!
Ineta—Your energies are infectious, in the best of ways.
Guille—Whenever I see your face, I just get this feeling that everything is going to be okay.
Coffee machine #2—Thank you for dispensing 2483 cups of mostly acceptable coffee.
To my people in the **Forest faculty PhD council**—Keep on fighting the good fight. One day Göran will let us have our traktamente.

If your name is not in here, I forgot you. I'm sorry. It's probably not because you're not important to me.[3] Let's just say the thesis gnomes ate your name. If you still feel bad about it, you can fill in your own name and motivation below.
__________—A sincere thanks for ______________________________ ________________. I fondly remember when we _________________. This wouldn't have been possible without you.

[2] A: When he broke into hives!

[3] Although, well, it might be.

II

Landscape and Urban Planning 227 (2022) 104519

Contents lists available at ScienceDirect

Landscape and Urban Planning

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

Landscape usage by recreationists is shaped by availability: Insights from a national PPGIS survey in Sweden

Carl Lehto [a,*], Marcus Hedblom [b], Erik Öckinger [a], Thomas Ranius [a]

[a] Swedish University of Agricultural Sciences, Department of Ecology, SLU Instutitionen för ekologi, Ulls väg 16, 756 51 Uppsala, Sweden
[b] Swedish University of Agricultural Sciences, Department of Urban and Rural Development, Sweden

HIGHLIGHTS

- Outdoor recreation is highly geographically aggregated to urban and periurban areas.
- Selection of landscape type follow availability - recreationists use what is nearby.
- Longer than preferable travel distances suggest possible recreational deficit.
- Landscape characteristics weak predictor of where recreation is conducted.

ARTICLE INFO

Keywords:
Landscape heterogeneity
Machine learning
Public Participatory GIS
Revealed preferences
Recreation
Travelling distance

ABSTRACT

Areas suitable for outdoor recreation are in decline due to urbanization and land-use intensification. To provide people with access to recreational areas, it is imperative to understand what characterizes areas attractive to recreationists. In this study we explore patterns of outdoor recreation visits on a national scale, using a large (n = 3853) Public Participatory GIS survey in Sweden. We analyze land cover of areas visited in comparison to landscape composition across a gradient from urban to rural areas. Additionally, we employ machine learning models to compare attributes of areas visited to random areas in the available landscape. We found that the geographical distribution of outdoor recreation was highly aggregated, with 57 % of recreation occurring in urban and periurban areas, which together cover 5 % of the total land area. Landscape characteristics were weak predictors of where outdoor recreation took place. The median travel distance to the area where recreation was conducted was 2 km, which is longer than what recreationists prefer according to previous studies. We argue that this is indicative of a recreational deficit in Sweden, with recreationists' preferences not being expressed due to lack of access to suitable areas close to home. This highlights the importance for physical planners to consider spatial accessibility when planning for outdoor recreation.

1. Introduction

Urbanization and intensified land use has led to a decrease in the supply of outdoor recreation opportunities globally (Hedblom, Andersson, & Borgström, 2017; IPBES, 2019). Recreational opportunities are difficult to plan for, since they encompass a wide range of activities, each with different demands on the landscape (Juutinen, Kosenius, Ovaskainen, Tolvanen, & Tyrväinen, 2017). To be able to provide recreational opportunities, it is important to understand what factors are most important in shaping recreational usage of landscapes. Previous studies on recreational preference have mostly relied on stated preference, i.e. the outcome when asking people to rate pictures or other descriptions of real or hypothetical landscapes. These studies have found effects of e.g. forest types (Gundersen & Frivold, 2008), biodiversity levels (Qiu, Lindberg, & Nielsen, 2013) and landscape heterogeneity (Filyushkina, Agimass, Lundhede, Strange, & Jacobsen, 2017). Preferences have been shown to vary between individuals, influenced by e.g. socio-demographic factors (van Zanten, Verburg, Koetse, & van Beukering, 2014), held beliefs and attitudes (Kearney & Bradley, 2011), cultural differences (Gosal et al., 2021), and group identity (Scott, Carter, Brown, & White, 2009). The outcomes have been used in other studies to make spatial predictions of where high recreational values are located (e.g. Norton, Inwood, Crowe, & Baker, 2012; Komossa, van der Zanden, Schulp, & Verburg, 2018).

* Corresponding author.
E-mail address: Carl.lehto@slu.se (C. Lehto).

https://doi.org/10.1016/j.landurbplan.2022.104519
Received 17 January 2022; Received in revised form 30 June 2022; Accepted 23 July 2022
Available online 30 July 2022

Fewer studies (e.g. Agimass, Lundhede, Panduro, & Jacobsen, 2018; De Valck et al., 2017; Kienast, Degenhardt, Weilenmann, Wäger, & Buchecker, 2012) have assessed revealed preference at the landscape level, i.e. the patterns of actual landscape use by recreationists. Observing actual usage is a more difficult undertaking, but an important complement to stated preference studies in order to see how held preferences are realized in recreational patterns. With the advent of new technologies it has become easier to obtain large amounts of spatial data on recreation, with researchers using approaches such as GPS tracking (Korpilo, Virtanen, & Lehvävirta, 2017), data scraping of social media (Yoshimura & Hiura, 2017), or online surveys with elements of public participatory GIS (PPGIS) (Brown & Fagerholm, 2015).

Using PPGIS, outdoor recreation has been analyzed with different research questions and on various, but mostly local, scales. Examples include estimating the usage of a city park (Korpilo, Virtanen, Saukkonen, & Lehvävirta, 2018), analyzing access to aquatic environments (Laatikainen, Tenkanen, Kyttä, & Toivonen, 2015), modeling recreation around small Swiss towns (Kienast et al., 2012), or examining how residents of a large city (Antwerp) utilize the surrounding region (De Valck et al., 2016). The local scale is relevant when studying patterns of recreation; for instance, Kienast et al. (2012) revealed significant model differences between each town they studied, suggesting that local conditions make these patterns unique for every town. Large-scale studies are still mainly lacking, but could potentially reveal broader patterns of outdoor recreation, i.e. how recreationists in general utilize landscapes available to them. This could inform higher-level public policy and help in the prioritization of land use for recreation, even in areas where no local-scale studies have been conducted.

The preferences for recreation is not only a question of what is required of the landscape, but also of where it should be located. Recreation has been shown to be highly influenced by geographic accessibility: recreational areas are visited more often the closer they are to peoples' home (Neuvonen, Sievänen, Tönnes, & Koskela, 2007; SEPA, 2019), and most people prefer shorter distances than they currently have to their closest recreational forest (Hörnsten & Fredman, 2000). The interplay between distance and landscape characteristics have been explored in several willingness-to-pay studies, often analyzing how far recreationists are willing to travel to experience areas with certain characteristics (Ezebilo, Boman, Mattsson, Lindhagen, & Mbongo, 2015; Giergiczny, Czajkowski, Żylicz, & Angelstam, 2015). Again, as with studies on landscape characteristics, studies on travel distances have mainly relied on asking recreationists about preferences (e.g. Hörnsten & Fredman, 2000) and have not assessed actual movement patterns of recreationists. Further, research on the availability of recreational areas has primarily focused on urban and periurban areas, with few studies including rural areas. This is possibly due to an assumption that rural areas have a higher availability of natural areas suitable for recreation. This assumption might be unfounded, with high-intensity agriculture and silviculture dominating many landscapes.

In this study, we attempt to fill in some of the above-mentioned knowledge gaps by analyzing a large-scale PPGIS survey on recreational habits across Sweden. We employ a novel approach, combining the ability of PPGIS to yield a large amount of spatial data with the flexibility of machine learning in the form of Boosted Regression Trees (BRT). BRT modeling is particularly useful when faced with large amounts of data and many possible predictors, yielding high predictive power paired with simple model selection (Elith, Leathwick, & Hastie, 2008), but has to our knowledge not been used in the research field of recreation. Using biophysical landscape characteristics (such as land cover, heterogeneity, and topology), combined with socio-demographic attributes (such as age, gender, and degree of education) and characteristics of recreational visits (season, duration and type of activity), we explore what factors affect the choice of location for recreation on a national scale. We investigate the difference in landscape composition and frequency of recreation on a gradient from rural areas to urban areas. We also investigate travel distances to recreational areas. More specifically, we aim to answer the following questions:

I. Does the availability and recreationists' selection for different land cover types vary between urban and rural areas?
II. How do typical travel distances from home to recreational areas vary in relation with type of outdoor recreation and time of the year?
III. What biophysical characteristics of landscapes (i.e. land cover, heterogeneity, topology, path and road density, forest characteristics, and protected areas) are most important in shaping where outdoor recreation is conducted, and how do these effects depend on individual attributes of the recreationist (i.e socio-demographics or type of recreation performed)?

2. Materials and methods

2.1. Study area

This study is based on survey data on outdoor recreational habits collected across Sweden. In Sweden, most outdoor recreation is performed in natural or semi-natural environments, with 50 % of residents spending time in nature on weekdays. The most common activities are walking, spending time in forests, and cycling (SEPA, 2019). Recreationists often utilize the right of public access, which allows access to almost all property except arable land to anyone, and is enshrined in law. Sweden is a relatively sparsely populated country (25.5 inhabitants/km^2). It has followed a similar trajectory to the rest of the world of increased urbanization, with 87 % living in urban areas as defined by the Swedish Bureau of Statistics (SCB, 2018), i.e. sites with more than 200 inhabitants with a distance of < 200 m from their closest neighbouring house within the urban area. In addition to being clustered around urban areas, the population density is also skewed towards the south of Sweden, and towards the coasts.

2.2. Survey design

The data used in this study was collected as part of a national survey on Swedish residents' recreational habits (SEPA, 2015). Twelve digital panel surveys were performed, totaling 8410 responses during the period of December 2013 – November 2014. Each survey was initially sent to 340 people drawn from a panel of 80 000 each day over the first week of every month. The sample was stratified to be representative of the Swedish population in regard to age, sex and region of Sweden. Extra invitations to participate in the survey were sent as needed during each month, weighted on the response rate of the group quotas. Participants were anonymous and able to participate in multiple months, but not more than one time each month. Respondent IDs were lost during data handling, thus the number of unique respondents is unknown; however the large panel size made repeat participants presumably rare.

The survey tasked the respondents with marking the location of their latest outdoor recreational visit on a map, and to provide details of the visit, such as the time spent, the distance from home to the location, how often they visit this location, and the type of activity. If they had visited a larger area they were instructed to mark the center point of the area. In the survey, outdoor recreation was defined as "any activity performed outdoors in a natural or cultural landscape for the purpose of well-being and experiencing nature". This broad definition encompassed almost all kinds of outdoor activities, which was reflected in the extensive list of activities the respondents could choose between (Supplementary materials S1). Simple activities such as walking, jogging or cycling made up the majority of responses, while more complicated activities such as roller-skating or horse riding were rarer. For our analysis, we chose to exclude responses where the performed activity restricted the ability to choose freely where to conduct the activity, such as alpine skiing, gardening, or golf.

Due to programming errors in the survey website the first two months (December and January) of the spatial data were lost. Additionally, not all survey participants chose to mark a location. The dataset was further reduced by removing 12 visits outside of Sweden; 482 visits longer than 24 h (to distinguish outdoor recreation from tourism (Bell, Tyrväinen, Sievänen, Pröbstl, & Simpson, 2007)); and 51 visits where the respondent had indicated it was their first visit to the location (i.e. this was not a location preferred due to experiences from earlier visits). The final sample size was 3853 (Fig. 1). The reduction in sample size did not lead to a significant geographical skew, moving the mean center of the dataset only 6.3 km southwest. Gender and age distributions were similar to national demographics in 2014 (Supplementary materials S2).

2.3. Availability and selection of land cover types across the urban–rural gradient

To evaluate availability and selection of different land cover types along an urban–rural gradient, we divided the dataset into four categories. The first category consisted of recreational visits within urban areas, using the definition of the Central Bureau for Statistics (SCB, 2018): any area of at least 200 residents with < 200 m to their closest neighbour. The second category consisted of visits within periurban areas, where periurban was defined using the definition by the National Forest Inventory of Sweden (2009): a buffer around each urban area (200–7500 m) with an increasing radius with increasing population size of the urban area. The third category consisted of all visits outside periurban areas but <10 km from any urban area, and the fourth category consisted of all visits more than 10 km from any urban area. Land cover data was extracted using the high-resolution satellite-based CadasterENV raster (Swedish Environmental Protection Agency, 2018). The 25 land cover classes were reclassified into 13 classes to aggregate classes that we believed were similar in recreational aspects (Supplementary materials S3). In each of the four urban–rural categories, the land cover composition was compared to the land cover at the locations of the recreational visits using the Manly-Chesson selection index. The index is calculated by dividing the fraction of each used land cover with the fraction of available land cover, to see which land cover types are selected for or against (Manly, McDonald, Thomas, McDonald, & Erickson, 2002).

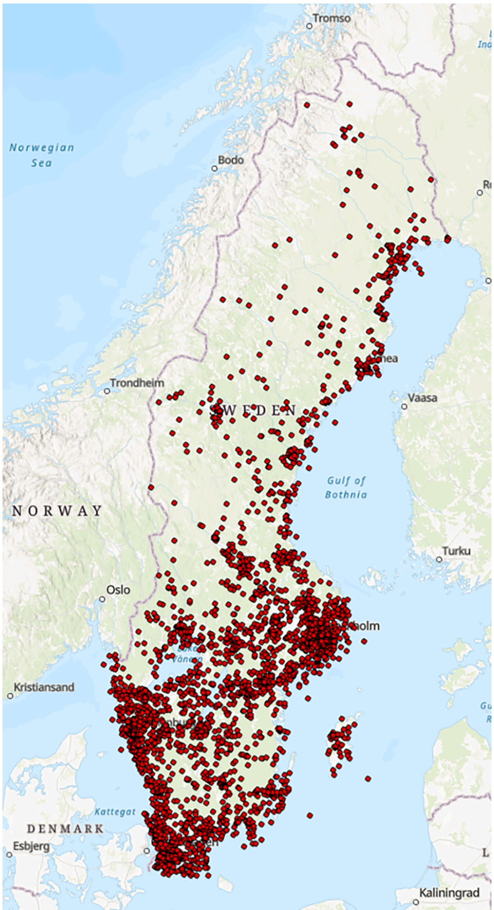


Fig. 1. Distribution of the recreational visits across Sweden.

2.4. Predictive modeling of outdoor recreation

2.4.1. Use-available framework

To explore how landscape characteristics and individual attributes of recreationists affect where outdoor recreation was conducted, we applied a use-available framework. This is a common approach in studies of animal habitat selection, where spatial data on the movements of animals (the use sample) is contrasted with locations drawn randomly from the surrounding landscape (the availability sample) (Northrup et al., 2013). Here, our use sample consisted of a single point from each survey respondent representing their last recreational visit. The availability sample was placed at a point randomly within twice the travel distance of each visited point. This approach was used since the starting point of the travel (the respondents' home) was unknown.

Since the exact extent of the area the respondent had experienced was unknown, we created five different models sampling landscape characteristics on different spatial scales. The first model only used the point given by the recreationist, while the second employed a circular buffer of 100 m around each point, reflecting an assumed minimum area the recreationists had experienced. The final three models employed a buffer with a varying radius, with the radius increasing with increasing time spent during the recreational visit. The buffer radius was constrained to reach its maximum at 120 min time spent on location, and the maximum radius was set to roughly yield a tripling of the area compared to the previous model (200 m, 340 m, and 570 m respectively for model 3–5). The reason for constraining the buffer radius in this way was due to the many outliers in regards to visit duration, which would have yielded unreasonably large buffers; 120 min corresponded to the third quantile of respondents.

2.4.2. Predictors

We used all sources of map data we believed could affect outdoor recreation, and that covered our entire study area (Table 1). We extracted map data within the buffers (or, at the point for model 1) using ArcGIS Pro 10.7. Land cover data was extracted from the CadasterENV satellite-based raster (Swedish Environmental Protection Agency, 2018). We reclassified the 25 land cover classes into 13 (Supplementary materials S3) to simplify model interpretation. Land cover was used both as a predictor by itself (fraction of each land cover class within the buffer), but was also used to estimate landscape heterogeneity (see *Estimation of landscape heterogeneity*). Forest data was included by extracting tree height and volumes of different tree species from the SLU forest map (SLU, 2015). Elevation was extracted from the Swedish National Land Survey ground topology map (Lantmäteriet, n.d.) using both the median and the standard deviation of height above sea level within the buffers as separate predictors. As a proxy for biodiversity, we calculated the overlap of protected areas (National parks, nature

Table 1
Predictors related to biophysical landscape characteristics included in the machine learning models.

Predictor	Description	Levels
Land cover (13 predictors)	Percentage cover of each land cover type. Data source CadasterENV raster.	% [0, 100]
Q index	Landscape heterogeneity defined via CadasterENV land cover classes, see *Estimation of landscape heterogeneity*.	Continuous [0, 1]
Shannon-Weaver index	Landscape heterogeneity defined via CadasterENV land cover classes	Continuous [0, 1.85]
Recreational quality dimensions (6 predictors)	The mean value of each recreational quality dimension, see *Estimation of landscape heterogeneity* and supplementary materials S4.	Continuous [0, 10]
Tree height	Average height of trees (m). Data source SLU Forest map.	Continuous [0, 23]
Spruce volume	Average volume of Norway spruce per square meter (m^3/ha). Data source SLU Forest map.	Continuous [0, 232]
Pine volume	Average volume of Scots pine per square meter (m^3/ha). Data source SLU Forest map.	Continuous [0, 167]
Deciduous tree volume	Average volume of deciduous trees (m^3/ha). Data source SLU Forest map.	Continuous [0, 41]
Total biomass volume	Average volume of all biomass (m^3/ha). Data source SLU Forest map.	Continuous [0, 196]
Elevation (2 predictors)	Median and standard deviation of elevation within buffer. Data source Swedish Geological Survey elevation map.	Continuous [0,1461]/[0, 61]
Path/road density (2 predictors)	Length of paths/roads per square meter (m/m^2). Data source Lantmäteriet path and road maps.	Continuous [0, 0.1]/[0, 0.2]
Protected area	Percentage cover of legally protected areas (Nature reserves, national parks, protected biotopes).	% [0, 100]

reserves, protected biotopes and Natura 2000 area). Using OpenStreetMap data (OpenStreetMap Foundation, n.d.), we calculated path and road density.

Predictors related to socio-demographic characteristics of the respondent, along with the type of recreational activity they were engaged in, the visit duration, and the season of the visit were also included in the models (Table 2).

2.4.3. Estimation of landscape heterogeneity

Since landscape heterogeneity has been shown to influence recreational preference (Filyushkina et al., 2017), we included it as a predictor using land cover classes. We calculated heterogeneity with two methods: firstly the Shannon-Weaver index, which has been commonly used to estimate landscape heterogeneity previously. Secondly we employed the index Q, which in contrast to Shannon-Weaver accounts for heterogeneity not reflecting only the proportion and number of classes within a patch, but also of the qualitative differences between different classes (Díaz-Varela, Roces-Díaz, & Álvarez-Álvarez, 2016). The advantage of Q is that it can better estimate perceived heterogeneity when certain land cover classes are more similar to each other (e.g. different forest classes) while others classes are more distinct (e.g. sea, alpine). For a description of how we estimated Q, see Supplementary materials S4.

2.4.4. Boosted regression trees

Statistical modeling was performed using boosted regression trees (BRT). BRT is a decision tree-based machine learning approach where a predictive model is created by iteratively building an ensemble of many decision trees, each with a low weight (Friedman, 2001). The method has several advantages over traditional regression methods such as GLMs or GAMs: it does not assume linear relationships between predictor variables and response variables, and can handle a large number

Table 2
Predictors related to the recreationist and the recreational visit included in the machine learning models.

Predictor	Description	Levels
Education	The highest education level obtained by respondent.	Primary education Secondary education Bachelor's degree or equivalent Master's degree or equivalent
Income	Gross household income per year	0–100000 SEK 100001–200000 SEK […] 1,000,001 SEK or more
Rural or Urban*	Whether the respondent lived in an urban or rural area.	Stockholm City with at least 100 000 inhabitants City with 50000–99999 inhabitants Town with 5000–49999 inhabitants Rural area
Boreal or Boreo-Nemoral + Nemoral	Whether the respondent lived north or south of the boreo-nemoral boundary, as defined by Rydin et al. (1999) (roughly equivalent to latitude 60° N in Sweden).	Boreal Boreo-Nemoral + Nemoral
Gender	The gender of the respondent.	Male Female
Ageclass	The age of the respondent, divided into four classes of equal number of respondents.	16–35 years 36–50 years 51–66 years 67–84 years
Disability	Whether the recreationist experienced themselves to have a disability that decreased their ability to conduct outdoor recreation to any degree	Yes No
Immigrant (3 predictors)	Whether the recreationist or either of their parents were born in Sweden.	Sweden Nordic country except Sweden Europe Rest of world
Activity	Type of recreational activity performed.	35 different activities (Table 1)
Season	Season of the year, defined by calendar month.	Spring (March, April, May) Summer (June, July, August) Autumn (September, October, November) Winter (December, January, February)
Visit duration	How long the recreationist spent during the visit	Discrete [5,1440] min

*N.B. that this definition of urban and rural is different from the definition used for the analysis described under *Availability and selection of land cover types across the urban–rural gradient.*

of predictors regardless of multicollinearity. Further, there is no need for model selection or specifying interaction effects in advance, while at the same time yielding models with high predictive power. The main disadvantage of BRT is the lower interpretability of the final models. However, with recent methodological advances, such as the Interpretable Machine Learning package for R (Molnar, 2018), these shortcomings can be overcome to a large degree.

All analysis and visualization was carried out using the gbm package (Greenwell, Boehmke, & Cunningham, 2020) in R version 4.0.3 (R Core Team, 2020). Boosted regression trees were constructed following the recommendations outlined by Elith et al. (2008) using a Bernoulli distribution with used area/available area as the response variable. When fitting boosted regression trees, three hyperparameters that affect model fitting are set: Tree complexity (how many splits are allowed in each tree); learning rate (how quickly the algorithm converges, with lower values leading to better models at the cost of computing time); and bag fraction (how large a fraction of the dataset to use in each iteration). We created models with combinations of four different tree complexities (1, 3, 5 and 7) and two bag fractions (0.5 and 0.75) and lowered the learning rate until a model of at least 1000 trees were fitted. Model performance was evaluated using cross-validated AUC. Feature importance, interactions and partial dependence plots were produced using the iml package (Molnar, 2018).

In total, the models were fitted with 44 predictors, except for the point model for which heterogeneity and averages of quality dimensions could not be assessed, and land cover was used as a single categorical predictor instead of 13 continuous predictors.

3. Results

3.1. The urban–rural gradient of outdoor recreation

57 % of the respondents most recent recreational visits were in urban and periurban areas (Table 3). Landscape composition changed along the urban–rural gradient. Urban areas had a higher proportion of built-up area and open area with vegetation, while areas more than 10 km from an urban area had a higher proportion of sea (Fig. 2). In total, across the whole dataset, 44 % of recreational visits were in forested land cover types, 13 % in built-up areas, 12 % in bodies of water, 11 % in arable land and 4 % in wetlands. Overall, the proportion of visits among land cover classes was strongly correlated with available land cover (Pearson's R-value = 0.94), suggesting that across the four urban–rural categories, selection for different land cover types was weak. However, for some land cover types there was a clear difference between the proportion of visits and the availability. The strongest selection (defined using the Manly-Chesson index: the proportion of recreational visits within each land cover class divided by that land cover class' proportion of the total area) was for arable land and built-up areas in rural areas, followed by freshwater and sea in urban areas (Fig. 2). Among forest types, temperate deciduous forests were most selected for, followed by deciduous forests, mixed forests and pine forests. Spruce forests and clearcuts were selected against.

3.2. Travel distance

The overall median distance from the respondents' home to the recreational area was 2 km, but the distance varied depending on the type of activity (Fig. 3). The median distance was longest for swimming and berry/mushroom picking, and shortest for walking and jogging. There were no significant differences in distances between men and women or any other socio-demographic characteristic. Travel distances varied over the year, with longer distances during summer and shorter during winter (median distance for July 3 km, for December 500 m; see Supplementary materials S5). Travel distances were positively correlated with duration of the recreational visit (linear regression, $p < 0.001$, $r^2 = 0.18$).

Table 3
The distribution of recreational visits and land area across four categories, representing a gradient from urban areas to rural areas. For definitions of the four categories, see Survey data.

Urban-rural category	Recreational visits (% of total)	Land area (% of Sweden)
Urban	27	1.4
Periurban	30	3.6
<10 km	36	48
>10 km	5.8	47

3.3. Predicting outdoor recreation using landscape characteristics

The BRT models performed poorly, with cross-validated AUC scores of 0.55 for the model only sampling the points (model 1) and 0.58–0.6 for the models sampling a buffer around the points (models 2–5, Supplementary materials S6). This suggests all models were only slightly better than chance (corresponding to AUC = 0.5) at distinguishing between the use sample and the availability sample. Which predictors had the largest effect on the outcome of each model was evaluated by calculating the relative influence of each predictor. Model 1 had only one influential predictor, land cover, which had a 95 % influence on model accuracy. Model 2–5 exhibited similar patterns to each other, with type of activity being most influential (18–19 %), followed by the same 7–8 predictors, each with low influence (Supplementary materials S7).

The relationship between each predictor and the probability that an area was selected for recreation was investigated through partial dependence plots. These evaluate the effect of a predictor by setting all other predictors to their median value, and examining how model outcomes change as the predictor of interest changes. Model 1′s only influential predictor, land cover, showed that open areas without vegetation, built-up areas, temperate deciduous forests, and deciduous forests increased the probability that an area would be selected for recreation the most, while the presence of sea, arable land, wetlands, freshwater, or clearcuts lowered the probability (Fig. 4). Note that both terrestrial and water habitats were included in these analyses, and thus the low probability for sea and freshwater simply reflect that most recreation activities in the dataset took place on land.

The partial dependence plots for the influential predictors of model 2–5 revealed almost flat responses, suggesting that model predictions were based on many weak effects (Supplementary materials S8). Interaction effects between predictors were analysed by calculating H-statistics. Model 1 lacked interaction effects due to its tree complexity being 1. The strongest interaction found in model 2–5 was between type of activity and other predictors, and accounted for 15–19 % of the variance of the model prediction. Investigating these interactions yielded no interpretable effects due to the weak main effects of the predictors. There were no clear interactions between landscape predictors and socio-demographic predictors.

4. Discussion

We found that recreation in Sweden was highly aggregated geographically, with 57 % of the recreation occurring in urban or periurban areas, despite these areas only constituting 5 % of the total land area. The median distance from the respondent's home to the site of recreation was 2 km, with the distances varying depending on activity. Further, there was a high correlation between the land cover types that were used and the availability of these types, indicating low levels of selection for most land cover types. Our predictive models had low accuracy, suggesting that the included predictors (land cover, heterogeneity, topology, path and road density, forest characteristics and protected areas) were not important for why an area was chosen for recreation.

4.1. Availability and utilization of land cover types across the urban–rural gradient

The utilization of land cover types for recreation was highly correlated with the land cover composition across the four urban–rural categories. This suggests that there was overall only a weak selection for land cover types, with most being used proportionally to their frequency within each urban–rural category. Selection was only observed for certain land cover types, with temperate deciduous forests and

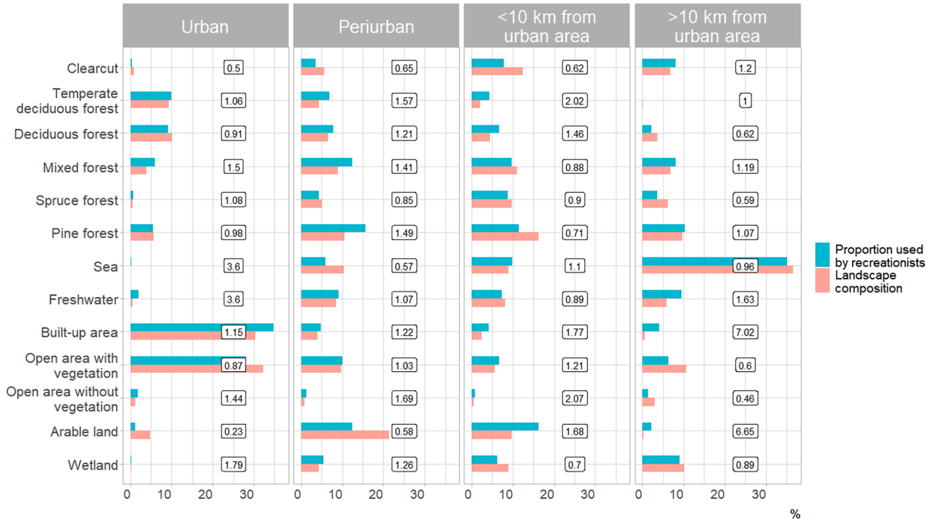


Fig. 2. Landscape composition compared to the proportion of visits by outdoor recreationists of land cover types along the urban–rural gradient. By dividing the used proportion with the landscape composition the Manly-Chesson selection index is calculated, which is presented in the boxes. A value greater than 1 implies selection for the land cover type, a value < 1 selection against.

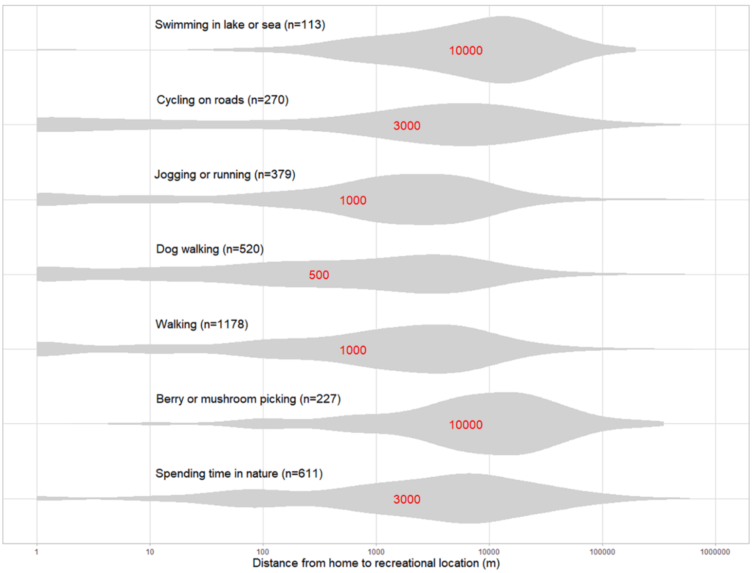


Fig. 3. The distribution of distances from home to the location where recreation occurred for the most common types of recreational activities. The area of each violin is equal. Numbers in red are median distance values; numbers within parentheses are number of responses.

deciduous forests used most, followed by mixed forests and pine forests, while spruce forests and clearcuts were selected against. These results are in agreement with previous stated preference studies of forest types (Gundersen & Frivold, 2008). Water environments in urban areas showed high levels of selection (i.e. the respondent selected a point situated in the water), despite their importance likely being underestimated in this analysis. This is due to recreation occurring close to water being counted as terrestrial, even though the purpose of the visit might have been to experience the water environment, e.g. by taking a walk along a river or a lake. Our result confirms the strong preference for

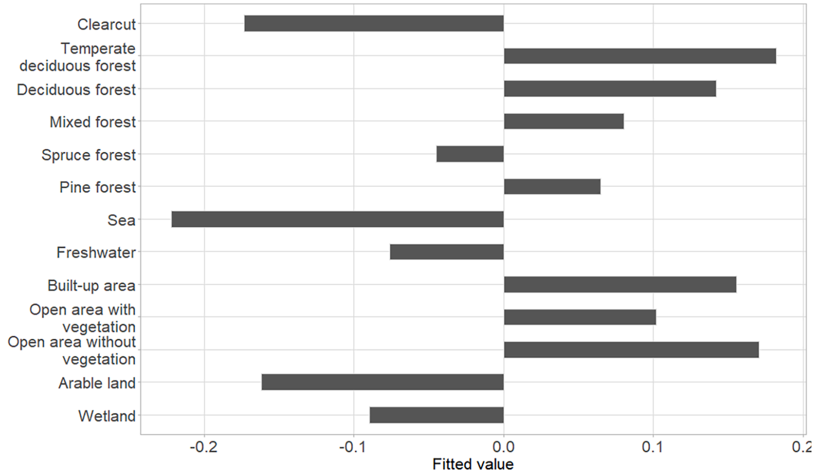


Fig. 4. Partial dependence plot for the land cover predictor of model 1, showing how land cover type affects the probability that an area is chosen for recreation. Positive values represent an increased probability.

water environments for recreation within urban areas, which has been shown in stated preference studies (Schneider, 2009). The high selection values for built-up areas and arable land in the most rural category could be explained by these land cover types being clustered around settlements, and thus frequently visited because they occur close to home.

4.2. Travel distances for recreational activities

To be able to plan for recreation, an understanding of what distances people are willing to travel to reach a recreational area is paramount. We found that the overall median distance from home to the recreational area was 2 km, with some variation between types of activities and seasons. To note here is that the survey asked the respondents to state the distance from their home to the point they had defined as the center of the recreational area they had visited, and it may have varied among the respondents to what extent this travel distance was regarded as a part of the recreational experience or not.

Studies of actual movement patterns of a broader range of recreational activities and landscapes are rare. A survey of Helsinki residents showed that two thirds of respondents traveled <100 m for their latest "close-to-home" recreational visit (Neuvonen et al., 2007). A study on recreational visits to forests in Denmark estimated the median travel distance to 4 km (Agimass et al., 2018), while an Australian study showed a mean of 5.5 km to pre-defined recreational facilities such as parks, beaches, or rivers (McCormack, Giles-Corti, Bulsara, & Pikora, 2006). All these studies were restricted in the types of locations studied, whereas we have analyzed all recreation conducted in any environment. It has also been shown that travel distance is a key factor in the choice of recreational area. For instance, forests closer to home are much more likely to be chosen (Agimass et al., 2018), and people with access to fewer recreational facilities travel longer distances (McCormack, Giles-Corti, Bulsara & Pikora, 2006). This is also supported by studies in Sweden: a survey on the frequency of recreational visits to the closest forest showed a dramatic decline in frequency when the distance from home exceeded 2 km (Hörnsten, 2000), while Grahn and Stigsdotter (2003) showed that the frequency of visits to the closest recreational area was halved if the distance increased from 300 m to 1000 m. These findings suggest that the longer distances in our study could be a sign of a recreational deficit, with people traveling further than they would prefer to reach their recreational areas. In another study, almost 50 % of Swedes reported some degree of deficit in their access to recreational areas (Petersson-Forsberg, 2014).

We observed that travel distances varied depending on which activity was performed. Travel distances in our study were shorter (500–1000 m) for walking, running and "spending time in nature", while they were longer for activities such as swimming and berry/mushroom picking. The former can be performed in most environments, imposing few specific demands of the landscape, while the latter require certain landscape features to be performed, which could explain the difference in travel distances. They are also activities that are more associated with longer visits, which presumably increases the willingness to travel further. The traveling distances varied over the year, reaching a minimum in the winter months and a maximum in July. This could be an effect of most people having vacation in July, increasing the time available for recreational activities.

4.3. Predicting recreational use from landscape characteristics

Our five models performed poorly, being only slightly better than chance at classifying the used points from the random availability sample. The models that employed a buffer around each point (model 2–5) were a minor improvement to the point-based model (model 1). The outcome was not sensitive to the scale of the buffer, with model 2–5 having similar predictive power. For model 1 there was only one influential predictor, land cover type, while model 2–5 each had sets of 7 or 8 predictors that were mainly the same for all models. The land cover predictor of model 1 aligned with previous stated preference research, with e.g. the rankings of different forest types being the same as in Gundersen and Frivold (2008). Model 2–5 had many influential predictors, but all with almost flat responses: they each influenced the predicted outcomes only to a small degree (Supplementary materials S8). Although the models performed slightly better, the weak effects of each variable made them hard to draw any meaningful conclusions from.

The poor predictive power of the machine learning models suggests that the landscape characteristics we investigated (land cover, heterogeneity, topology, path and road density, forest characteristics and protected areas) are weak predictors of actual landscape usage. This came as a surprise, as previous stated preference research has revealed a multitude of effects of different aspects of landscape characteristics, e.g.

indicating what kinds of forests are preferred (Gundersen & Frivold, 2008), the role of landscape heterogeneity (Dronova, 2017; Filyushkina et al., 2017), or the importance of different kinds of infrastructure such as paths and roads (De Valck et al., 2017). Our modeling results are contrary not only to these studies, but also to some studies of revealed preferences. Kienast et al. (2012) showed that biophysical landscape characteristics, such as land cover, could explain patterns of recreation in landscapes around Swiss towns. Their models also showed that travel distance was the most influential factor; similar results could be seen in patterns of usage of a national park in France (Tardieu & Tuffery, 2019). In a study on urban forests in Germany, the strongest predictors of recreational supply were related to human infrastructure, such as monuments, and bluespace (Baumeister, Gerstenberg, Plieninger, & Schraml, 2020). However, similar to our study, they found weak predictive powers of forest characteristics. A study on recreationists in Hamburg showed no correlation between the characteristics of sites and the frequency of visit, and also a disconnect between the preferences stated by the recreationists and which site they visited (Boll, von Haaren, & von Ruschkowski, 2014). Similarly, Bagstad et al. (2016) found no correlations between perceived aesthetic values of landscapes and modeled values based on biophysical characteristics. Taken together, this paints a muddled picture of the relationship between recreationists and the landscape, with some disconnect between results of different studies, warranting further research.

Given that earlier studies on stated preferences have shown preferences for many biophysical characteristics included in our models (e.g. Gundersen & Frivold, 2008), it was unexpected that these preferences were not manifested in our analyses. One possible explanation is linked to the long travel distances observed in this study: it could be that the recreationists' preferred landscape is not accessible enough, and that they instead choose a location that is closer to home, even though it is not their most preferred option. There are however other possible explanations for the lack of clear patterns: Firstly, in Sweden, recreation can be performed on any land, public or private, while in other countries recreation to a higher extent is restricted to certain areas, which might be more homogenous in their attributes. This makes it less likely to find strong effects of certain biophysical characteristics in Sweden, especially if people enjoy variation, and choose areas because they have different characteristics to what they have visited previously. Secondly, the land cover map data used here is coarse in its categorization: each land cover class contains a range of different environments, for example the class "Open area with vegetation" includes both urban lawns and semi-natural grasslands. The SLU forest map adds some nuance, at least for forested environments, by supplying information on tree height, tree species composition and volume; still, this might not be enough to properly characterize how the landscape is experienced by recreationists. We used all the spatial data that was available on a national scale, however, there are several other aspects that might be important for recreationists, which mainly can be studied at a smaller spatial scale. For instance, recreational infrastructure, such as campgrounds (Donovan, Cerveny, & Gatziolis, 2016) and perceived safety (Lis & Iwankowski, 2021) has been shown to influence outdoor recreation.

Due to the weak main effects, it was difficult to draw any conclusions from the interaction effects between predictors of the models. Type of activity performed was the predictor with most interaction effects, which is not surprising, since this has previously been shown to alter how landscapes are used (De Valck et al., 2016). In our models there were only weak interactions between landscape characteristics and gender, age, income, disability and level of education. Previous research has shown that preference can be influenced by socio-demographic factors (van Zanten et al., 2014), cultural differences (Gosal et al., 2021) or group identity (Scott et al., 2009). Since the main effects of the landscape characteristics were weak, and we do not expect the interaction effects to be stronger than the main effects, we cannot say in what manner individual attributes or type of activity affects the choice of recreational location.

Despite our under-performing models, we believe that overall our methodology is sound. A weakness is our estimate of what landscape the recreationists experienced, where the respondent only provided the center point of the area they had visited. Since we did not know how large an area they had visited, we sampled circular buffers of various sizes around this point. For the large fraction of the data set where the recreationists moved over a larger area (e.g. walking, cycling) our approach sampled a smaller part of their experience. We argue that this approach is valid, in that we are contrasting a part of the landscape experienced by the recreationists to a landscape they did not experience. An improvement would have been to collect data on the exact route each recreationist had taken. We further lacked exact information on where the recreationist had traveled from, which would have improved our estimate of what landscape was available, instead of having to rely on the destination point combined with the travel distance.

We believe BRT modeling to be a very well suited tool for analyzing a complex phenomenon such as recreation. It has generally high predictive power, combined with flexibility and the ability to handle any number of predictors (regardless of collinearity). It does not require the specification of interaction effects in advance, nor assume linear relationships between predictors and response. Its main problem of producing models that are harder to interpret can be overcome, and is in our view worth the drawback.

5. Conclusions

We have found that recreationists in Sweden travel farther to recreational areas than what previous research has suggested is preferred. At the same time, we found only weak signals of recreationists having selected the area due to its biophysical characteristics. Thus, recreationists' preferences are not manifested, and one explanation for that is that a low availability of closely situated areas are limiting their choice. This indicates a possible recreational deficit, making it important for policy-makers to take into account the need for recreational opportunities in physical planning, even in a sparsely populated country such as Sweden.

We found that a large proportion of recreation occurs on a small proportion of the total land area (i.e. urban and periurban areas). This is because the population is clustered towards urban areas, combined with the fact that most recreation occurs close to home. This has implications for planning: recreational opportunities can be improved for half of the Swedish population by focusing on these areas, however to improve them for the other half would affect much larger areas.

The outcomes from studies of stated preferences and revealed preferences seem contradictory. To achieve a better understanding of this, a combination of stated preference and revealed preference could be applied in future studies by asking recreationists about their preferences, while at the same time studying their actual recreational patterns.

Conclusively, it is important to take spatial accessibility into account, both when performing research and during physical planning. Recreationists use the landscape that is available to them, which in our study were on average one or two kilometers from home for the most common activities.

CRediT authorship contribution statement

Carl Lehto: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Marcus Hedblom:** Conceptualization, Methodology, Validation, Investigation, Data curation, Writing – original draft, Writing – review & editing, Supervision. **Erik Öckinger:** Validation, Writing – original draft, Writing – review & editing, Supervision. **Thomas Ranius:** Conceptualization, Methodology, Validation, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

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

Acknowledgments

We would like to thank Anna Filyushkina and Peter Fredman for their valuable contributions. The survey was funded by the Swedish Environmental Protection Agency.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.landurbplan.2022.104519.

References

Agimass, F., Lundhede, T., Panduro, T. E., & Jacobsen, J. B. (2018). The choice of forest site for recreation: A revealed preference analysis using spatial data. *Ecosystem Services, 31*, 445–454. https://doi.org/10.1016/j.ecoser.2017.11.016

Bagstad, K. J., Reed, J. M., Semmens, D. J., Sherrouse, B. C., & Troy, A. (2016). Linking biophysical models and public preferences for ecosystem service assessments: A case study for the Southern Rocky Mountains. *Regional Environmental Change, 16*(7), 2005–2018. https://doi.org/10.1007/s10113-015-0756-7

Baumeister, C. F., Gerstenberg, T., Plieninger, T., & Schraml, U. (2020). Exploring cultural ecosystem service hotspots: Linking multiple urban forest features with public participation mapping data. *Urban Forestry & Urban Greening, 48*, 126561. https://doi.org/10.1016/j.ufug.2019.126561

Bell, S., Tyrväinen, L., Sievänen, T., Pröbstl, U., & Simpson, M. (2007). Outdoor Recreation and Nature Tourism: A European Perspective. *Living Reviews in Landscape Research*, *1*. 10.12942/lrlr-2007-2.

Boll, T., von Haaren, C., & von Ruschkowski, E. (2014). The Preference and Actual Use of Different Types of Rural Recreation Areas by Urban Dwellers—The Hamburg Case Study. *PLoS ONE, 9*(10), e108638.

Brown, G., & Fagerholm, N. (2015). Empirical PPGIS/PGIS mapping of ecosystem services: A review and evaluation. *Ecosystem Services, 13*, 119–133. https://doi.org/10.1016/j.ecoser.2014.10.007

De Valck, J., Broekx, S., Liekens, I., De Nocker, L., Van Orshoven, J., & Vranken, L. (2016). Contrasting collective preferences for outdoor recreation and substitutability of nature areas using hot spot mapping. *Landscape and Urban Planning, 151*, 64–78. https://doi.org/10.1016/j.landurbplan.2016.03.008

De Valck, J., Landuyt, D., Broekx, S., Liekens, I., De Nocker, L., & Vranken, L. (2017). Outdoor recreation in various landscapes: Which site characteristics really matter? *Land Use Policy, 65*, 186–197. https://doi.org/10.1016/j.landusepol.2017.04.009

Díaz-Varela, E., Roces-Díaz, J. V., & Álvarez-Álvarez, P. (2016). Detection of landscape heterogeneity at multiple scales: Use of the Quadratic Entropy Index. *Landscape and Urban Planning, 153*, 149–159. https://doi.org/10.1016/j.landurbplan.2016.05.004

Donovan, G. H., Cerveny, L. K., & Gatziolis, D. (2016). If you build it, will they come? *Forest Policy and Economics, 62*, 135–140. https://doi.org/10.1016/j.forpol.2015.11.002

Dronova, I. (2017). Environmental heterogeneity as a bridge between ecosystem service and visual quality objectives in management, planning and design. *Landscape and Urban Planning, 163*, 90–106. https://doi.org/10.1016/j.landurbplan.2017.03.005

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology, 77*(4), 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Ezebilo, E. E., Boman, M., Mattsson, L., Lindhagen, A., & Mbongo, W. (2015). Preferences and willingness to pay for close to home nature for outdoor recreation in Sweden. *Journal of Environmental Planning and Management, 58*(2), 283–296. https://doi.org/10.1080/09640568.2013.854196

Filyushkina, A., Agimass, F., Lundhede, T., Strange, N., & Jacobsen, J. B. (2017). Preferences for variation in forest characteristics: Does diversity between stands matter? *Ecological Economics, 140*, 22–29. https://doi.org/10.1016/j.ecolecon.2017.04.010

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5). https://doi.org/10.1214/aos/1013203451

Giergiczny, M., Czajkowski, M., Żylicz, T., & Angelstam, P. (2015). Choice experiment assessment of public preferences for forest structural attributes. *Ecological Economics, 119*, 8–23. https://doi.org/10.1016/j.ecolecon.2015.07.032

Gosal, A. S., Giannichi, M. L., Beckmann, M., Comber, A., Massenberg, J. R., Palliwoda, J., … Ziv, G. (2021). Do drivers of nature visitation vary spatially? The importance of context for understanding visitation of nature areas in Europe and North America. *Science of The Total Environment, 776*, 145190. https://doi.org/10.1016/j.scitotenv.2021.145190

Grahn, P., & Stigsdotter, U. A. (2003). Landscape planning and stress. *Urban Forestry & Urban Greening, 2*(1), 1–18. https://doi.org/10.1078/1618-8667-00019

Greenwell, B., Boehmke, B., & Cunningham, J. (2020). gbm: Generalized Boosted Regression Models (Version 2.1.8). Retrieved from https://CRAN.R-project.org/package=gbm.

Gundersen, V. S., & Frivold, L. H. (2008). Public preferences for forest structures: A review of quantitative surveys from Finland, Norway and Sweden. *Urban Forestry & Urban Greening, 7*(4), 241–258. https://doi.org/10.1016/j.ufug.2008.05.001

Hedblom, M., Andersson, E., & Borgström, S. (2017). Flexible land-use and undefined governance: From threats to potentials in peri-urban landscape planning. *Land Use Policy, 63*, 523–527. https://doi.org/10.1016/j.landusepol.2017.02.022

Hörnsten, L. (2000). *Outdoor recreation in Swedish forests: Implications for society and forestry (Swedish University of Agricultural Sciences)*. Uppsala: Swedish University of Agricultural Sciences. Retrieved from http://urn.kb.se/resolve?urn=urn:Nbn:Se:Slu:Epsilon-e-5406.

Hörnsten, L., & Fredman, P. (2000). On the distance to recreational forests in Sweden. *Landscape and Urban Planning, 51*(1), 1–10. https://doi.org/10.1016/S0169-2046(00)00097-9

IPBES. (2019). Global assessment report on biodiversity and ecosystem services of the Intergorvernmental Science-Policy Platform on Biodiversity and Ecosystem Services (p 1148). Retrieved from doi: 10.5281/zenodo.3831673 .

Juutinen, A., Kosenius, A.-K., Ovaskainen, V., Tolvanen, A., & Tyrväinen, L. (2017). Heterogeneous preferences for recreation-oriented management in commercial forests: The role of citizens' socioeconomic characteristics and recreational profiles. *Journal of Environmental Planning and Management, 60*(3), 399–418. https://doi.org/10.1080/09640568.2016.1159546

Kearney, A. R., & Bradley, G. A. (2011). The Effects of Viewer Attributes on Preference for Forest Scenes: Contributions of Attitudes, Knowledge, Demographic Factors, and Stakeholder Group Membership. *Environment and Behavior, 43*(2), 147–181. https://doi.org/10.1177/0013916509353523

Kienast, F., Degenhardt, B., Weilenmann, B., Wäger, Y., & Buchecker, M. (2012). GIS-assisted mapping of landscape suitability for nearby recreation. *Landscape and Urban Planning, 105*(4), 385–399. https://doi.org/10.1016/j.landurbplan.2012.01.015

Komossa, F., van der Zanden, E. H., Schulp, C. J. E., & Verburg, P. H. (2018). Mapping landscape potential for outdoor recreation using different archetypical recreation user groups in the European Union. *Ecological Indicators, 85*, 105–116. https://doi.org/10.1016/j.ecolind.2017.10.015

Korpilo, S., Virtanen, T., & Lehvävirta, S. (2017). Smartphone GPS tracking—Inexpensive and efficient data collection on recreational movement. *Landscape and Urban Planning, 157*, 608–617. https://doi.org/10.1016/j.landurbplan.2016.08.005

Korpilo, S., Virtanen, T., Saukkonen, T., & Lehvävirta, S. (2018). More than A to B: Understanding and managing visitor spatial behaviour in urban forests using public participation GIS. *Journal of Environmental Management, 207*, 124–133. https://doi.org/10.1016/j.jenvman.2017.11.020

Laatikainen, T., Tenkanen, H., Kyttä, M., & Toivonen, T. (2015). Comparing conventional and PPGIS approaches in measuring equality of access to urban aquatic environments. *Landscape and Urban Planning, 144*, 22–33. https://doi.org/10.1016/j.landurbplan.2015.08.004

Lantmäteriet. (n.d.). *National Elevation Model*. Retrieved from https://www.lantmateriet.se/en/geodata/geodata-products/product-list/terrain-model-download-grid-1/.

Lis, A., & Iwankowski, P. (2021). Why is dense vegetation in city parks unpopular? The mediative role of sense of privacy and safety. *Urban Forestry & Urban Greening, 59*, 126988. https://doi.org/10.1016/j.ufug.2021.126988

Manly, B. F. J., McDonald, L. L., Thomas, D. L., McDonald, T. L., & Erickson, W. P. (2002). *Resource Selection by Animals, 231*.

McCormack, G. R., Giles-Corti, B., Bulsara, M., & Pikora, T. J. (2006). Correlates of distances traveled to use recreational facilities for physical activity behaviors. *International Journal of Behavioral Nutrition and Physical Activity, 10*.

Molnar, C. (2018). iml: An R package for Interpretable Machine Learning. *Journal of Open Source Software*, *3*(26), 786. 10.21105/joss.00786.

National Forest Inventory of Sweden. (2009). *Forestry statistics 2009*. Umeå: Swedish University of Agricultural Sciences. Retrieved from Swedish University of Agricultural Sciences website: https://pub.epsilon.slu.se/3405/1/Skogsdata2009.pdf.

Neuvonen, M., Sievänen, T., Tönnes, S., & Koskela, T. (2007). Access to green areas and the frequency of visits – A case study in Helsinki. *Urban Forestry & Urban Greening, 6*(4), 235–247. https://doi.org/10.1016/j.ufug.2007.05.003

Northrup, J. M., Hooten, M. B., Anderson, C. R., & Wittemyer, G. (2013). Practical guidance on characterizing availability in resource selection functions under a use–availability design. *Ecology, 94*(7), 1456–1463. https://doi.org/10.1890/12-1688.1

Norton, L. R., Inwood, H., Crowe, A., & Baker, A. (2012). Trialling a method to quantify the 'cultural services' of the English landscape using Countryside Survey data. *Land Use Policy, 29*(2), 449–455. https://doi.org/10.1016/j.landusepol.2011.09.002

OpenStreetMap Foundation. (n.d.). *OpenStreetMap*. Retrieved from https://www.openstreetmap.org/copyright.

Petersson-Forsberg, L. (2014). Swedish spatial planning: A blunt instrument for the protection of outdoor recreation. *Journal of Outdoor Recreation and Tourism, 5–6*, 37–47. https://doi.org/10.1016/j.jort.2014.03.003

Qiu, L., Lindberg, S., & Nielsen, A. B. (2013). Is biodiversity attractive?—On-site perception of recreational and biodiversity values in urban green space. *Landscape and Urban Planning, 119*, 136–146. https://doi.org/10.1016/j.landurbplan.2013.07.007

R Core Team. (2020). *R: A language and environment for Statistical Computing*. Vienna: Austria. Retrieved from https://www.R-project.org/.

SCB. (2018). *SCB:s avgränsningar av koncentrerad bebyggelse* (p. 32). Stockholm: Central Bureau of Statistics. Retrieved from Central Bureau of Statistics website: https://

www.scb.se/hitta-statistik/statistik-efter-amne/miljo/markanvandning/tatorter/pong/publikationer/scbs-avgransningar-av-koncentrerad-bebyggelse/.
Schneider, I. E. (2009). Urban Water Recreation: Experiences, Place Meanings, and Future Issues. In L. A. Baker (Ed.), *The Water Environment of Cities* (pp. 125–140). Boston, MA: Springer US. 10.1007/978-0-387-84891-4_7.
Scott, A., Carter, C., Brown, K., & White, V. (2009). 'Seeing is Not Everything': Exploring the Landscape Experiences of Different Publics. *Landscape Research, 34*(4), 397–424. https://doi.org/10.1080/01426390903009289
SEPA. (2015). *Friluftsliv 2014: nationell undersökning om svenska folkets friluftslivsvanor, No. 6691*. Stockholm: Swedish Environmental Protection Agency.
SEPA. (2019). *Friluftsliv 2018: nationell undersökning av svenska folkets friluftslivsvanor* (No. 6887). Stockholm: Swedish Environmental Protection Agency. Retrieved from Swedish Environmental Protection Agency website: https://www.naturvardsverket.se/978-91-620-6887-5.
SLU. (2015). SLU Forest map. Retrieved from https://www.slu.se/en/Collaborative-Centres-and-Projects/the-swedish-national-forest-inventory/foreststatistics/slu-forest-map/.
Swedish Environmental Protection Agency. (2018). *CadasterENV Sweden - A multi-scale and multi-purpose land cover monitoring system*. 6.
Tardieu, L., & Tuffery, L. (2019). From supply to demand factors: What are the determinants of attractiveness for outdoor recreation? *Ecological Economics, 161*, 163–175. https://doi.org/10.1016/j.ecolecon.2019.03.022
van Zanten, B. T., Verburg, P. H., Koetse, M. J., & van Beukering, P. J. H. (2014). Preferences for European agrarian landscapes: A meta-analysis of case studies. *Landscape and Urban Planning, 132*, 89–101. https://doi.org/10.1016/j.landurbplan.2014.08.012
Yoshimura, N., & Hiura, T. (2017). Demand and supply of cultural ecosystem services: Use of geotagged photos to map the aesthetic value of landscapes in Hokkaido. *Ecosystem Services, 24*, 68–78. https://doi.org/10.1016/j.ecoser.2017.02.009

III

Seeing through their eyes: Revealing recreationists' landscape preferences through viewshed analysis and machine learning

Carl Lehto[1], Marcus Hedblom[2], Anna Filyushkina[1], Thomas Ranius[1]
[1]Department of Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden.
[2]Department of Urban and Rural Development, Uppsala, Sweden.

Abstract

Planning for outdoor recreation requires knowledge about the needs and preferences of recreationists. While previous research has mainly relied on stated preferences, recent advances in spatial data collection and analysis have enabled the assessments of actual usage patterns. In this study, we explored how landscape characteristics interact with the attributes of recreationists to determine their area choice for recreation. Using a public participation GIS (PPGIS) approach we asked residents of a Swedish city in the boreal region to draw typical recreational routes and identify favourite places for recreation on a digital online map (1389 routes, 385 individuals). We employed a novel methodology, where LiDAR data was used to calculate what was visible along all routes and at favourite places (viewsheds) in order to more realistically capture the landscape that each recreationist had experienced. Using machine learning modelling, we compared landscape characteristics of experienced areas with areas available to each recreationist. Our novel approach yielded accurate models that revealed that water environments, recreational infrastructure and deciduous forests increased the probability of choosing an area for recreation, while urban environments, noise, forest clearcuts and young forests had the opposite effect. Characteristics of the recreationists such as age, gender, level of education, or of the activity, such as type of activity performed, did not meaningfully influence area choice. Our findings suggest that it is possible to improve the conditions for recreation by developing recreational infrastructure, maintaining recreation opportunities close to waters, and adapting forest management in areas important for recreation.

Keywords: PPGIS; Outdoor recreation; Viewsheds; Machine learning; Green space; Blue space

1 Introduction

Urban and near-urban green spaces are in decline globally (Richards & Belcher, 2020). A reason for this is the undervaluation of green space in decision-making processes due to the challenges of incorporating cultural ecosystem services, such as the provision of opportunities for outdoor recreation (Fish, Church & Winter, 2016). To address this issue, there have been suggestions to incorporate green space indicators into physical planning, such as residents having at least a certain amount of green space within an accessible distance (Ekkel & de Vries, 2017). While such efforts emphasize the importance of quantity, it's also crucial to consider the quality of green space. Higher quality has been linked to both increased visitation (Kajosaari et al., 2024) and improved health outcomes for recreationists (Nguyen, Astell-Burt, Rahimi-Ardabili & Feng, 2021). To be able to assess the quality of greenspace however, which characteristics that are attractive for recreationists need to be understood.

Research on landscape preferences of recreationists has shown that, for instance, forest characteristics (Gundersen, Köhler, & Myrvold, 2019), landscape heterogeneity (Filyushkina, Agimass, Lundhede, Strange, & Jacobsen, 2017), and presence of water (White et al., 2010) can affect willingness to access a specific area. Most research has relied on *stated preferences*, mainly studied by showing recreationists pictures of landscapes and asking them to rate them. The results from such studies have been synthesized to map the supply of recreational landscapes (e.g. Paracchini et al., 2014; Walz & Stein, 2018). With recent technological advancements, particularly the widespread adoption of smartphones, an increasing number of studies have been performed on *revealed preferences*, i.e. how recreationists actually utilize landscapes. Various methods have been employed, such as GPS tracking (Korpilo, Virtanen, & Lehvävirta, 2017), data scraping from social media (Karasov, Vieira, Külvik, & Chervanyov, 2020), and implementations of public participation geographic information systems (PPGIS), wherein online surveys are deployed to collect spatial data from respondents (Brown & Fagerholm, 2015).

Most PPGIS studies correlate landscape utilization with remote sensing data, such as land cover maps. Often the studies do not control for effects stemming from spatial heterogeneity (e.g. Kienast, Degenhardt, Weilenmann, Wäger, & Buchecker, 2012; Baumeister, Gerstenberg, Plieninger, & Schraml, 2020; De Valck et. al 2016), such as what has been termed *distance-decay*, which means that areas further away are less likely to be visited (De Valck & Rolfe, 2018). How easily accessible an area is has shown to have a strong influence on to what degree it is used for recreation, with recreationists tending to utilize landscapes that are in close proximity (Grahn & Stigsdotter, 2003; Hörnsten, 2000; Neuvonen, Sievänen, Tönnes, & Koskela, 2007; Lehto, Hedblom, Öckinger, & Ranius, 2022). To be able to tease apart the effect of preference with that of accessibility it is important to control for this. A further methodological obstacle is defining what landscape the recreationist perceived, with a common approach being sampling a buffer around respondent’s locations (e.g. Baumeister et al. 2020). An alternative (or complement) is to calculate viewsheds, using topography to estimate what landscape was visible to the recreationist (e.g. Schirpke, Tasser & Tappeiner, 2013; Yoshimura & Hiura 2017). This approach increases the realism of the analysis, but has rarely been utilized in PPGIS studies, probably due to its higher computational cost.

Landscape preferences of recreationists have been found to be heterogeneous, with variation in preference due to the type of preferred activity (De Valck et al., 2017), socio-demographic factors (van Zanten, Verburg, Koetse, & van Beukering, 2014), held beliefs (Kearney & Bradley 2011) and attitudes (such as nature relatedness: Nisbet, Zelenski & Murphy, 2009; Elbakidze et al., 2022; Flowers, Freeman, & Gladwell, 2016), cultural differences (Gosal et al., 2021), user typology (Komossa, van der Zanden, & Verburg, 2019), age, gender (Gunnarsson, Knez, Hedblom, & Sang, 2017), or group identity (Scott, Carter, Brown, & White, 2009). However, heterogeneity of preference has mainly been shown in studies of stated preferences (e.g. De Valck et al., 2017), while only a few studies have revealed differences in actual patterns of recreational usage of landscapes (e.g. De Valck et al., 2016; Kienast et al., 2012).

The goals with this study were:

1. Exploring which landscape characteristics (e.g. land cover, heterogeneity, topography, recreational infrastructure, forest characteristics) are important determinants of the choice of area for recreation.
2. Investigating to what degree the preference for these landscape characteristics depends on attributes of the recreationist (age, gender, level of education, nature relatedness), or attributes related to the recreational visit (type of activity, frequency of visit, time spent, time of week/year).
3. Furthering the field of PPGIS analysis of recreation by developing and implementing a more advanced approach, based on the inclusion of viewsheds, network analysis, and machine learning.

To achieve these goals, we employ a PPGIS survey to collect spatial data on *typical routes* and *favourite places* of recreationists in and around the city of Umeå, Sweden. The reason for including both modes of recreation was that we expected the routes to give a more complete picture of where daily recreation is performed, while the favourite places to a higher degree would exhibit which landscape characteristics are preferable (Frick, Degenhardt & Buchecker; 2007).

We employ a novel methodology, in which we firstly control for the effect of accessibility using network analysis, to properly compare the areas used by the recreationists to areas that were available to them. Secondly, we capture the perceived recreational experience in a more realistic manner through estimating what landscape was visible to the recreationist using LiDAR data. Finally, we use flexible machine learning modeling in the form of Boosted Regression Trees, capable of handling a large number of map covariates.

2 Materials and Methods

2.1 Survey

2.1.1 Study area

The study was performed in Umeå municipality, Sweden (Figure 1). It covers an area of approximately 2300 km² with a population of 131 000, yielding a population density of 56/km² (Umeå municipal government, 2022). Its seat, the city of Umeå, is the 13th most populous city in Sweden and harbors the largest university in northern Sweden. The climate is cold continental, with freezing winters and mild summers. The surrounding landscape is dominated by managed forest land (mainly coniferous), but with some remnants of unmanaged forest as well as arable land, wetlands, and lakes. Sweden has a right of public access that encompasses almost all land, both public and private, which means that there are very few restrictions on where people can engage in outdoor recreation.

Figure 1. Land cover types of the study area (Umeå municipality, Sweden). Blue is water, light blue is wetland, green is forest, yellow is arable land, and grey is built-up areas.

2.1.2 Survey design

An invitation to participate in our survey was sent to 3,000 residents over 18 years of age of the Umeå postal area via mail in September 2021, with a reminder sent three weeks later. The list of recipients was acquired from the Swedish state person address registry, which provided a stratified sample designed to be proportional to the population of Umeå with regards to gender and age. The invitation contained a link to the digital survey, which was implemented using the online survey tool Maptionnaire.

In the survey, the respondents were asked to provide background personal data (age, gender, and level of education). Furthermore, they were asked to assess to what degree they were a *nature-oriented* and an *urban-oriented* person, using two respective sliders with a range between 0 and 100, where 0 represented "Not at all" and 100 "Fully". The terms were not defined further to the survey respondents, and was included as it had affected perception of green space in a previous study (Gunnarsson et al, 2017).

The main part of the survey was divided into two sections. In the first section, the respondents were asked to summarize their outdoor recreation by drawing typical routes on a map of Umeå municipality. The respondents were asked to only draw what they experienced as the recreational route, and not including the travel route. For each route drawn, follow-up questions were asked, such as what type of activity was performed, the mode of transportation used to reach the area, and the frequency and duration of visits. This procedure was done separately for summer and winter recreation. The second section of the survey tasked the respondents with marking the locations of their favourite places when engaging in recreation. A favourite place was defined as a place "holding any specific importance, such as a place of beauty or somewhere you often stop and spend time in". The participants were also asked to mark a location as close to their home location as they felt comfortable with. Prior to deployment, the survey was tested on a convenience sample of 45 friends and colleagues to assess its clarity and adjusted accordingly. Supplementary materials S1 contains an English translation of the survey.

Since the survey did not handle sensitive information, we assessed it as not falling under any of the criteria listed in the Swedish Ethical review act (2003:460), and thus did not need authorization from the Swedish Ethical Review Authority. Collected data was handled in accordance with the General Data Protection Regulation (Regulation (EU) 2016/679). At the start and the end of the survey, the respondents were provided information on how the collected data would be handled and consent was asked.

2.1.3 Summary of responses

Of the 3,000 invited participants, 658 (22 %) opened the link to the digital survey, and 285 (10 %) finished the entire survey. Data from respondents who drew at least one route or placed one favourite place was kept and used in the analysis. The routes were manually screened to assess data quality, with 15 erroneous routes removed. The basis for removal was that the route had crossed itself, had many acute angles, or had unrealistically long distances between vertices. The final sample was 1389 routes within Umeå municipality (947 summer, 442 winter) from 358 individuals (mean 3.88 routes/individual, std. dev 4.1). For the favourite places within Umeå municipality, the final sample consisted of 275 from 181 individuals. The routes and favourite places are visualized in supplementary materials S2.

Of the respondents, 47 % were male and 53 % female. Median age was 48 (std. dev. 17), which is similar to the Umeå average (49 +- 18, Umeå kommuns demografidatabas 2023). The respondents were more educated than the Swedish average; 70 % had attended higher education for at least one year, compared to the Swedish average of 45 % (SCB, 2021).

2.2 Modelling recreational choice

With the collected spatial data, we analyzed which factors were most important in the choice of location for outdoor recreation. We did this in three steps: first we defined which areas the respondents had available to them; then we sampled various map data in both the used area and the available area; and finally we trained a machine learning model to compare the characteristics of the areas visited by the recreationists with those that were available to them.

2.2.1 Use-available framework

The routes and favourite places were analyzed in a use-available framework, where characteristics of the use sample are compared to those of the availability sample (Northrup, Hooten, Anderson & Wittemyer, 2013). Here, our use sample consisted of the routes and favourite places marked by the survey respondents. To construct the availability sample for the routes, first a spatial network analysis was performed. This analysis used path and road map data to determine which areas could have been reached in the same time it took to reach the beginning of the route from the home of the respondent, utilizing the same mode of transportation as the respondent (on foot, by bike, or by car/public transportation). A random point was then placed along the edge of this area, and at that point a copy of the route was placed rotated 180 degrees (Figure 2). The route was rotated to minimize the risk of overlap between the performed and random route. For the respondents that drew a route but did not mark their home location (101 people, 204 routes), the geographic median of all other home locations was used. The placement of the random routes were constrained so that the starting point of a route was not placed in water, and the entire route was always inside the municipal borders.

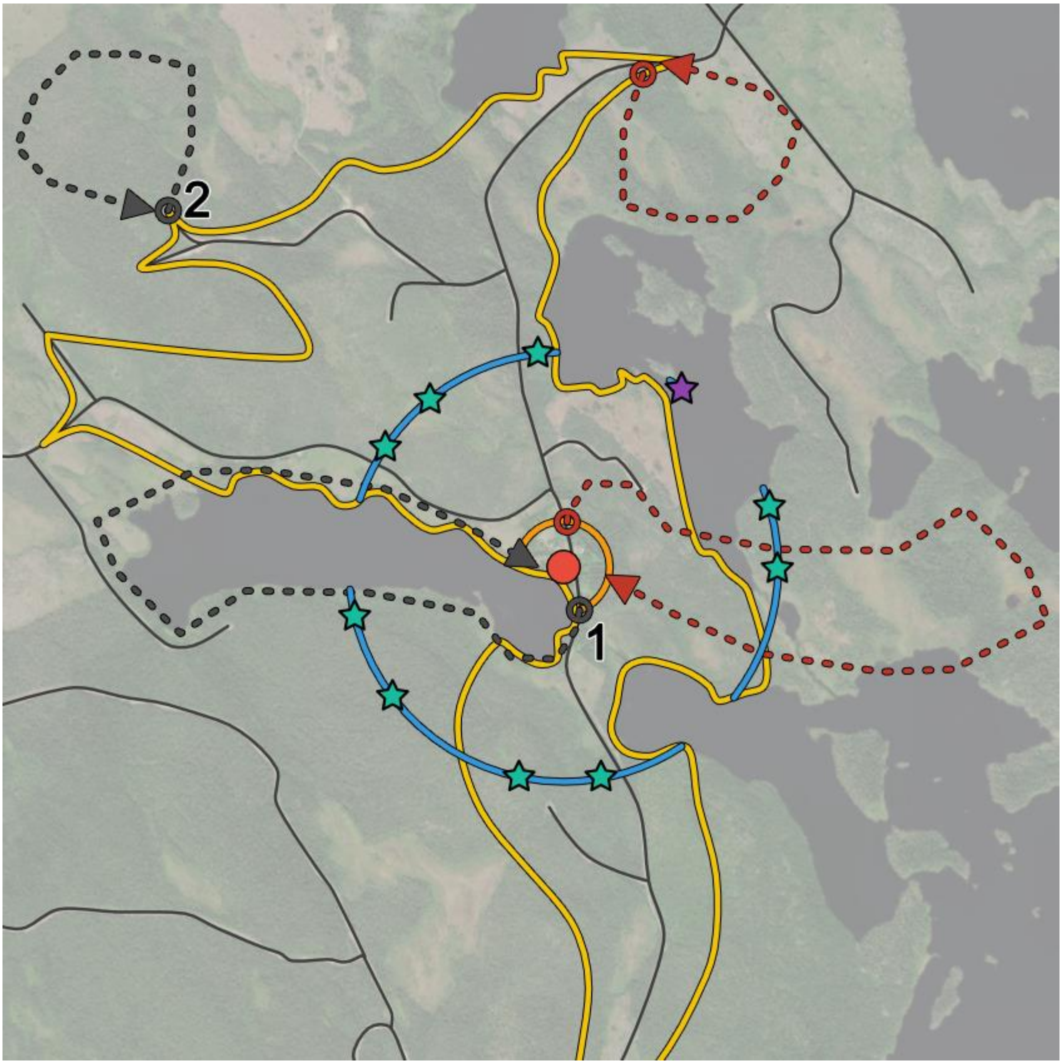


Figure 2. An example of how the availability sample was created for one respondent. The respondent had placed the home location (red dot), drawn two routes (grey dashed lines), and placed one favourite place (purple star). For the routes, the availability sample was created by copying the shape of each route, flipping it (red dashed lines), and placing it in a random position that could be reached in the same amount of time as the performed route, taking into account the mode of transportation. All equidistant locations from the home location are represented by the orange circle. The availability sample was randomly placed at a terrestrial point along this circle. For the favourite place, nine random locations (teal stars) were placed at an equal distance from the home location (blue circle) as the favourite place (purple star).

To construct the availability sample for the favourite places, random points were placed around each home location within a distance equal to the distance to the favourite place. Network analysis was not applied here, since information on the mode of transportation was not requested for the favourite places. We evaluated the sensitivity of the model predictions to the size of the availability sample by creating models with either one, three or all nine random points included, as suggested by Northrup et al. (2013). The random points were constrained to not be placed in water or outside the municipal borders. This sensitivity analysis was not performed for the routes, due to the assumption that a 1:1 matched sample was enough due to the larger sample size.

2.2.2 Defining the area experienced by the respondents

To define the spatial extent of what recreationists experienced, a combination of two approaches was employed. First, a buffer with a radius of 50 m was created around each favourite place and along each route, representing the immediate surroundings (cf. Baumeister et al. 2020). This distance was chosen as a conservative estimate of a 'perceptual horizon', ensuring our analysis captures the core of the recreational experience without extending into possibly unexperienced areas. Secondly, a viewshed was calculated at each favourite place and along each route, representing the area that was visible. The viewsheds were constructed using LiDAR data (Lantmäteriet, 2023), which provide high resolution heightmaps of both the ground terrain and any obstacles that block vision (trees, buildings etc.). On the one hand, treating trees as complete visual barriers yields unrealistically small viewsheds, since vision is often only partially obscured by foliage. On the other hand, not accounting for trees would instead lead to unrealistically large viewsheds. As a compromise, we treated trees outside the 50 m buffer as a visual barrier and assumed full visibility within this distance (Figure 3). For the favourite places, the viewshed was calculated from the point of the place, while for the routes it was calculated every 100 m along the routes, and then summed into a total viewshed. The viewsheds were calculated from a height of 1.5 m, with a maximum sight distance of 1 km. The distance between calculations was chosen partly for computational cost reasons, but also so that the 50 m buffer where we assume full visibility would exactly lie tangent with the next calculated viewshed.

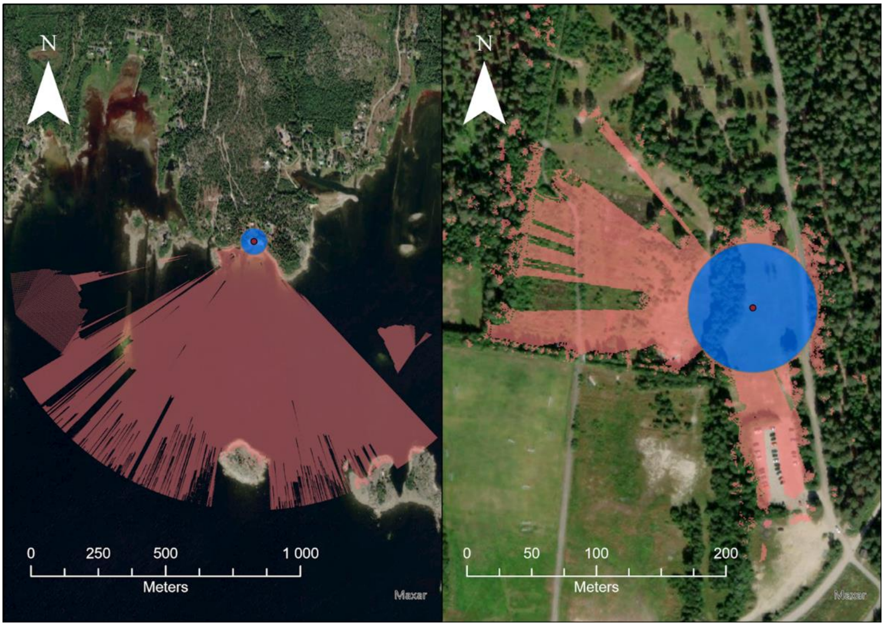


Figure 3. Example of the sampled landscape around two favourite places. The red point is the favourite place provided by the survey respondent, the blue circle is the 50 m buffer, and red areas are the calculated visible landscape when standing at the point (viewshed).

2.2.3 Model predictors

Landscape characteristics were sampled using several map sources. In addition to landscape predictors, characteristics of the respondents and of the activity were included as predictors (Table 1). Most landscape predictors were sampled in both the viewshed and the buffer, while those assumed to be more related to the immediate experience (e.g. noise) were exclusively sampled in the buffer.

Land cover was extracted from the CadasterENV Sweden map (Swedish Environmental Protection Agency, 2023) and reclassified from 25 original classes into 13 classes (Supplementary materials S3). The fractions of each land cover of the buffer and the viewshed, respectively, were used as predictors. They were also used to estimate landscape heterogeneity of viewsheds and buffers by calculating the Shannon-Wiener diversity index, which reflect how many land cover types there are and how evenly the area is divided into these types (Shannon, 1948).

The Swedish Agricultural University forest map added nuance to the land cover maps in forested areas by supplying estimates of the mean tree height and volumes of different tree species, as well as total tree biomass volume (SLU 2015). Conservation value of land was included as a predictor by combining several sources of map data: formally protected areas (nature reserves and protected forest biotopes) sourced from the Swedish Environmental Protection Agency, woodland key habitats from the Swedish Forestry Agency (i.e. forests with high conservation values; Timonen et al. (2010)), and areas with high conservation values identified by the Umeå municipal government. The predictor used in the model was the percentage overlap between the buffer and any of these maps.

We included noise level as a predictor using three maps of estimated average daily noise levels (Lden) due to road traffic, railroad traffic, and industry, respectively (Umeå Municipal government, 2016). These were combined by taking the highest estimated noise level at each point of the three maps, and then calculating the average across the buffer. To include recreational infrastructure, data on amenities (shelters, toilets, and fireplaces) from the municipal government was used as a predictor by calculating the average distance to the nearest recreational amenity, while paths and roads were extracted from OpenStreetMap, and densities of each were calculated within the buffer. As topographical measures, we used the median, standard deviation, and range (largest difference) of elevation above sea level within buffer and viewshed.

Table 1

Predictors Used in the Machine Learning Models

Predictor	Description	Unit
Land cover (13 predictors) [a,b]	Composition of reclassified land cover types	%
Shannon-Wiener diversity [a,b]	Landscape heterogeneity, calculated using the reclassified land cover classes	Unitless
Tree height [a,b]	Average height of trees	m
Spruce volume [a,b]	Average standing volume of Norway spruce	m^3/ha
Pine volume [a,b]	Average standing volume of Scots pine	m^3/ha
Birch volume [a,b]	Average standing volume of birch	m^3/ha
Biomass volume [a,b]	Average volume of all vegetation	m^3/ha
Elevation (3 predictors) [a,b]	Median, standard deviation and range of elevation	m
Noise [a]	A-weighted day noise level	Lden dB(A)
Area of conservation concern [a]	Overlap of buffer with areas of high nature conservation values	%
Path/road density [a]	Density of paths and roads within buffer	m/m^2
Distance to amenities	Average distance to the closest recreational amenity	m
Age	The age of the respondent	years
Gender	The gender of the respondent	Male; Female; Other
Education	Highest level of finished education	Elementary School; Secondary School; University 2 yrs or less; University > 2 yrs; Folk high school
Urban person	To what extent the person self-identified as an "Urban person"	Unitless [0-100]
Nature person	To what extent the person self-identified as a "Nature person"	Unitless [0-100]
Activity *	Type of activity engaged in	Walking; Walking with dog; Jogging/running; Cycling; Ice skating; Cross-country skiing
Season *	Time of year	Summer; Winter
Transportation *	The mode of transportation used to get to the route from home	On foot; Bicycle; Car; Public transportation
Weekday/Weekend *	Whether the route primarily is performed during weekdays, the weekend, or both	Weekday; Weekend; Both
Time usually spent *	The average visit duration	Minutes
Visit frequency *	How often the route is performed	Times per year

* Predictor only used for route model

[a] Predictor was sampled within the 50 m buffer

[b] Predictor was sampled within the viewshed

2.2.4 Modeling: Boosted regression trees

Statistical modeling was performed using boosted regression trees (BRT), also known as gradient boosting machines, or generalized boosting models. BRT is a machine learning approach that can be used both for regression or classification, where the predictive model is created by iteratively building an ensemble of many decision trees (Friedman, 2001). The method has several advantages: it does not assume linear relationships between predictor variables and response variables; it can handle a large number of predictors regardless of multicollinearity; and it eschews the need for model selection or pre-specifying interaction effects. The main disadvantage of BRT is a lower interpretability of the final models, being more of a "black box" than traditional regression models such as GAMs or GLMs. However, with recent methodological advances (e.g. the Interpretable Machine Learning package for R applied here; Molnar, 2018), these shortcomings are mitigated. For a more detailed exploration of BRT, see Elith, Leathwick & Hastie (2008).

All analyses and visualizations were carried out in R version 4.0.3 (R Core Team, 2020). Boosted regression trees were constructed using a Bernoulli distribution with Use/Available as the response variable, and hyperparameters were set using a grid-search to find the optimal values, with models being evaluated on their cross-validated accuracy. Feature (predictor) importance, interaction effects and accumulated local effect (ALE) plots were evaluated using the iml package (Molnar, Casalicchio & Bischl, 2018). A more detailed account of the modeling is presented in supplementary materials S4.

3 Results

3.1 Model validity

The route model yielded a cross-validated accuracy of 0.78, meaning that 78 % of the time the model correctly differentiated between an actual route by a recreationist and a randomly placed route. The model for favourite places was similarly accurate regardless of the size of the availability sample: the accuracy was 0.83, 0.82, and 0.84 respectively for 1, 3, or 9 random points per used point. As the three models also were consistent in predictor effects, we concluded that an availability sample with one random point was sufficient, and present results only from that model.

To evaluate BRT model effects, the first step is to calculate a feature importance table, which ranks each predictor (i.e. "feature") according to its influence on model accuracy. The effect size is normalized across all predictors to produce a relative influence in percent for each predictor. This relative influence provides information on how important each predictor is for model accuracy, but does not inform on the specific relationship between the predictor and the response variable. To understand how the likelihood that a route or favourite place was used by a recreationist related to the value of the predictor, we produced accumulated local effects (ALE) plots, which are 2D representations of this relationship. As BRT models can have multidimensional interactions between predictors, ALE plots are valid only when the predictor is not strongly affected by such effects. This can be investigated using the H statistic, which estimates how much of a predictors relative influence is due to interactions with other predictors. In the route model, the fractions of built-up area within the viewshed and within the buffer were the only predictors showing an H statistic > 10 %. The H statistic can be decomposed in a second stage to see which other predictors that the predictor is interacting with; decomposing the H statistic for these predictors showed that they mainly interacted with each other. For the favourite place model, only the distance to recreational amenities showed H > 10 %, which when decomposed revealed only weak interactions with many other predictors in the model. The presented ALE plots (section 3.2) for individual predictors are thus mostly unaffected by interactions, and accurately depict how each predictor affects model output.

3.2 Predictors' effect on the choice of recreational location

When evaluating predictors of BRT models, a common rule of thumb is to only investigate predictors that have a relative influence larger than expected by chance, which is the inverse of the number of predictors. Due to the large number of predictors in our models (53 and 51) this cut-off was low (~1.9 %), with 23 predictors having a higher influence than the cut-off for the route model and 12 for the favourite places model. We created ALE plots for all predictors above the cut-off, but present only the interpretable ALE plots (i.e. mainly plots with a clear direction of the relationship between the variables, and for which the relationship could not easily be explained as an artifact due to confounding variables) for the two models in Figure 4 and 5, while the plots for the remaining predictors are included in supplementary materials S5.

3.2.1 Predictors affecting selection of routes

For the route model, proximity to recreational amenities (shelters, fireplaces, and toilets) had the strongest positive effect, where shorter distances increased the probability that a route was used (Figure 5). The amount of built-up area in the viewshed had a strongly negative effect, while for the amount of built-up area in the buffer, the relationship was inversely u-shaped. Path density, deciduous forest in the buffer, open area without vegetation (both buffer and viewshed), tree height, and freshwater in the viewshed were all positively related with route use. Noise and clearcuts in the viewshed showed negative correlations.

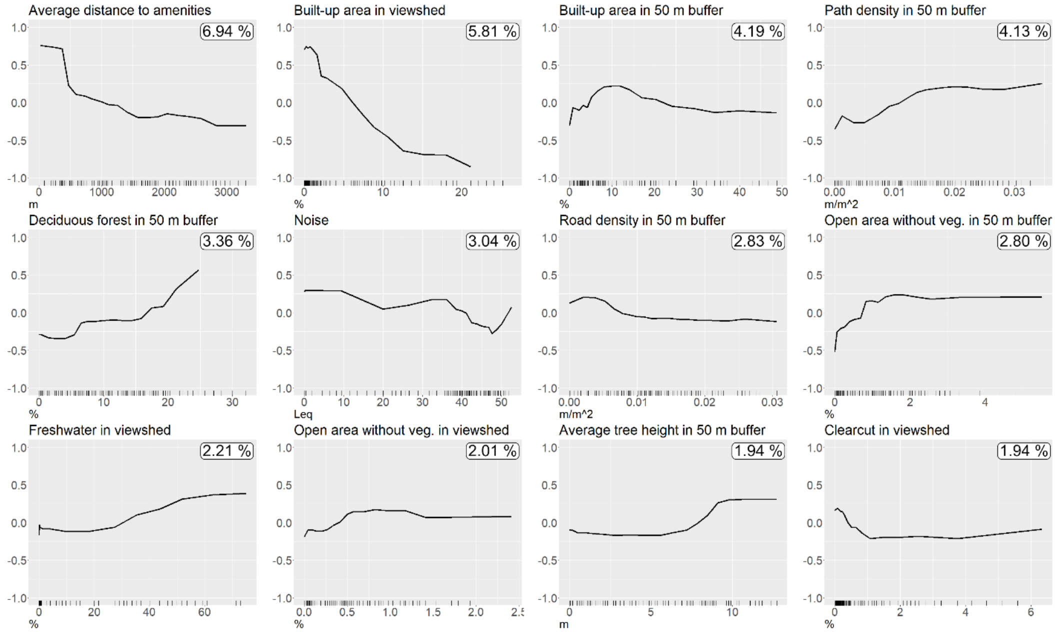


Figure 4. Accumulated local effects for 12 of the most influential predictors in a model comparing landscape characteristics of routes used by recreationists to random routes. A higher value on the y-axis represents a higher likelihood that it is a used route. The relative influence of each predictor on model outcomes is shown in the boxes in the upper right corner of each graph. Above the x-axis is a rug plot, which shows the distribution of values within the dataset, with each notch representing one percentile. The x-axis has been cut off at 95 % of the range of each variable within the dataset to remove outliers.

3.2.2 Predictors affecting selection of favourite places

In the model predicting favourite places, two predictors were considerably more influential than the others and strongly positively correlated with being a favourite place: the amount of freshwater in the viewshed, and proximity to the nearest amenity (Figure 6). Freshwater within the 50 m buffer was also positive, along with the standard deviation of elevation. In the viewshed, the fraction of sea and median elevation were positive, while the fraction built-up area and pine forest was negative. Moreover, viewshed size was positively correlated with being a favourite place.

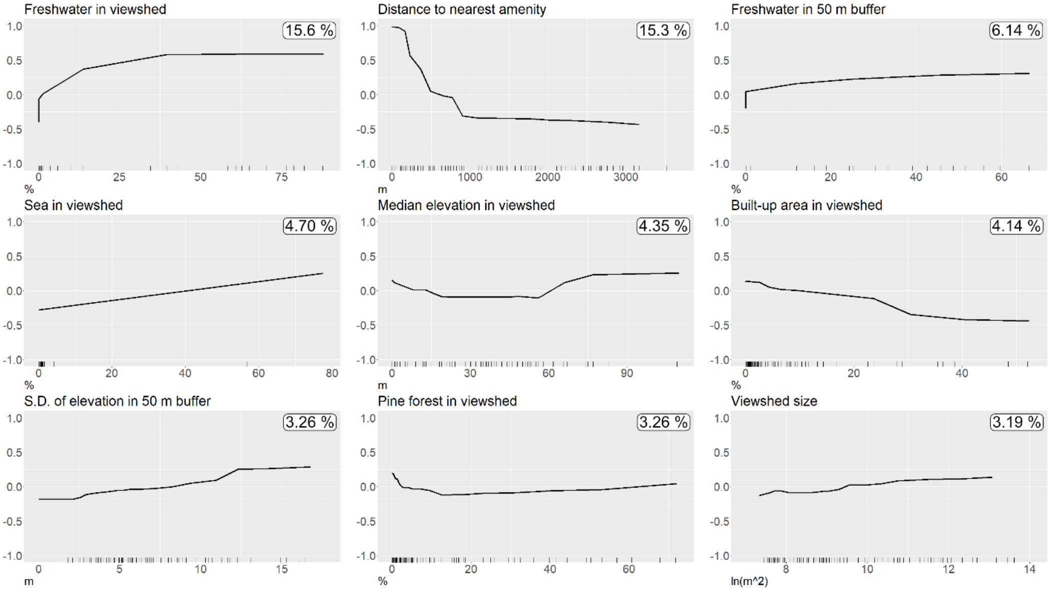


Figure 5. Accumulated local effects for the 9 most influential predictors of the model comparing landscape characteristics of favourite places stated by recreationists to random places. A higher value on the y-axis represents a higher likelihood that it is a favourite place. The relative influence of each predictor on model outcomes is shown in the boxes in the corner of each graph. Above the x-axis is a rug plot, which shows the distribution of values within the dataset, with each notch representing one percentile. The x-axis has been cut off at 95 % of the range of each variable within the dataset to remove outliers.

4. Discussion

Here we employed a novel approach to analysis of PPGIS survey data. By including viewsheds and controlling for accessibility when analyzing landscapes around favourite places and along recreational routes, we created high-accuracy models that revealed which landscape characteristics are important to recreationists. Environments with recreational infrastructure, water elements and deciduous forests were preferred, while noisy, built-up areas, young forests, and clearcuts were avoided. The analysis revealed that the routes and the favourite places had some commonalities in what features were important, showing preference for recreational infrastructure and avoidance of urban areas. There were also some differences, with the route model emphasizing forest attributes, whereas the favourite place model was more influenced by water elements and topography. Contrary to expectation, we did not find more pronounced landscape preferences in the favourite places model, with both models instead being similarly accurate.

4.1 What landscape characteristics matter for recreationists?

4.1.1 Preference for recreational infrastructure

Proximity to a shelter, fireplace, or toilet had a strong positive effect on the likelihood of an area being used by recreationists in both models. Correlations between recreational infrastructure and visits have been revealed also in previous studies (De Valck et al., 2017; Giergiczny, Czajkowski, Żylicz, & Angelstam, 2015; Kienast et al., 2012). Such a pattern can be either because the recreational infrastructure attracts visitors, or because it is built in already popular places. A study of American national parks suggested the former (Donovan, Cerveny, & Gatziolis, 2016). We argue that our results also support this view, as the large number of predictors leaves little room for the effect of recreational infrastructure to be only a proxy for other qualities in these areas. The magnitude of the effect in our study underscores the importance of incorporating recreational infrastructure in physical planning.

The density of paths and trails showed a positive effect in the route model. The role of paths and trails have rarely been studied, but Gundersen & Vistad (2016) have highlighted the importance of path quality for recreationists, finding that less developed paths (in terms of size, paving, signage etc.) were preferred in an experimental setting (rating photographs), but that more developed paths was utilized more frequently when studying actual behavior. We did not study the effect of path quality, but conclude that the occurrence of paths attract recreationists.

4.1.2 Importance of water elements

We found a strong preference for recreation close to water elements in the favourite place model. Moreover, in the route model the land cover class "Open area without vegetation" had a strong positive effect, which in the study area mostly represents rocky and sandy coastal areas. These results are consistent with previous studies of both stated (De Valck et al., 2017; Kaplan & Kaplan, 1989; White et al., 2010) and revealed preferences (Baumeister et al., 2020; Kienast et al., 2012; Laatikainen, Piiroinen, Lehtinen, & Kyttä, 2017). Our results highlight the importance of preserving water environments for recreational use. Recreational opportunities around water elements can be further improved by establishing recreational infrastructure in the vicinity or adapting management to increase viewsheds towards them.

4.1.3 Preferences for specific forest characteristics

Forest-related predictors were important in both models, but especially in the route model. Deciduous forest and taller trees had a positive effect, while clearcuts had a negative effect. The fraction of deciduous forest within the buffer around routes ranged between 0 and 30 % in the dataset, and within this range we observed a linear positive effect. A preference for deciduous forest stands has been shown in previous preference research (Gundersen et al., 2019). The observed effect may arise from an inherent preference for deciduous trees, or it could be attributed to the predominance of conifers in the boreal landscape, where the introduction of deciduous trees contributes to greater heterogeneity—which has been shown to be preferable (Filyushkina et al., 2017). Regardless of mechanism, our results show that increasing the fraction of deciduous trees increases recreational values.

The observed preference for taller trees and avoidance of clearcuts is consistent with previous studies on stated preferences (Gundersen et al., 2019). The positive effect of tree height leveled off at around 10 m, suggesting that recreationists avoid young forests formed after clearfelling, which thus typically are even-aged. This result supports the claims of higher recreational values when applying methods such as continuous cover forestry rather than even-aged forestry with clearcuts, which currently is the prevalent method in Fennoscandia (Pukkala, Lähde, & Laiho, 2012).

4.1.4 Avoidance of noise and urban areas

Both models indicated that recreationists avoid urban areas, and the route model further revealed a negative impact of noise. Noise, especially from anthropogenic sources such as traffic, have been shown to negatively affect perceptions of natural environments in experimental settings (Benfield, Rainbolt, Troup, & Bell, 2020; Li et al., 2018) and in situ (Krog, Engdahl & Tambs, 2010). A national survey conducted in Sweden found that approximately 50 % of recreationists perceived negative effects of noise during outdoor activities (Naturvårdsverket, 2019). Results such as these have spurred research on the role of what has been termed ‘soundscapes’, and has been incorporated in PPGIS methodology to map where people experience positive and negative sounds (Korpilo, Nyberg, Vierikko, Nieminen, Arciniegas & Raymond, 2023). A related concept that has been shown to be important to recreationists is ‘tranquility’, which denotes not only an absence of noise, but also how restorative a landscape is perceived (Purves & Wartmann, 2023). In a large Danish study where recreationists were asked to map ‘good locations’ they experienced along their walking routes, 40 % of these were described as having a tranquil quality (Christiansen, Top Klein-Wengel, Koch, Høyer-Kruse & Schipperijn, 2023). Our result shows that noise also affects actual landscape usage, i.e. that recreationists choose areas that have less noise. Our findings also demonstrate the utility of spatial noise modelling, and underscores the recent efforts made to map and protect “quiet areas” (Cerwén & Mossberg, 2019).

4.1.5 Preference for a varied topography and viewshed size

The influence of predictors related to elevation and viewshed size revealed that topography was important: people preferred a landscape of varying height that also yielded a large view, but avoided views of urban areas or clearcuts. We found that both a high elevation and a low elevation were positive, which we interpret as representing both a preference for height and for close-to-sea areas, both of which yield large views with long sight lines. Earlier studies have also found a preference for views (Gundersen et al., 2019; Kaplan & Kaplan, 1989; Kienast et al., 2012). Thus, views is an important aspect to consider in landscape planning for recreation, and viewshed analysis is a possible route to identify them. Existing views can be improved by management, for instance by opening up views towards water elements, obscuring views towards buildings, or through the construction of lookouts.

4.2 Are recreationists’ preferences homogeneous?

Our models had very weak interaction effects, implying that characteristics of the recreationists did not influence which landscape characteristics they sought out (Fig. 4). This was also the case for predictors related to the activity (e.g. type of activity, time spent on location, frequency of visit) or the season (winter/summer). In contrast, other studies have provided evidence for the influence of socio-demographic factors on landscape preference and utilization. Kienast et al. (2012) found that older people tended to visit places with more distinct characteristics compared to younger people. However, socio-demographic characteristics appear to have a weaker explanatory power in determining landscape preferences compared to environmental attitudes, nature relatedness or ideology (Eriksson et al., 2012; Juutinen, Kosenius, Ovaskainen, Tolvanen, & Tyrväinen, 2017; Ode Sang, Knez, Gunnarsson, & Hedblom, 2016; Scott et al., 2009). This could explain the weak effects seen here, as we only included one questions on attitudes and ideology, namely on to what degree the respondent identified as an “urban person” and “nature person”.

The type of activity has also been found to affect preferences and behaviors (De Valck et al., 2016, 2017; Korpilo et al., 2017). However, only some variation related to preferences related to characteristics of the landscape have been observed (De Valck et al., 2016), and the main effects appear to be driven by preferences for different types of recreational infrastructure (Abildtrup, Garcia, Olsen, & Stenger, 2013; De Valck et al., 2017). User typology, i.e. defining archetypes of recreationists related to their typical patterns of recreational use (e.g. preferred activity, willingness to travel, visit frequencies etc.), has been suggested as an approach to analyze the heterogeneous preferences of recreationists (Komossa et al., 2019). As the BRT models employed here can handle multiple predictors interacting concurrently, our methodology should be able to identify such user groups. Yet, the results did not identify such groups, implying that preferences for the landscape characteristics we used as predictors do not vary, or only vary a little, between user types.

Interestingly, we did not see any difference in preference for landscape characteristics between winter and summer recreation. Seasonal effects on outdoor recreation have rarely been studied, but a small study in Utah revealed that the winter landscape was perceived as drastically different, and that recreational patterns and experiences changed (Gatti, Brownlee, & Bricker, 2022). Moreover, a study on tourists' perceptions in Finland showed that forest characteristics were less important when snow was present, whereas the presence of long sight lines was more important (Tyrväinen, Silvennoinen, & Hallikainen, 2017). Our lack of pattern was surprising, given that our study area is dramatically changed in winter, with most water elements freezing, deciduous trees losing their leaves, and large quantities of snow blanketing the landscape, making some types of activities possible (e.g. cross-country skiing and ice skating) while others become more difficult (e.g. cycling). A possible issue with our methodology was the timing of survey deployment: since data was collected in September, summer recreation would have been easier to recall for the respondents.

4.3 Improved methodology for PPGIS landscape preference analysis

Here, we further PPGIS research through a novel combination of three approaches: firstly, through the calculation of viewsheds at the locations of recreation; secondly by performing a network analysis to define what landscape was available to each recreationist; and thirdly via the inclusion of flexible machine learning methods.

Viewshed analysis has been employed previously in outdoor recreation research, such as when modeling the aesthetic value of landscapes using crowdsourced photographs (Karasov et al., 2020; Tenerelli, Püffel, & Luque, 2017; Yoshimura & Hiura, 2017). Here we included viewsheds to estimate what each recreationist experienced. Viewshed analysis has some prerequisites, such as a digital surface model (DSM) with high spatial accuracy (Lagner, Klouček, & Šímová, 2018), and it has relatively high demands on computing power. We believe that continued development of viewshed analysis in recreation research will yield results closer to the ground truth, and recommend researchers to experiment with possible implementations.

Most PPGIS studies we could find have not properly controlled for accessibility (e.g. Kienast et. al 2012, Baumeister et. al 2020, De Valck et. al 2016), or have done so only to a certain extent (Agimass, Lundhede, Panduro, & Jacobsen, 2018). We believe our approach of using network analysis (see 2.2.1, Fig. 2) is a good solution to this issue.

Machine learning has been suggested to be particularly useful in ecosystem service research (Scowen, Athanasiadis, Bullock, Eigenbrod & Willcock, 2021). The modelling performed here, using boosted regression trees, is not novel in itself, (Friedman, 2001), but as with most innovations in statistical methods, adoption by researchers is slow (Sharpe, 2013). Our results here are a showcase for how this type of modelling can be advantageous compared to traditional alternatives such as GLMs. Here, we had few prior hypotheses on which landscape characteristics to be most predictive, and for many predictors non-linear relationships were expected. Since collinearity of predictors is not an issue for model fitting, and there is no need for model selection or pre-specifying interactions, we were able to add all available map data that could be relevant to the analysis. This type of modelling is thus very useful for exploratory studies. The main drawback is that the models can be harder to interpret than e.g. GLMs, but with newer tools (e.g. the *iml* R package used here; Molnar, 2018) these issues can be overcome.

We can compare the outcome from this study with our previous study on Swedish recreationists (Lehto et al., 2022). That study analyzed a spatial dataset on recreation in a similar manner to here, but without letting the respondents draw full routes, without viewsheds, and with less adequate control for what landscape was accessible to the recreationist, which resulted in much weaker models despite a much larger sample size.

The response rate of our survey was rather low, with 20 % starting the survey and 9 % filling it out in full. Decreasing response rates to surveys is a trend (Stedman 2019), especially for web-based surveys (Daikeler 2020). On the other hand, surveys with a strong local connection, as here, usually have higher response rates (Stedman 2019). There might be a degree of self-selection bias in that people who were more interested in outdoor recreation chose to finish the survey to a higher degree. However, our sample relatively closely matched the population under study regarding age and gender, so for those characteristics we were able to compare, our sample was representative for the population as a whole.

5. Conclusions

Our study has successfully developed new methods for studying revealed preferences among recreationists. We have improved on existing PPGIS methods of outdoor recreation in three ways. Firstly, we included viewsheds, which brings the analysis closer to the recreationists' experience by attempting to see the landscape through their eyes. Secondly, we controlled for accessibility by using network analysis, to properly compare the area used by the recreationist with an area that was equally accessible. Thirdly we employed flexible machine learning methods, capable of handling a large number of map covariates.

Our results yield actionable results on recreation. Water environments, recreational infrastructure, and deciduous forests were selected for, while urban environments, noise, forest clearcuts, and young forest were avoided. These outcomes suggest that increased recreational infrastructure could improve the conditions for recreation, especially in proximity to water. This also ties into policy, showing the importance of providing access to and hindering the exploitation of water environments and minimizing noise pollution. To manage forests for recreation, deciduous trees should be favored and clearcuts avoided, whereas felling trees to create viewsheds toward water could be positive for recreation.

References

Abildtrup, J., Garcia, S., Olsen, S. B., & Stenger, A. (2013). Spatial preference heterogeneity in forest recreation. *Ecological Economics*, *92*, 67–77. doi - 10.1016/j.ecolecon.2013.01.001

Agimass, F., Lundhede, T., Panduro, T. E., & Jacobsen, J. B. (2018). The choice of forest site for recreation: A revealed preference analysis using spatial data. *Ecosystem Services, 31*, 445–454. doi - 10.1016/j.ecoser.2017.11.016

Baumeister, C. F., Gerstenberg, T., Plieninger, T., & Schraml, U. (2020). Exploring cultural ecosystem service hotspots: Linking multiple urban forest features with public participation mapping data. *Urban Forestry & Urban Greening*, *48*, 126561. doi - 10.1016/j.ufug.2019.126561

Benfield, J. A., Rainbolt, G. A. N., Troup, L. J., & Bell, P. A. (2020). Anthropogenic noise source and intensity effects on mood and relaxation in simulated park environments. *Frontiers in Psychology*, *11*. Retrieved from https://www.frontiersin.org/articles/10.3389/fpsyg.2020.570694

Brown, G. (2012). An empirical evaluation of the spatial accuracy of public participation GIS (PPGIS) data. *Applied Geography*, *34*, 289–294. doi - 10.1016/j.apgeog.2011.12.004

Brown, G., & Fagerholm, N. (2015). Empirical PPGIS/PGIS mapping of ecosystem services: A review and evaluation. *Ecosystem Services*, *13*, 119–133. doi - 10.1016/j.ecoser.2014.10.007

Cerwén, G., & Mossberg, F. (2019). Implementation of Quiet Areas in Sweden. *International Journal of Environmental Research and Public Health*, *16*(1), 134. doi - 10.3390/ijerph16010134

Christiansen, L. B., Klein-Wengel, T. T., Koch, S., Høyer-Kruse, J., & Schipperijn, J. (2023). Recreational walking and perceived environmental qualities: a national map-based survey in Denmark. *International Journal of Health Geographics*, *22*(1), 21.

Daikeler, J., Bošnjak, M., & Lozar Manfreda, K. (2020). Web versus other survey modes: an updated and extended meta-analysis comparing response rates. *Journal of Survey Statistics and Methodology, 8*(3), 513-539.

De Valck, J., & Rolfe, J. (2018). Spatial heterogeneity in stated preference valuation: Status, challenges and road ahead. *International Review of Environmental and Resource Economics, 11(4)*, 355-422.

De Valck, J., Broekx, S., Liekens, I., De Nocker, L., Van Orshoven, J., & Vranken, L. (2016). Contrasting collective preferences for outdoor recreation and substitutability of nature areas using hot spot mapping. *Landscape and Urban Planning*, *151*, 64–78. doi - 10.1016/j.landurbplan.2016.03.008

De Valck, J., Landuyt, D., Broekx, S., Liekens, I., De Nocker, L., & Vranken, L. (2017). Outdoor recreation in various landscapes: Which site characteristics really matter? *Land Use Policy*, *65*, 186–197. doi - 10.1016/j.landusepol.2017.04.009

Donovan, G. H., Cerveny, L. K., & Gatziolis, D. (2016). If you build it, will they come? *Forest Policy and Economics*, *62*, 135–140. doi - 10.1016/j.forpol.2015.11.002

Ekkel, E. D., & de Vries, S. (2017). Nearby green space and human health: Evaluating accessibility metrics. *Landscape and Urban Planning*, *157*, 214–220. doi - 10.1016/j.landurbplan.2016.06.008

Elbakidze, M., Dawson, L., Milberg, P., Mikusiński, G., Hedblom, M., Kruhlov, I., … Grodzynskyi, M. (2022). Multiple factors shape the interaction of people with urban greenspace: Sweden as a case study. *Urban Forestry & Urban Greening*, *74*, 127672. doi - 10.1016/j.ufug.2022.127672

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*(4), 802–813. doi - 10.1111/j.1365-2656.2008.01390.x

Eriksson, L., Nordlund, A., Olsson, O., & Westin, K. (2012). Beliefs about urban fringe forests among urban residents in Sweden. *Urban Forestry & Urban Greening*, *11*(3), 321–328. doi - 10.1016/j.ufug.2012.02.004

Filyushkina, A., Agimass, F., Lundhede, T., Strange, N., & Jacobsen, J. B. (2017). Preferences for variation in forest characteristics: Does diversity between stands matter? *Ecological Economics*, *140*, 22–29. doi - 10.1016/j.ecolecon.2017.04.010

Fish, R., Church, A., & Winter, M. (2016). Conceptualising cultural ecosystem services: A novel framework for research and critical engagement. *Ecosystem Services*, *21*, 208-217.

Flowers, E. P., Freeman, P., & Gladwell, V. F. (2016). A cross-sectional study examining predictors of visit frequency to local green space and the impact this has on physical activity levels. *BMC Public Health*, *16*(1), 420. doi - 10.1186/s12889-016-3050-9

Frick, J., Degenhardt, B., & Buchecker, M. (2007). Predicting local residents' use of nearby outdoor recreation areas through quality perceptions and recreational expectations. *Forest Snow and Landscape Research*, *81*(1-2), 31-41.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5). doi - 10.1214/aos/1013203451

Gatti, E. T. J., Brownlee, M. T. J., & Bricker, K. S. (2022). Winter recreationists' perspectives on seasonal differences in the outdoor recreation setting. *Journal of Outdoor Recreation and Tourism*, *37*, 100366. doi - 10.1016/j.jort.2021.100366

Gerstenberg, T., Baumeister, C. F., Schraml, U., & Plieninger, T. (2020). Hot routes in urban forests: The impact of multiple landscape features on recreational use intensity. *Landscape and Urban Planning*, *203*, 103888. doi - 10.1016/j.landurbplan.2020.103888

Giergiczny, M., Czajkowski, M., Żylicz, T., & Angelstam, P. (2015). Choice experiment assessment of public preferences for forest structural attributes. *Ecological Economics*, *119*, 8–23. doi - 10.1016/j.ecolecon.2015.07.032

Gosal, A. S., Giannichi, M. L., Beckmann, M., Comber, A., Massenberg, J. R., Palliwoda, J., … Ziv, G. (2021). Do drivers of nature visitation vary spatially? The importance of context for understanding visitation of nature areas in Europe and North America. *Science of The Total Environment*, *776*, 145190. doi - 10.1016/j.scitotenv.2021.145190

Grahn, P., & Stigsdotter, U. A. (2003). Landscape planning and stress. *Urban Forestry & Urban Greening*, *2*(1), 1–18. doi - 10.1078/1618-8667-00019

Gundersen, V., Köhler, B., & Myrvold, K. M. (2019). Seeing the forest for the trees: a review-based framework for better harmonization of timber production, biodiversity, and recreation in boreal urban forests. *Urban Science*, *3*(4), 113. doi - 10.3390/urbansci3040113

Gundersen, V., & Vistad, O. (2016). Public opinions and use of various types of recreational infrastructure in boreal forest settings. *Forests*, *7*(12), 113. doi - 10.3390/f7060113

Gunnarsson, B., Knez, I., Hedblom, M., & Sang, Å. O. (2017). Effects of biodiversity and environment-related attitude on perception of urban green space. *Urban Ecosystems*, *20*(1), 37–49. doi - 10.1007/s11252-016-0581-x

Haines-Young, R., & Potschin-Young, M. (2018). Revision of the Common International Classification for Ecosystem Services (CICES V5.1): A policy brief. *One Ecosystem*, *3*, e27108. doi - 10.3897/oneeco.3.e27108

Hörnsten, L. (2000). *Outdoor recreation in Swedish forests: implications for society and forestry* (Swedish University of Agricultural Sciences). Swedish University of Agricultural Sciences, Uppsala. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:slu:epsilon-e-5406

Juutinen, A., Kosenius, A.-K., Ovaskainen, V., Tolvanen, A., & Tyrväinen, L. (2017). Heterogeneous preferences for recreation-oriented management in commercial forests: the role of citizens' socioeconomic characteristics and recreational profiles. *Journal of Environmental Planning and Management*, *60*(3), 399–418. doi - 10.1080/09640568.2016.1159546

Kajosaari, A., Hasanzadeh, K., Fagerholm, N., Nummi, P., Kuusisto-Hjort, P., & Kyttä, M. (2024). Predicting context-sensitive urban green space quality to support urban green infrastructure planning. *Landscape and Urban Planning*, *242*, 104952.

Kaplan, R., & Kaplan, S. (1989). *The Experience of Nature: A Psychological Perspective*. CUP Archive.

Karasov, O., Vieira, A. A. B., Külvik, M., & Chervanyov, I. (2020). Landscape coherence revisited: GIS-based mapping in relation to scenic values and preferences estimated with geolocated social media data. *Ecological Indicators*, 111, 105793.

Kearney, A. R., & Bradley, G. A. (2011). The effects of viewer attributes on preference for forest scenes: contributions of attitudes, knowledge, demographic factors, and stakeholder group membership. *Environment and Behavior*, *43*(2), 147–181. doi - 10.1177/0013916509353523

Kienast, F., Degenhardt, B., Weilenmann, B., Wäger, Y., & Buchecker, M. (2012). GIS-assisted mapping of landscape suitability for nearby recreation. *Landscape and Urban Planning*, *105*(4), 385–399. doi - 10.1016/j.landurbplan.2012.01.015

Komossa, F., van der Zanden, E. H., & Verburg, P. H. (2019). Characterizing outdoor recreation user groups: A typology of peri-urban recreationists in the Kromme Rijn area, the Netherlands. *Land Use Policy*, *80*, 246–258. doi - 10.1016/j.landusepol.2018.10.017

Korpilo, S., Virtanen, T., & Lehvävirta, S. (2017). Smartphone GPS tracking—Inexpensive and efficient data collection on recreational movement. *Landscape and Urban Planning*, *157*, 608–617. doi - 10.1016/j.landurbplan.2016.08.005

Korpilo, S., Nyberg, E., Vierikko, K., Nieminen, H., Arciniegas, G., & Raymond, C. M. (2023). Developing a Multi-sensory Public Participation GIS (MSPPGIS) method for integrating landscape values and soundscapes of urban green infrastructure. *Landscape and Urban Planning*, *230*, 104617. https://doi.org/10.1016/j.landurbplan.2022.104617

Krog, N. H., Engdahl, B., & Tambs, K. (2010). Effects of changed aircraft noise exposure on experiential qualities of outdoor recreational areas. *International Journal of Environmental Research and Public Health*, *7*(10), 3739–3759. doi - 10.3390/ijerph7103739

Laatikainen, T., Piiroinen, R., Lehtinen, E., & Kyttä, M. (2017). PPGIS approach for defining multimodal travel thresholds: Accessibility of popular recreation environments by the water. *Applied Geography*, *79*, 93–102. doi - 10.1016/j.apgeog.2016.12.006

Lagner, O., Klouček, T., & Šímová, P. (2018). Impact of input data (in)accuracy on overestimation of visible area in digital viewshed models. *PeerJ*, *6*, e4835. doi - 10.7717/peerj.4835

Lantmäteriet. (2023). *National Elevation Model*. Retrieved from https://www.lantmateriet.se/en/geodata/geodata-products/product-list/terrain-model-download-grid-1/

Lehto, C., Hedblom, M., Öckinger, E., & Ranius, T. (2022). Landscape usage by recreationists is shaped by availability: Insights from a national PPGIS survey in Sweden. *Landscape and Urban Planning*, *227*, 104519. doi - 10.1016/j.landurbplan.2022.104519

Li, J., Burroughs, K., Halim, M. F., Penbrooke, T. L., Seekamp, E., & Smith, J. W. (2018). Assessing soundscape preferences and the impact of specific sounds on outdoor recreation activities using qualitative data analysis and immersive virtual environment technology. *Journal of Outdoor Recreation and Tourism*, *24*, 66–73. doi - 10.1016/j.jort.2018.08.001

Molnar, C., Casalicchio, G., & Bischl, B. (2018). iml: An R package for interpretable machine learning. *Journal of Open Source Software*, *3*(26), 786.

Naturvårdsverket. (2019). *Friluftsliv 2018: nationell undersökning av svenska folkets friluftslivsvanor* (No. 6887). Stockholm: Naturvårdsverket. Retrieved from Naturvårdsverket website: https://www.naturvardsverket.se/978-91-620-6887-5

Neuvonen, M., Sievänen, T., Tönnes, S., & Koskela, T. (2007). Access to green areas and the frequency of visits – A case study in Helsinki. *Urban Forestry & Urban Greening*, *6*(4), 235–247. doi - 10.1016/j.ufug.2007.05.003

Nguyen, P. Y., Astell-Burt, T., Rahimi-Ardabili, H., & Feng, X. (2021). Green space quality and health: a systematic review. *International journal of environmental research and public health*, *18*(21), 11028.

Nisbet, E. K., Zelenski, J. M., & Murphy, S. A. (2009). The nature relatedness scale: Linking individuals' connection with nature to environmental concern and behavior. *Environment and Behavior*, *41*(5), 715–740. doi - 10.1177/0013916508318748

Northrup, J. M., Hooten, M. B., Anderson, C. R., & Wittemyer, G. (2013). Practical guidance on characterizing availability in resource selection functions under a use–availability design. *Ecology*, *94*(7), 1456–1463. doi - 10.1890/12-1688.1

Ode Sang, Å., Knez, I., Gunnarsson, B., & Hedblom, M. (2016). The effects of naturalness, gender, and age on how urban green space is perceived and used. *Urban Forestry & Urban Greening*, *18*, 268–276. doi - 10.1016/j.ufug.2016.06.008

Paracchini, M. L., Zulian, G., Kopperoinen, L., Maes, J., Schägner, J. P., Termansen, M., … Bidoglio, G. (2014). Mapping cultural ecosystem services: A framework to assess the potential for outdoor recreation across the EU. *Ecological Indicators*, *45*, 371–385. doi - 10.1016/j.ecolind.2014.04.018

Pukkala, T., Lähde, E., & Laiho, O. (2012). Continuous cover forestry in Finland – Recent research results. In T. Pukkala & K. von Gadow (Eds.), *Continuous Cover Forestry* (pp. 85–128). Dordrecht: Springer Netherlands. doi - 10.1007/978-94-007-2202-6_3

Purves, R. S., & Wartmann, F. M. (2023). Characterising and mapping potential and experienced tranquillity: From a state of mind to a cultural ecosystem service. *Geography Compass*, *17*(11), e12726.

R Core Team. (2020). *R: a language and environment for Statistical Computing*. Vienna, Austria. Retrieved from https://www.R-project.org/

Richards, D. R., & Belcher, R. N. (2020). Global changes in urban vegetation cover. *Remote Sensing*, *12*(1), 23. doi - 10.3390/rs12010023

SCB. (2021). Utbildningsnivån i Sverige. Retrieved 7 July 2023, from Statistiska Centralbyrån website: https://www.scb.se/hitta-statistik/sverige-i-siffror/utbildning-jobb-och-pengar/utbildningsnivan-i-sverige/

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

Scott, A., Carter, C., Brown, K., & White, V. (2009). 'Seeing is not everything': Exploring the landscape experiences of different publics. *Landscape Research*, *34*(4), 397–424. doi - 10.1080/01426390903009289

Scowen, M., Athanasiadis, I. N., Bullock, J. M., Eigenbrod, F., & Willcock, S. (2021). The current and future uses of machine learning in ecosystem service research. *Science of the Total Environment*, *799*, 149263.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27(3)*, 379-423.

Sharpe, D. (2013). Why the resistance to statistical innovations? Bridging the communication gap. *Psychological Methods*, *18*(4), 572–582. doi - 10.1037/a0034177

SLU. (2015). *SLU Forest map*. Retrieved from https://www.slu.se/en/Collaborative-Centres-and-Projects/the-swedish-national-forest-inventory/foreststatistics/slu-forest-map/

Stedman, R. C., Connelly, N. A., Heberlein, T. A., Decker, D. J., & Allred, S. B. (2019). The end of the (research) world as we know it? Understanding and coping with declining response rates to mail surveys. *Society & Natural Resources*, *32*(10), 1139-1154.

Swedish Environmental Protection Agency. (2023). CadasterENV Sweden. https://www.naturvardsverket.se/verktyg-och-tjanster/kartor-och-karttjanster/nationella-marktackedata/

Tenerelli, P., Püffel, C., & Luque, S. (2017). Spatial assessment of aesthetic services in a complex mountain region: combining visual landscape properties with crowdsourced geographic information. *Landscape Ecology*, *32*(5), 1097–1115. doi - 10.1007/s10980-017-0498-7

Timonen, J., Siitonen, J., Gustafsson, L., Kotiaho, J. S., Stokland, J. N., Sverdrup-Thygeson, A., & Mönkkönen, M. (2010). Woodland key habitats in northern Europe: concepts, inventory and protection. *Scandinavian Journal of Forest Research*, *25*(4), 309–324. doi - 10.1080/02827581.2010.497160

Tyrväinen, L., Silvennoinen, H., & Hallikainen, V. (2017). Effect of the season and forest management on the visual quality of the nature-based tourism environment: a case from Finnish Lapland. *Scandinavian Journal of Forest Research*, *32*(4), 349–359. doi - 10.1080/02827581.2016.1241892

Umeå municipal government. (2016). Umeå Municipality. Retrieved 7 July 2023, from https://www.umea.se/byggaboochmiljo/boendemiljobullerochluftkvalitet/buller/bullerberakningar.4.250f9659174ae4b97941386.html

Umeå municipal government. (2022). Umeå municipality. Retrieved 7 July 2023, from https://www.umea.se/kommunochpolitik/kommunfakta/statistikochanalyser/befolkningochbefolkningsprognoser.4.2bd9ced91726ea4d7b49.html van Zanten, B. T., Verburg, P. H., Koetse, M. J., & van Beukering, P. J. H. (2014). Preferences for European agrarian landscapes: A meta-analysis of case studies. *Landscape and Urban Planning, 132*, 89–101. doi - 10.1016/j.landurbplan.2014.08.012

Walz, U., & Stein, C. (2018). Indicator for a monitoring of Germany's landscape attractiveness. *Ecological Indicators, 94,* 64–73. doi - 10.1016/j.ecolind.2017.06.052

White, M., Smith, A., Humphryes, K., Pahl, S., Snelling, D., & Depledge, M. (2010). Blue space: The importance of water for preference, affect, and restorativeness ratings of natural and built scenes. *Journal of Environmental Psychology, 30*(4), 482–493. doi - 10.1016/j.jenvp.2010.04.004

Yoshimura, N., & Hiura, T. (2017). Demand and supply of cultural ecosystem services: Use of geotagged photos to map the aesthetic value of landscapes in Hokkaido. *Ecosystem Services*, *24*, 68–78. doi - 10.1016/j.ecoser.2017.02.009

Supplementary materials

S1. English translation of the survey

Page 1: Introductory page

Survey on recreational activities from the Swedish University of Agricultural Sciences

Hi!
In this survey, you will have the opportunity to summarize the locations in the landscape that you use for outdoor recreation by drawing typical routes you usually take within Umeå municipality. You will first be asked to draw your outdoor recreation during the summer, and then draw it during the winter, while also marking particularly important places. The goal of this research is to gain knowledge about the characteristics that are important for an area to be utilized for outdoor recreation.
Outdoor recreation in this survey refer to all types of activities you engage in outdoors during your leisure time, both in natural environments and urban settings.
The survey is completely anonymous and cannot be linked to you as an individual. If you have any concerns about how we handle the data you provide in the survey, you can download our data protection policy using the button below.
Click on the right arrow at the bottom to start the survey, which is expected to take approximately fifteen minutes.
[Button that leads to the data protection policy]
[Button that leads to page 2]

Page 2: Demographics etc.

A few questions about you

We start with a few questions about you.
How old are you?
Which gender do you identify as?

What is your highest completed level of education?
To what degree would you identify yourself as a nature-oriented person?
To what degree would you identify yourself as an urban-oriented person?
[Button that leads to page 1] [Button that leads to page 3]

Page 3: Summer recreation

Outdoor recreation in summer

To the right, you see a map (if you are taking the survey on a mobile device, you will only see this text). The map is currently centered over Umeå, but you can move it by clicking and dragging. You can zoom in and out using the plus and minus buttons in the upper left corner, or by using the scroll wheel on your mouse. To center the map on a specific address, click on the magnifying glass. You can also switch between map and satellite view by clicking on the layers button in the upper right corner.
You are now asked to indicate the locations you visit when engaging in outdoor activities during the summer by drawing a typical route you usually take for each place. Click on the "Draw Route" button below (you may need to scroll down to see the button), and then click on the map to start drawing. Begin drawing where you feel your outdoor activity begins, so do not include what you perceive as a transportation route to the place you visit. Draw as accurately as you can.
After each route you draw, you will be asked a few questions about how you use this location. You can draw as many routes as you like, but do not draw multiple routes that pertain to the same location. If you take different routes on different visits to a certain location, draw a typical route or the most recent route you took at the place.
When you feel you have finished drawing, click on the right arrow at the bottom to proceed.
[Button that reads "draw route"]
[Button that leads to page 2] [Button that leads to page 4]

Page 4: Place of origin

Place of origin

On this page you mark the location where you most often start out from when you're beginning your outdoor recreation. If this point is your home and you are uncomfortable with sharing your exact home location, you can place the point at an approximate position as close to your home as you are comfortable.
You can only place one place of origin.
[Button that reads "Place of origin"]
[Button that leads to page 3] [Button that leads to page 5]

Page 5: Winter recreation

Outdoor recreation in winter

To the right, you see a map (if you are taking the survey on a mobile device, you will only see this text). The map is currently centered over Umeå, but you can move it by clicking and dragging. You can zoom in and out using the plus and minus buttons in the upper left corner, or by using the scroll wheel on your

mouse. To center the map on a specific address, click on the magnifying glass. You can also switch between map and satellite view by clicking on the layers button in the upper right corner.
You are now asked to indicate the locations you visit when engaging in outdoor activities during the winter by drawing a typical route you usually take for each place. Click on the "Draw Route" button below (you may need to scroll down to see the button), and then click on the map to start drawing. Begin drawing where you feel your outdoor activity begins, so do not include what you perceive as a transportation route to the place you visit. Draw as accurately as you can.
After each route you draw, you will be asked a few questions about how you use this location. You can draw as many routes as you like, but do not draw multiple routes that pertain to the same location. If you take different routes on different visits to a certain location, draw a typical route or the most recent route you took at the place.
When you feel you have finished drawing, click on the right arrow at the bottom to proceed.
[Button that reads "draw route"]
[Button that leads to page 4] [Button that leads to page 6]

Page 6: Favourite places

Favourite places

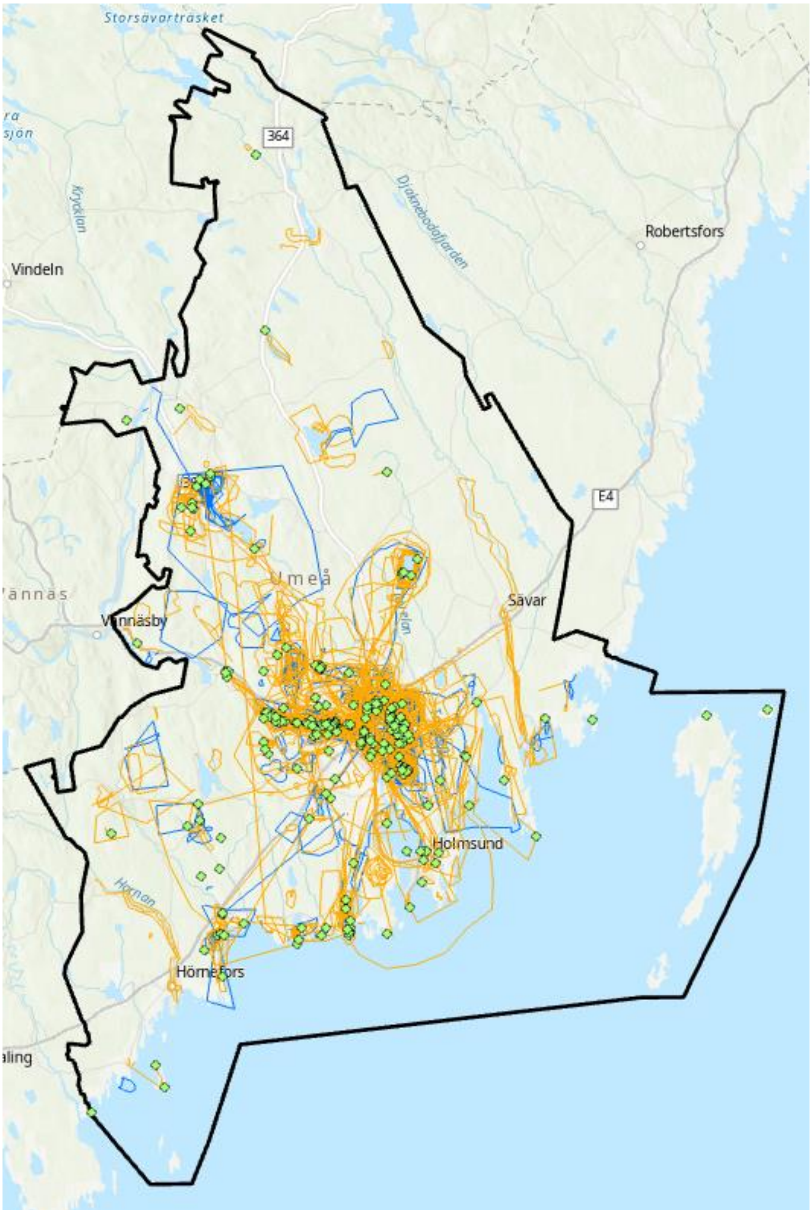
If you have specific locations along the routes you previously drew that are particularly important to you, for example, places you find exceptionally beautiful or where you tend to stop for a while for some reason, you can mark these with the help of the "Favorite Place" button. After clicking on a location, you will be asked a few questions about your experience of that place.
When you feel you have finished drawing, click on the right arrow at the bottom to proceed to the end of the survey.
[Button that reads "Favourite place"]
[Button that leads to page 5] [Button that leads to page 7]

Page 7: End of survey

Thank you!

By clicking "Done!" at the bottom, you consent to the collected information being used in accordance with our privacy policy (which can be accessed by clicking the button below), and your responses will be submitted.
The survey results will be compiled in a scientific article, available on this website. It is expected to be completed sometime in 2022.
Thank you very much for helping us with our research!
[Button that leads to the Data protection policy]
[Button that reads "Done!"]

S2. Visualization of the dataset



Green dots are favourite places, orange lines are routes in summer, blue lines are routes in winter. Black outline is Umeå municipality. North is up, for map scale see figure 1. Map projection is SWEREF 99TM.

S3: Reclassification of CadasterENV

List of the land cover classes of the CadasterENV map and the reclassification used in all analyses. Non-vegetated is mostly represents by rocky/sandy beaches in the study area. Vegetated other open land consists mainly of lawns in urban areas, road verges, and some meadows/pasture land, as long as they are not part of an agricultural rotation which would classify them as arable land.

Class code	Class name	Reclassification
2	Open wetland	Wetland
3	Arable land	Arable land
41	Non-vegetated other open land	Non-vegetated other open land
42	Vegetated other open land	Vegetated other open land
51	Buildings	Built-up area
52	Artificial non-vegetated surface	Built-up area
53	Roads or railways	Built-up area
61	Inland waters	Inland waters
62	Marine waters	Marine waters
111	Pine forest not on wetlands	Pine forest
112	Spruce forest not on wetlands	Spruce forest
113	Mixed coniferous forest not on wetlands	Mixed forest
114	Mixed forest not on wetlands	Mixed forest
115	Deciduous forest	Deciduous forest
116	Temperate deciduous forest not on wetlands	Temperate deciduous forest
117	Deciduous forest with temperate deciduous forest not on wetland	Temperate deciduous forest
118	Temporarily non-forest not on wetland	Clearcut
121	Pine forest on wetland	Wetland
122	Spruce forest on wetland	Wetland
123	Mixed coniferous forest on wetland	Wetland
124	Mixed forest on wetland	Wetland
125	Deciduous forest on wetland	Wetland
126	Temperate deciduous forest on wetland	Wetland
127	Deciduous forest with temperate deciduous forest on wetland	Wetland
128	Temporarily non-forest	Clearcut

S4: BRT modelling

Boosted Regression Trees modelling is dependent on setting four main hyperparameters (settings) before fitting a model. These parameters control the learning process and significantly affect the model's performance. The hyperparameters are:

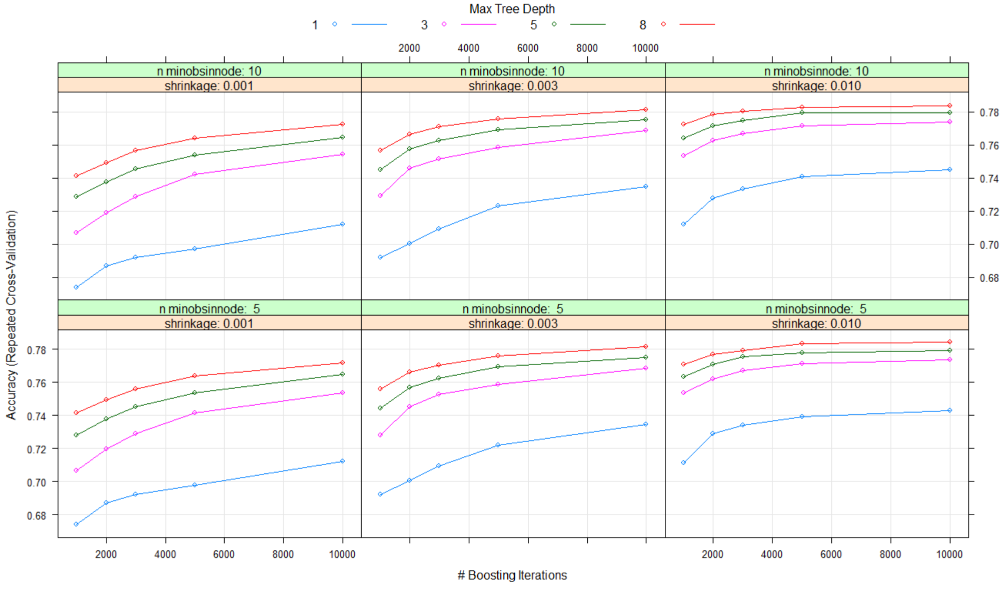
- Learning rate (also known as shrinkage): Determines how quickly the model learns. A smaller learning rate requires more trees to model all the relationships, potentially leading to a more complex model.
- Number of trees: The total count of sequential trees to be built. More trees can capture more complex patterns but also risk overfitting.
- Tree depth (also called interaction depth): Specifies the maximum depth of each tree. Deeper trees can model more complex relationships (higher dimensions of interactions) but also may overfit the data.

- Minimum samples in leaf nodes: The minimum number of samples a leaf node must have. This parameter can help prevent overfitting by ensuring that each leaf node has a sufficient number of observations.

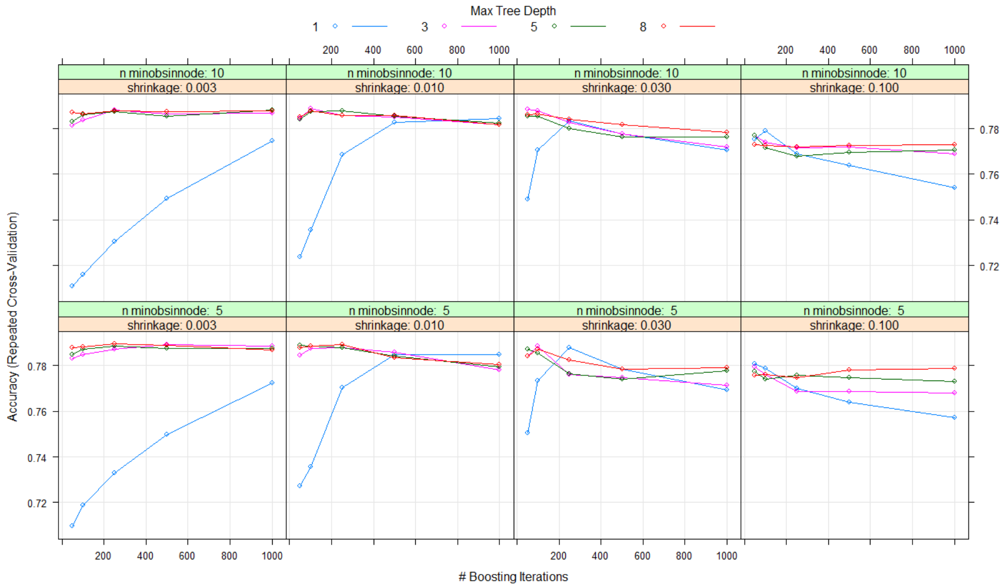
A grid search for the best hyperparameters involves systematically working through multiple combinations of values for these hyperparameters, fitting a model for each combination, and then evaluating their performance using a predefined metric (here using 10-fold cross-validated accuracy). The grid search is computationally expensive but ensures that the best set of parameters within the defined grid is found for the given modeling task.
Which hyperparameter settings are optimal is related to dataset size and number of features, with learning rates typically needed to be set lower and number of trees higher for larger datasets. We tested the following combinations of hyperparameters for the two models:

Hyperparameter	**Route**	**Favourite places**
Learning rate	0.001, 0.003, 0.01	0.1, 0.03, 0.01, 0.003
Tree depth	1, 3, 5, 8	1, 3, 5, 8
Number of trees	1000, 2000, 3000, 5000, 10000	100, 250, 500, 1000
Minimum samples in leaf nodes	5, 10	5, 10

Below are figures showing the cross-validated accuracy of models with different hyperparameter settings. The top graph refers to the route model, while the bottom refers to the favourite places model.



Best route model: trees = 10000, tree depth = 8, learning rate = 0.01 and minimum samples in leaf nodes = 5

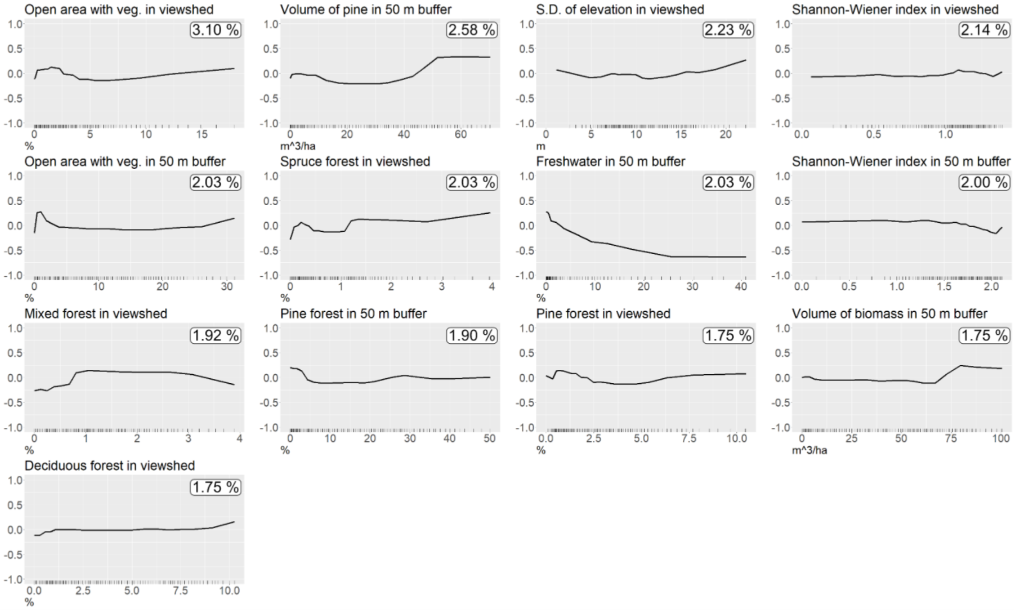


Best favourite places model: number of trees = 250, tree depth = 8, learning rate = 0.003 and minimum samples in leaf nodes = 5.

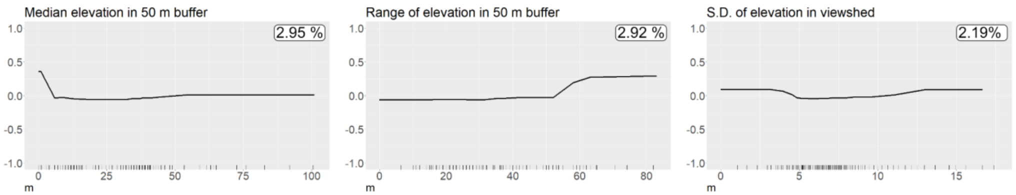
For code used to generate this and the subsequent analysis in the paper, see the data and code repository at https://zenodo.org/records/10910508

S5: Remaining influential effects of the two models

Route model



Favourite place model



The negative effect of freshwater within buffer in the route model is probably an artifact due to the way we created the availability sample: very few of the used routes were drawn directly in water, and although the start of the random routes were placed in terrestrial areas, the rest of the route was not guaranteed to be terrestrial. Most other effects are too weak to draw conclusions from, or pertain to land cover classes that are difficult to interpret (e.g. Open area with vegetation, which includes both urban lawns and semi-natural grasslands) or are very rare in the dataset (mixed forest).

IV

Received: 16 December 2022 | Accepted: 29 October 2023
DOI: 10.1002/pan3.10574

PEOPLE AND NATURE | BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

Favourite places for outdoor recreation: Weak correlations between perceived qualities and structural landscape characteristics in Swedish PPGIS study

Jonathan Stoltz[1] | Carl Lehto[2] | Marcus Hedblom[3]

[1]Department of People and Society, Swedish University of Agricultural Sciences (SLU), Lomma, Sweden
[2]Department of Ecology, Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden
[3]Department of Urban and Rural Development, Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden

Correspondence
Jonathan Stoltz
Email: jonathan.stoltz@slu.se

Funding information
Svenska Forskningsrådet Formas, Grant/Award Number: 2019-01916

Handling Editor: Robert Fish

Abstract

1. Connections between outdoor recreation and various health and well-being benefits are well established. However, questions remain regarding which landscape characteristics that best predict places in the landscape that correspond to people's needs and preferences. The perceived sensory dimensions (PSDs) model proposes eight basic perceived qualities that people commonly seek in outdoor environments to support complementary needs: a *Natural*, a *Cultural*, a *Cohesive*, a *Diverse*, a *Sheltered*, an *Open*, a *Serene* and a *Social* quality respectively.
2. These PSDs have increasingly been suggested as a tool for green space assessment and planning. How strongly they correlate with objective landscape characteristics is, however, still an open question. We surveyed recreationists in Sweden, tasking them with noting their favourite places on a map ($n=275$), and to report the degree to which the PSDs were present. The qualities typically reported as most prominent at these places were *Open*, *Serene* and *Sheltered*, while the least prominent were *Social* and *Cultural*.
3. A cluster analysis further revealed that favourite places could be classified into two main groups regarding perceived qualities. One associated with presumably more restorative qualities, emphasising *Natural* and *Serene* settings, the other instead suggesting a more outward-directed experience, strong in the perceived *Social* and *Cultural* dimensions.
4. Machine learning models, however, revealed weak links between objective landscape characteristics and perceived qualities, with stronger correlations found with attributes connected to personal characteristics, such as the degrees to which a person identifies as nature or urban oriented.
5. Although largely confirming the basic relations between the PSDs suggested by earlier studies, our results cast some doubt on the way they often have been understood and used, as describing an 'objective' truth of places, rather than representing qualities that largely emanate from the individual experience. Our

results instead confirm previous reports of weak general links between objective landscape measures and perceived qualities.

KEYWORDS
cultural ecosystem services, landscape preferences, outdoor recreation, perceived sensory dimensions

1 | INTRODUCTION

Urbanisation continues to increase (UN, 2019) while noncommunicable, lifestyle-dependent and often stress-related, illness dominate globally (WHO, 2021). Meanwhile, outdoor recreation and experiences of nature and greenery are recognised as important contributors to human health and well-being (e.g. Aerts et al., 2018; Bratman et al., 2019; Egorov et al., 2016; Hartig et al., 2014; McMahan & Estes, 2015). However, people's needs vary over time and between individuals, highlighting the potential need for diverse landscape features and biodiversity to accommodate different recreational styles (e.g. Marselle et al., 2021). This presents a challenge for landscape and urban planners, necessitating practical guidelines and models that can be used to predict how well the surrounding landscape supports general recreational needs. One such model is the perceived sensory dimensions framework (Adevi & Grahn, Grahn & Stigsdotter, 2010; Stoltz & Grahn, 2021) which attempts to define a set of basic perceived qualities, or perceived sensory dimensions (PSDs), that people commonly seek in recreational outdoor spaces. More than 60 studies employing this framework in various ways have been conducted in different parts of the world, including examples from the Nordic Countries (Lindholst et al., 2015), Estonia (Maikov, 2013), Serbia (Vujcic & Tomicevic-Dubljevic, 2017), Canada (Lockwood, 2017), Iran (Memari et al., 2017), Malaysia (Mansor et al., 2017) and China (Gao et al., 2019).

In a review and synthesis of several previous studies, Stoltz and Grahn (2021) proposes a model summarising the PSDs as eight basic perceived qualities: *Natural*, *Cultural*, *Cohesive*, *Diverse*, *Sheltered*, *Open*, *Serene* and *Social*, interrelated as in Figure 1a (ibid.). They suggest these qualities to support complementary recreational needs, relevant to both activity and rest. Stoltz (2022; fig. 1b) furthermore proposes an evolutionary model, linking the PSDs to different habitat conditions during the evolution and development of the human species, to explain how they support different stages of restoration and rehabilitation from high stress levels and cognitive fatigue. This model proposes a unified restorative pathway based on the PSDs, synthesising the two main theoretical approaches explaining nature-based restoration from an evolutionary perspective, the *attention restoration theory* (ART; Kaplan, 1995; Kaplan & Berman, 2010) and the *stress reduction theory* (SRT; Ulrich et al., 1991) respectively. It suggests *Serene*, *Sheltered*, *Natural* and *Cohesive* environments of primary importance to support early stages of such restoration, when stress levels are high and/or attentional capacities low, whereas *Diverse*, *Open*, *Cultural* and *Social* settings increase in importance at subsequent stages, as fundamental attentional capacities have been restored and stress levels lowered. This model is supported by empirical evidence presented by,

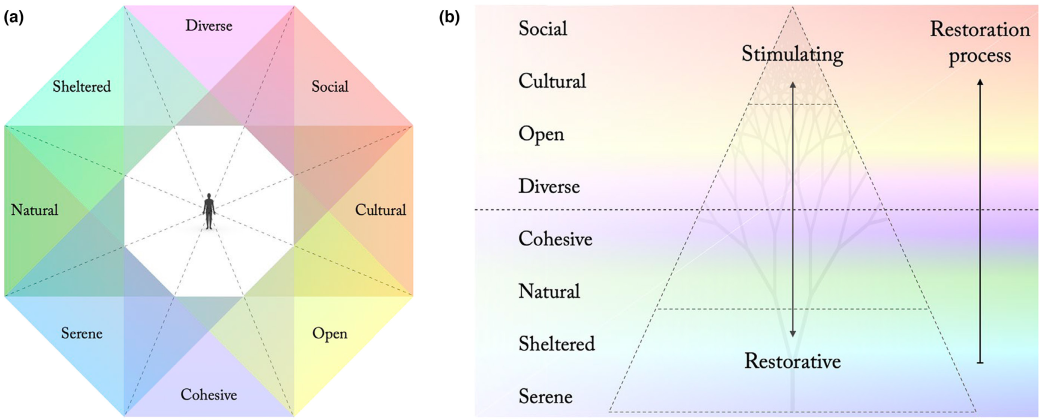


FIGURE 1 Schematic relations between the eight perceived sensory dimensions, where (a) correlations between perceived qualities are stronger the closer they are in the model (Stoltz & Grahn, 2021), and (b) in relation to supportive influence on restoration from high stress levels and cognitive fatigue, according to an evolutionary model (Stoltz, 2022).

among others, Grahn et al. (2010), Memari et al. (2017), Stigsdotter et al. (2017) and Pálsdóttir et al. (2018).

If certain landscape characteristics adequately could predict such perceived qualities of recreational areas, this could be useful for planning purposes when health and well-being outcomes and general recreational needs are considered. However, Stoltz and Grahn (2021) emphasise the necessity of an ecological approach to perception (Chemero, 2009; Gibson, 1979) when analysing the PSDs, viewing them as classes of *affordances* (ibid.); perceivable and utilisable behavioural possibilities offered in the environment, equally influenced by the physical world and the specific needs and abilities of the individual. To the extent people are similar, similar environmental attributes could be expected to reinforce each PSD. However, since humans also exhibit great individual variation, the PSDs cannot be seen as solely definable in terms of specific landscape features as individual characteristics most likely also will shape perceived affordances for the PSDs. The question is then to what extent the PSDs can be understood as universally shaped by certain landscape characteristics, and to what extent they are shaped by individual characteristics that vary in the population. Previous research by, for example, Leslie et al. (2010) has indicated a general lack of agreement between perceived qualities and objective landscape measures, which might be due to the influence of such individual characteristics in how environments are perceived. One example is provided by Gunnarsson et al. (2017) who reported that individuals considering themselves as mainly 'nature-oriented' rated areas with high objectively estimated biodiversity more in line with actual biodiversities than people who considered themselves as mainly 'urban-oriented'. Thus, attitudes and knowledge influence how people perceive the same objective landscape features.

Direct general connections between the PSDs and various structural landscape characteristics have been studied, if in a limited fashion, in urban (e.g. Skärbäck et al., 2014; Stoltz & Schaffer, 2018), rural (Adevi & Grahn, 2012; de Jong et al., 2012) and forest settings (Stigsdotter et al., 2017; Stoltz et al., 2016). In a Swedish survey study ($n = 121$) of urban green spaces, Qiu and Nielsen (2015) concluded that experiences of the PSDs were related to the diversity of biotopes offered by an urban green space and that larger green spaces containing more biotopes supported the experience of more PSDs. This appears in line with the suggestion by Stoltz (2022) that the PSDs can be connected to different habitat conditions during our evolution and development as a species. They also found experiences of the PSDs to be consistent across gender, age and frequency and type of recreational use, granting some legitimacy to the framework in assessment and mapping of recreational experiences (Qiu and Nielsen, 2015).

Björk et al. (2008) operationalised PSDs using a mix of objective landscape variables (land cover, noise, other map data), following the parameters used in a Swedish report by Skärbäck et al. (2009). Based on these models, they suggest that the presence of mapped PSDs within 300 m of residence correlates positively with well-being and propensity to exercise. The same GIS model was used by Annerstedt van den Bosch et al. (2015), who present a follow-up of the survey participants who have moved since the last study and compare the landscape conditions they moved from to what they moved to, as well as how they felt before and after. Based on their result, they suggest that moving to *Serene* environments might decrease the risk of mental illness. However, neither of these studies empirically validate their predictive PSD models, that is, show that the selected landscape variables cause people in general to perceive a particular PSD as stronger. de Jong et al. (2011, 2012) chose a different strategy and instead of structural landscape data constructed area-aggregated measures, derived from large public health surveys in which participants were asked about their perceptions of the PSDs in their close-by living environment. These results indicate that people tend to perceive their neighbourhood in a similar way as other people living within the same 1-km^2, which could suggest an underlying structural basis in the landscape for these perceptions. However, it also seems possible that such similarities, at least in part, could be attributed to some individual factors uniting people living in similar areas.

The main motivation for our study here was to investigate experiences of the PSDs at people's favourite places for outdoor recreation, and whether these could be predicted by a comprehensive set of landscape variables. As the PSDs have become more and more widely used both practically and in various research studies around the world, often with the assumption that they, more or less, directly reflect underlying objective landscape features, we wanted to test this assumption against a comprehensive set of landscape data together with a smaller set of individual characteristics. We also wanted to characterise people's favourite places in terms of general landscape types. To accomplish this, we gathered survey data on people's experiences at their favourite places during outdoor recreation. We employed a novel methodology, where we calculated what landscape was visible from the favourite places, using LiDAR data, to capture a closer approximation of the actual recreation experience. We utilised a large amount of map data as covariates, which was made possible by employing a flexible machine learning algorithm in the form of boosted regression trees (BRT) for modelling. Our main research questions were:

1. Which are the general landscape types at people's favourite places for outdoor recreation?
2. Which perceived qualities, PSDs, do people report at these places, and in which combinations?
3. Can the strength of these perceived qualities be accurately predicted by objective landscape characteristics at the site independent of individual characteristics, such as gender, age, educational background or nature/urban orientation?

2 | MATERIALS AND METHODS

To address our research questions, a digital survey was employed to residents of a large Swedish city. The collected data were then analysed by training a machine learning model on the characteristics of favourite places in the landscape.

2.1 | Survey

2.1.1 | Study area

The study area (Figure 2) consisted of Umeå Municipality, in Västerbotten County, Sweden. It covers an area of approximately 2300 km^2 and has an estimated population of 130,000, with a population density of 56.21/km^2 (2020). Its seat, the city of Umeå, is known for its university and many birch trees, giving it the nickname the 'Town of Birches' ('Björkarnas stad'). It is located near the coast of the Gulf of Bothnia, at the 63rd parallel. Climate is cold continental, with freezing winters and mild summers. Between the end of April and mid-August, the sun sets, but it does not get completely dark even at midnight. The Ume River that passes through the city widens into a fjord before flowing into the sea. The surrounding landscape is a mix of forests (mainly coniferous), arable land, some wetlands and lakes.

2.1.2 | Survey design

An invitation to participate in the survey was sent out to 3000 residents over the age of 18 in the city of Umeå via mail in September 2021. The list of recipients was acquired from the Swedish state person address registry, which provided a stratified sample designed to be proportional to the population of Umeå with regard to gender and age. A reminder to answer the survey was sent out 3 weeks later. The survey invitation contained a link to the digital survey, which was implemented using the GIS-based survey tool Maptionnaire (Mapita). In the survey, the respondents were asked to provide some basic demographic data (age, gender and level of education). They were also asked two separate questions to assess to what degree they would consider themselves as being 'nature-oriented' and 'urban-oriented' respectively. These terms were not further defined for the respondents and the questions were included as they have been shown to reflect factors with potential effect on greenspace use (Gunnarsson et al., 2017). Both questions had a slider with a range between 0 and 100, where 0 represented 'Not at all' and 100 'Fully'.

The main part of the survey was divided into two parts. The first part tasked the respondents with summarising their outdoor recreation within Umeå municipality by drawing typical routes they take, and providing details (e.g. type of activity, frequency and duration of visits etc.) of each route. These data were collected for a separate study and will not be further discussed here. The second part of the survey tasked the respondents with marking the location of their favourite places when performing recreation. A favourite place was defined as a place 'holding any specific importance, such as a place of beauty or somewhere you often stop and spend time in'. For each place, they were also asked to assess eight statements, each corresponding to one PSD (Table 1). These were based on the definitions of the PSDs described by Stoltz and Grahn (2021) and were phrased as simple one sentence statements, intended to capture the essence of each PSD. As such, they were very similar to the statements used by, for example, Björk et al. (2008), de Jong et al. (2011, 2012) and Stoltz et al. (2016).

For each statement, the respondents were presented with a slider that ranged from 0 to 100, where 0 corresponded to 'Not at all' and 100 to 'Fully'. The slider's starting position was in the middle (Stoltz et al., 2016). The participants were also asked to mark their home location on the map. Prior to deployment, the survey was tested on a convenience sample of 45 friends and colleagues, after which minor changes in wording of questions were made.

Since the survey did not handle sensitive information, we assessed it as not falling under any of the criteria listed in the Swedish Ethical review act (2003:460), and thus did not need authorisation from the Swedish Ethical Review Authority. The collected data were handled in accordance with GDPR, and the respondents were provided information on how the collected data would be handled at the start and the end of the survey, where consent was asked.

2.1.3 | Summary of responses

Of the 3000 invited participants, 658 opened the link to the digital survey, and 285 finished the entire survey. One hundred and ninety-five individuals placed one or more favourite places, yielding a total sample of 318. Favourite places outside Umeå municipality were removed ($n=26$). Favourite places where the respondent had not interacted with any of the eight PSD sliders were also removed ($n=17$). For respondents that had interacted with at least one of the sliders, untouched sliders were interpreted as having been left in the middle deliberately and counted as 50. Final sample consisted of 275 favourite places, originating from 181 individuals. The gender distribution of the respondents was 47% male and 53% female. Median age was 45, with a standard deviation of 17, which is similar to the Umeå average (49 ± 18, Umeå kommuns demografidatabas 2023). The respondents were more educated than the Umeå average, with 69% having attended higher education in some capacity, compared to the Umeå average of 38% (Statistics Sweden, 2021).

2.2 | Correlations between PSDs and cluster analyses

To see how reported PSDs at favourite places were correlated to each other, a correlation matrix was produced. To investigate if favourite places could be sorted into different clusters regarding PSDs, *K*-means clustering was performed. *K*-means clustering is a commonly used unsupervised machine learning algorithm that partitions a data set into a given number (*K*) of different clusters, where each observation belongs to the cluster with the nearest mean (Hartigan, 1975). The algorithm iteratively updates the cluster centroids and assigns

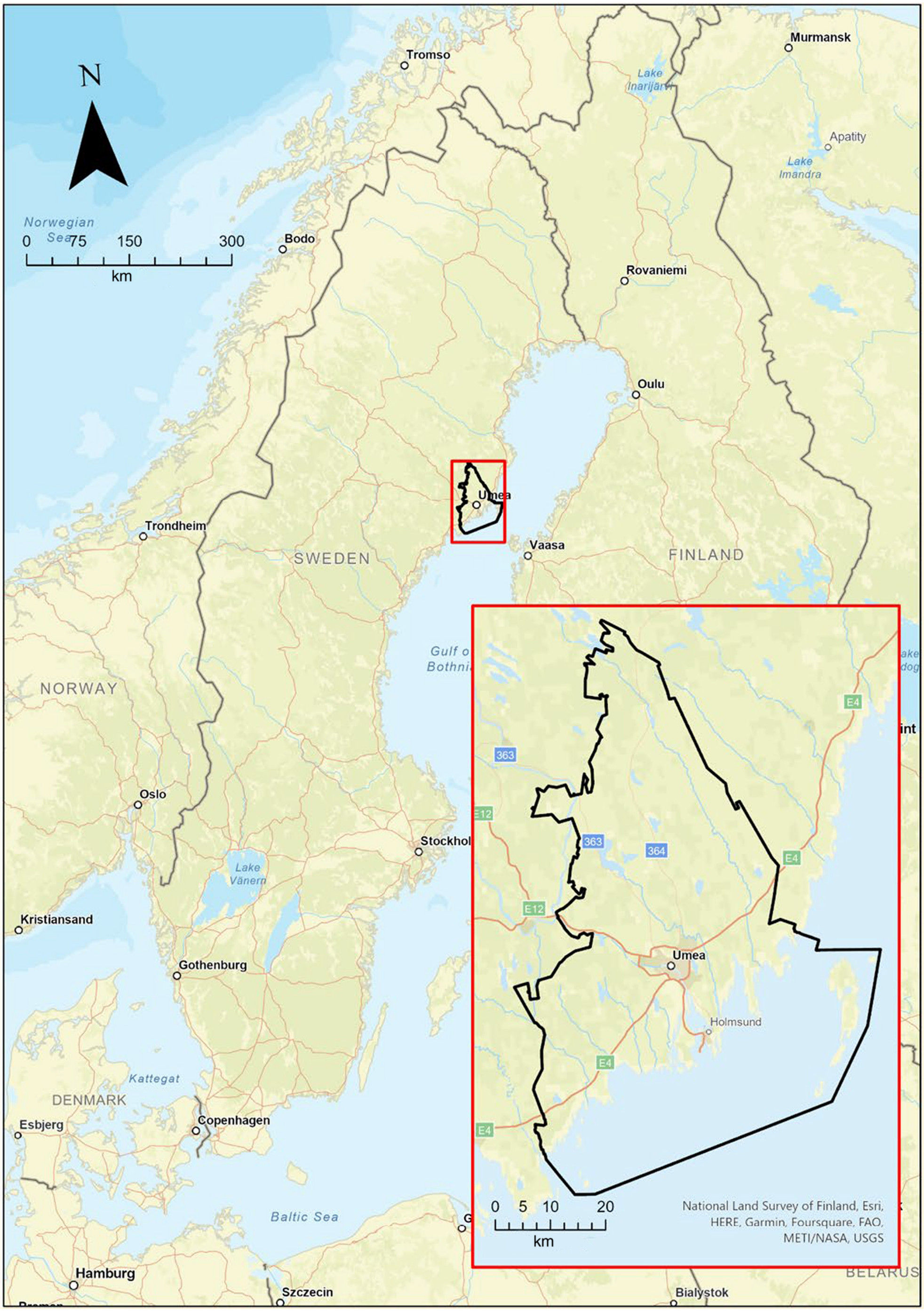


FIGURE 2 The study area of Umeå municipality, located in northern Sweden.

each observation to the nearest centroid until convergence. The resulting clusters can be used for exploratory data analysis, pattern recognition and other data mining applications. A prerequisite to employ the method is to choose the number of clusters (*K*), and as we did not have any prior hypotheses on the number of groups, we employed two methods to estimate how many clusters existed in the data set: the Caliński-Harabasz index (Caliński & Harabasz, 1974) and the silhouette method (Rousseeuw, 1987), which both are algorithms that estimate how well a given data set clusters.

2.3 | Predictive PSD modelling using landscape characteristics

To evaluate whether PSDs could be predicted by landscape or individual characteristics, eight machine learning models were created, one for each PSD.

TABLE 1 Statements for each PSD in the survey, answered using sliders from 0 ('Not at all') to 100 ('Fully').

PSD	The place evokes a sense of...
Natural	... wild and untouched nature
Cultural	... being shaped by humans
Open	... openness, with opportunities for vistas
Social	... a social space, with opportunities to interact with other people
Cohesive	... a cohesive whole, of being a world in itself
Diversity	... diversity and variation
Sheltered	... shelter
Serene	... serenity

Abbreviation: PSD, perceived sensory dimension.

2.3.1 | Converting points to experienced landscape

To define the extent of each place a combination of two approaches was employed. First, a circular buffer with a radius of 50 m was created around each point which represented the immediate surroundings the respondent experienced. Second, using a high-resolution digital surface model (DSM), a viewshed was calculated that represented the area that was visible from a height of 1.5 m when standing at the point using ArcGIS Pro 3.0.0. The viewshed was calculated with a maximum sight distance of 1 km for computational reasons. Trees and vegetation were assumed to be total sight blockers, except for within the 50 m buffer. Figure 3 shows two examples of the sampled landscape around two favourite places in our study.

2.3.2 | Model predictors

Several different landscape characteristics were sampled using different map sources (Table 2). Some landscape predictors were sampled in both the viewshed and the buffer, while others were exclusive to the buffer. Land cover was extracted from the CadasterENV Sweden map (Swedish Environmental Protection Agency, 2018) and reclassified from 25 original classes into 13 for easier model interpretation (Supplementary Materials S1). Each land cover type's cover in % of the buffer and the viewshed was used as a predictor, but they were also used to estimate landscape heterogeneity. This was done by calculating the Shannon–Wiener diversity index (Shannon, 1948). The SLU forest map (SLU, 2015) added nuance to the land cover maps in forested areas by supplying estimates of tree height and volumes of different tree species, as well as total biomass volume.

Biodiversity was included in the model by combining several sources of map data: the extent of all formally protected areas

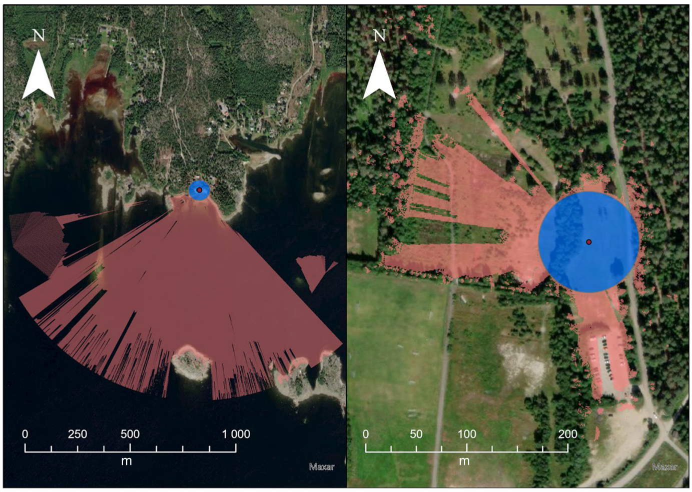


FIGURE 3 Examples of the sampled landscape around two favourite places. Centre point is the favourite place provided by the survey respondent. The blue circle is the 50 m buffer around this point, and red areas the calculated visible landscape, viewshed (360 degrees), when standing at the point.

(national parks, nature reserves, protected biotopes) was merged with maps of woodland key habitats (forests with high biodiversity values, see e.g. Timonen et al., 2010). Maps provided by the municipal government on areas with high biodiversity values were also included, with the predictor used in the model being the percentage overlap between the buffer and any of these maps. The municipal government provided three maps of estimated average noise levels due to road traffic, railroad traffic and industry respectively. These were combined by taking the highest estimated noise level at each point of the three maps, and then calculating the average across the buffer. Paths and roads were extracted from OpenStreetMap (Openstreetmap Foundation, n.d.), and lengths of each were calculated within the buffer. Data on recreational infrastructure (shelters, toilets and fireplaces) was also supplied by the municipal government and used as a predictor by calculating the distance from the point to the nearest recreational infrastructure. Topography was considered by calculating the median, standard deviation and the range (largest difference) of elevation above sea level within the buffer and the viewshed. Table 2 shows a summary of the landscape predictors used in the machine learning models. In addition to these landscape predictors, demographic data were included as predictors (Table 3). Age was excluded from the demographic variables due to 63 missing responses.

2.3.3 | Boosted regression trees

Modelling was performed using BRT. BRT is a machine learning approach where a predictive model is created by iteratively building an ensemble of many decision trees (Friedman, 2001). The method has several advantages: It does not assume linear relationships between predictor variables and response variables; it can handle a large number of predictors regardless of multicollinearity; and it avoids

TABLE 2 Summary of landscape predictors used in the machine learning models.

Predictor	Description
Land cover (13 predictors)[a,b]	Composition of land cover types. Data source CadasterENV
Shannon-Wiener diversity[a,b]	Landscape heterogeneity, calculated using the land cover classes
Tree height[a,b]	Average height of trees (m). Data source SLU Forest map
Spruce volume[a,b]	Average volume of Norway spruce (m^3/ha). Data source SLU Forest map
Pine volume[a,b]	Average volume of Scots pine (m^3/ha). Data source SLU Forest map
Birch volume[a,b]	Average volume of birch (m^3/ha). Data source SLU Forest map
Deciduous tree volume[a,b]	Volume of deciduous trees except birch (m^3/ha). Data source SLU Forest map
Biomass volume[a,b]	Total volume of all vegetation (m^3/ha). Data source SLU Forest map
Elevation (3 predictors)[a,b]	Median, standard deviation and range of elevation. Data source LIDAR DSM Lantmäteriet
Noise[a]	Average noise(db) over 24 h. Data source Umeå municipality noise estimates based on models of road and railroad traffic and industry
High biodiversity area[a]	Overlap of buffer with areas of high biodiversity (%). Areas included national parks, nature reserves, woodland key habitats and areas of high conservation value in the municipal inventory
Path/road length[a]	Length of paths/roads within buffer (m). Data source OpenStreetMap
Amenity distance	Distance (m) from point to the closest recreational amenity (shelter, toilet, fireplace). Data source Umeå municipality

[a]Predictor was sampled within the 50 m buffer.
[b]Predictor was sampled within the viewshed.

TABLE 3 Individual characteristics used as predictors in the machine learning models.

Variable	Description	Values
Gender	The gender of the respondent	Man, woman, other
Education	Highest level of finished education	Elementary school Secondary school Folk high school Folk high school University > 2 years
Urban oriented	To what extent the person identifies as urban-oriented in terms of general environmental preference	Discrete [0,100]
Nature oriented	To what extent the person identifies as - nature-oriented in terms of general environmental preference	Discrete [0,100]

the need for model selection or pre-specifying interaction effects in advance. The main disadvantage of BRT is the lower interpretability of the final models, having more aspects of being a 'black box' than traditional regression models such as GAMs or GLMs. However, with recent methodological advances, such as the Interpretable Machine Learning package for R (Molnar, 2018), these shortcomings can be mitigated to a larger extent.

All analyses and visualisations were carried out using the *gbm* package (Greenwell et al., 2020) in R version 4.0.3 (R Core Team, 2020). Boosted regression trees were constructed following the recommendations outlined by Elith et al. (2008), using a Gaussian distribution with each PSD as the response variable. When fitting BRT, three hyperparameters that affect model fitting are set: (1) tree complexity (how many splits are allowed in each tree); (2) learning rate (how quickly the algorithm converges, with lower values leading to better models at the cost of computing time); and (3) bag fraction (how large a fraction of the data set to use in each iteration). We created models with combinations of five different tree complexities (Adevi & Grahn, 2012; Aerts et al., 2018; Annerstedt van den Bosch et al., 2015; Beery et al., 2015; Bratman et al., 2019) and two bag fractions (0.5 and 0.75) and lowered the learning rate until a model of at least 1000 trees was fitted (ibid.). Model performance was evaluated using cross-validated R^2-values. Feature importance, interaction effects and partial dependence plots were evaluated using the *iml* package (Molnar, 2018).

3 | RESULTS

3.1 | General landscape types at favourite places

Our first research question related to which general landscape types that are found at people's favourite places. Figure 4 shows the coverage (%) for general land cover classes within the 50 m circular buffer at people's favourite places. The four forest classes (spruce, deciduous, pine and mixed) were grouped together. Forest and water dominated at favourite places, with forest being the most common land cover type (Figure 4). This seems quite in line with the supply in the study area as whole, which consists of a mix of forests (mainly coniferous), arable land, some wetlands and lakes. Figure 5 shows the locations of people's favourite places within the study area.

3.2 | Distribution and combinations of PSDs at favourite places

Our second research question related to the distribution and combinations of PSDs at people's favourite places.

3.2.1 | Distribution of PSDs

Figure 6 shows the distribution of ratings (0–100) for the presence of each PSD at the favourite places. The *Natural*, *Cultural*, and *Social* PSDs showed a larger variation in response than the other PSDs, which mainly elicited responses at or above 50. *Serene*, *Open* and *Sheltered* were the most pronounced qualities at favourite places (mean 74, 71 and 69 respectively), while *Cultural* and *Social* were the weakest (mean 50 and 46 respectively).

3.2.2 | Correlations between PSDs at favourite places

Table 4 shows correlations between PSDs at favourite places in our study. Green highlights positive and red negative correlations between perceived qualities. More saturated colour indicates stronger correlation. Correlations weaker than ±0.1 are not highlighted with any colour.

3.2.3 | Cluster analysis of PSDs at favourite places

The Calinski–Harabasz index and the silhouette analysis both suggested that the data set contained two clusters. Using k-means

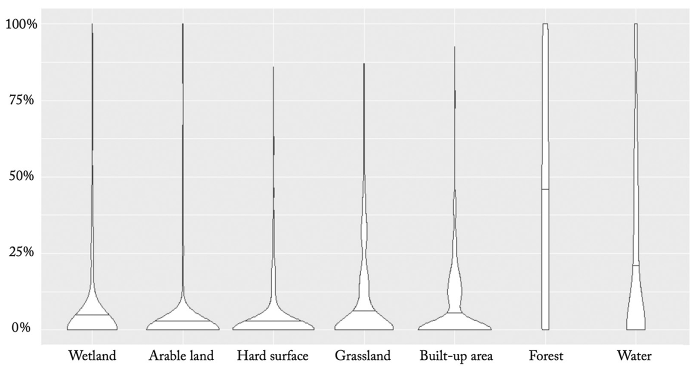


FIGURE 4 Violin diagrams showing the distributions of land cover percentages (0–100) of the 50 m circular buffers around favourite places in the study area. The width of each violin is a smoothed density plot, corresponding to the amount of data for each value along the *y*-axis. The median value is represented by the horizontal line within each diagram.

clustering with two clusters yielded groupings whose main differences were between the *Natural–Cultural* and *Serene–Social* axes of the PSD model: Group 1 is characterised by a stronger influence of the *Social* and *Cultural* PSDs, while Group 2 emphasises *Natural* and *Serene*. The remaining qualities (*Sheltered-Open*, *Diverse-Cohesive*) all clustered weakly with *Social* and *Cultural* in Group 1 (Figure 7).

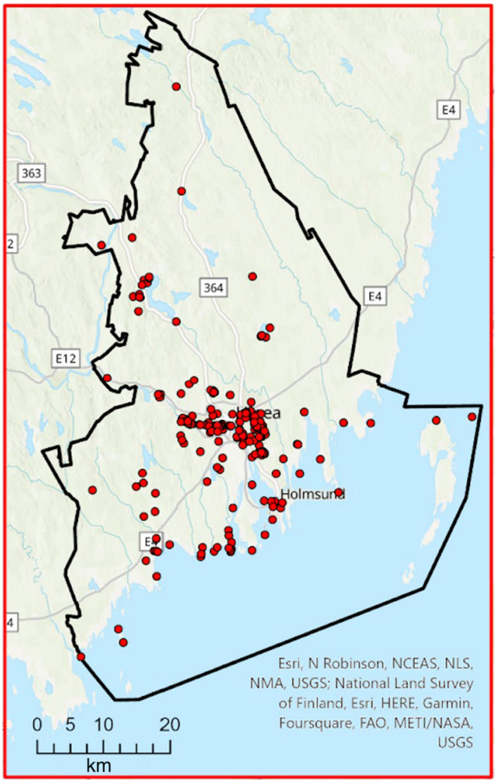


FIGURE 5 Locations of people's favourite places in the study, all within the borders of Umeå municipality, Sweden (see Section 2.1.1).

3.3 | Predicting PSDs using landscape characteristics

Our third research question regarded whether the reported strength of the PSDs could be accurately predicted by objective landscape characteristics at the site independent of individual characteristics. Overall, the eight models predicting PSDs achieved low predictive power. The strongest model was for the *Natural* PSD ($R^2 = 0.27$), followed by *Social* ($R^2 = 0.19$) and *Cultural* ($R^2 = 0.14$). The *Open*, *Cohesive*, *Diverse*, *Sheltered* and *Serene* models had little explanatory power ($R^2 < 0.1$) and were deemed too weak to draw any meaningful conclusions from. Which predictors had the largest effect on the outcome of the models were evaluated by calculating the relative influence of each predictor, a measure of how important each predictor is for model performance. When interpreting BRT models, a rule of thumb is that predictors with a relative influence higher than the inverse of the number of predictors (in our models 1/53 ≈ 1.9%) are worth looking at. However, with weak models and many predictors as in our study, this rule is less applicable.

To investigate the specific effects of each predictor, partial dependence plots are created. These evaluate how model outcomes change when the predictor of interest varies, while keeping all other predictors at their median value. Our models had one to two predictors that were responsible for most of each respective model's performance, followed by many predictors with low relative influence. Figures 8–10 show partial dependence plots of the six most influential predictors for our three models with explanatory power, R^2, greater than 0.1. Above the *x*-axis of each predictor is a rug plot, showing the distribution of values within the data set, with each notch representing 1% of the data set. The graphs show the entire range of values for each predictor within the data set, but as the machine learning algorithm fits few trees where there are little data. Interpretation should thus be focused on sections with higher data densities, approximately highlighted with rectangles in

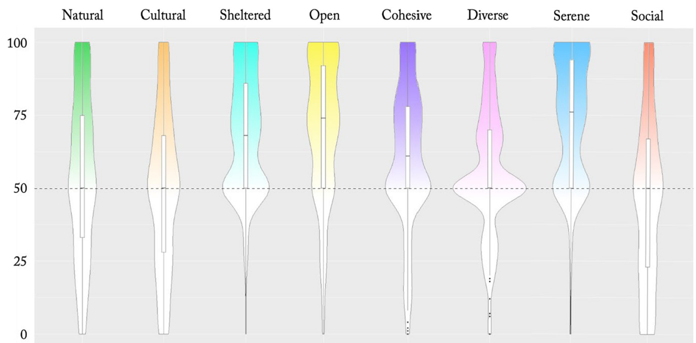


FIGURE 6 Violin diagrams showing distribution of ratings (0–100) for the strength of each perceived sensory dimension at people's favourite places. The width of each violin is a smoothed density plot, corresponding to the amount of data for each value along the *y*-axis. Within each violin is a box plot showing quantiles, with the median value as a line. Outliers are marked by dots.

	Cohesive	Serene	Natural	Sheltered	Diverse	Social	Cultural
Open	0.17	0.10	0.16	0.10	0.03	0.24	0.14
Cultural	0.00	−0.17	−0.28	0.09	0.17	0.43	
Social	0.02	−0.11	−0.07	0.14	0.29		
Diverse	0.42	0.15	0.27	0.44			
Sheltered	0.42	0.48	0.20				
Natural	0.48	0.35					
Serene	0.38						

Note: More saturated colour indicates stronger correlation.
Abbreviation: PSD, perceived sensory dimension.

TABLE 4 Correlation coefficients between PSDs at favourite places ($n = 275$).

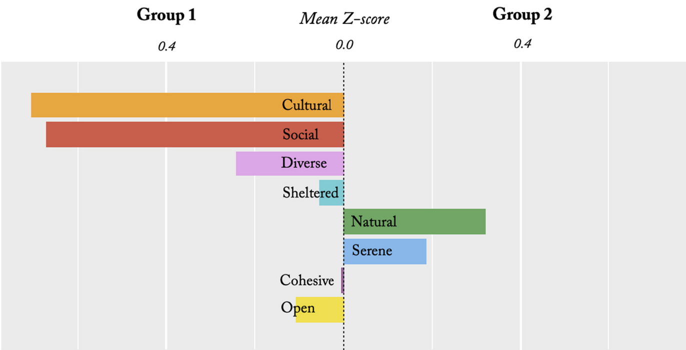


FIGURE 7 Differences between perceived sensory dimensions (PSDs) in the two groups of favourite places suggested by the performed cluster analyses. The higher the score, the greater the difference for this PSD between the two groups.

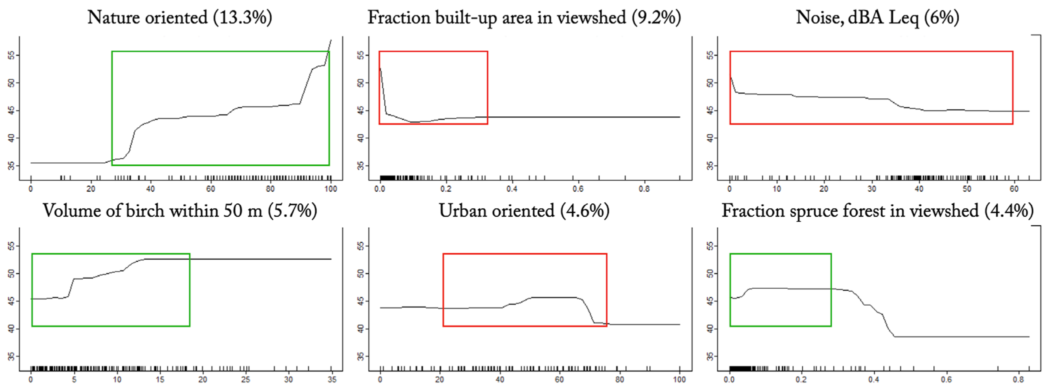


FIGURE 8 The natural perceived sensory dimension (PSD) model. Partial dependence plots for the six most influential predictors show how the PSD value (*Y*-axis) was predicted to change with each predictor. Relative influence of each predictor within parentheses (%).

green (positive), red (negative) or blue (U-shaped) colour, depending on the observed trend for the variable in relation to PSD strength. The full list of influential predictors for each model can be found in Supplementary Material S2.

In the model for the *Natural* PSD (Figure 8), higher ratings were positively correlated with identifying as a nature-oriented person, the volume of birch within 50 m and the fraction of spruce forest in the viewshed. The fraction of built-up areas in the viewshed, the amount of noise and identifying as an urban-oriented person were all negatively correlated. In the *Cultural* PSD model (Figure 9), the fraction of built-up area, the Shannon diversity index and identifying as an urban-oriented person were all positively correlated with perceiving the quality at favourite places. Distance to recreational infrastructure and identifying as a nature-oriented person were

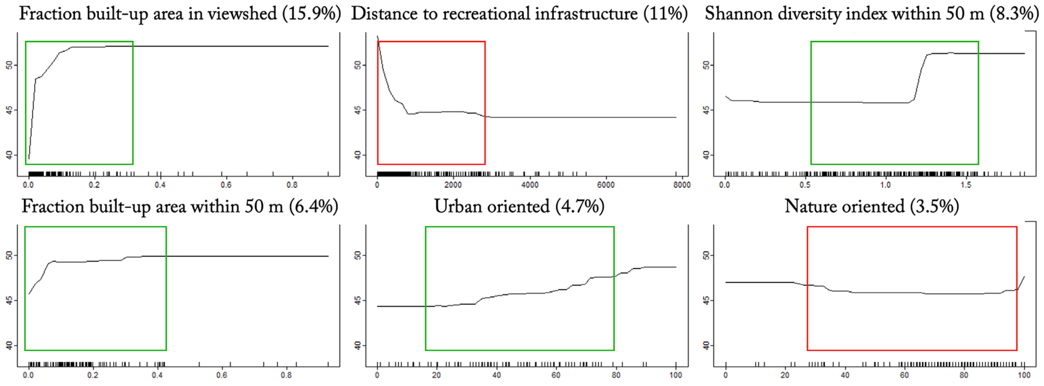


FIGURE 9 The cultural perceived sensory dimension (PSD) model. Partial dependence plots for the six most influential predictors show how the PSD value (*Y*-axis) was predicted to change with each predictor. Relative influence of each predictor within parentheses (%).

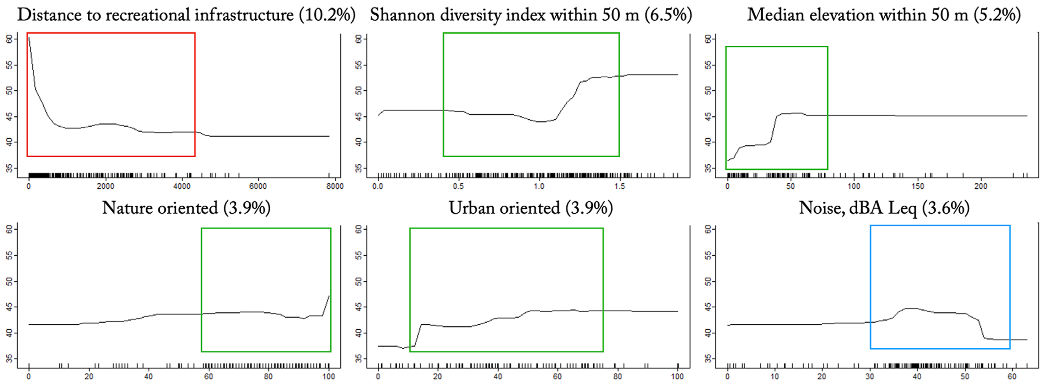


FIGURE 10 The social perceived sensory dimension (PSD) model. Partial dependence plots for the six most influential predictors show how the PSD value (*Y*-axis) was predicted to change with each predictor. Relative influence of each predictor within parentheses (%).

negatively correlated. Finally, in the *Social* PSD model (Figure 10), the Shannon diversity index, the median elevation and the degree of identifying as urban-oriented were positively correlated with perceiving the quality, while increased distance to recreational infrastructure was negatively correlated. Noise interestingly seems to indicate a *U*-shaped relation to this quality, suggesting perhaps that a certain amount of noise is a tolerable or maybe even a positive attribute of *Social* environments, whereas there are limits above which the quality diminishes.

4 | DISCUSSION

The main goals with our study were to (1) determine which general landscape types describe people's favourite places for outdoor recreation; (2) which PSDs that people experience at these places, and in which combinations; and (3) to determine the degree to which biophysical landscape characteristics could predict these PSDs and whether individual characteristics could be an important factor for such models to consider as well. We approached these questions by training machine learning models on a large set of landscape data surrounding favourite places against survey data with locations of favourite places for outdoor recreation and basic individual characteristics, including gender, age, educational background and the degrees to which a person identifies as nature and urban oriented respectively.

4.1 | General landscape types at favourite places

Regarding our first research question, forest and water were the dominating land cover types at people's favourite places in our study, with forest by far being the most common (Figure 4). It thus does not seem like favourite places in our study differ dramatically from what is provided by the landscape in large, at least not at this rather coarse level of analysis. Future studies might want to explore direct links between landscape features and favourite places for recreation in more detail, using higher resolution landscape data for the predictive models.

4.2 | Distribution and combinations of PSDs at favourite places

Our second research question was related to which perceived qualities, PSDs, that people experience at their favourite places and in which combinations.

4.2.1 | Distribution of PSDs

The most pronounced PSDs at favourite places in our study ($n=275$) were *Serene*, *Open* and *Sheltered*. *Cohesive* and *Diverse* showed a similar trend, with values mainly above 50. *Natural*, *Cultural* and *Social* showed a larger variation in response than the other PSDs with both high and low values being represented at favourite places. That *Open* stands out as the most pronounced perceived quality at favourite places in our material could indicate that this quality is (a) overall common in the available landscape or (b) particularly important to people and thus actively sought out, or both. Since we do not have comparable direct assessments for PSDs at places that are not considered favourite places for recreation, we are not able to assess the degrees to which (a) or (b) might be the case here. The *Open* PSD is associated with long, unbroken sightlines and plenty of space to roam freely without physical obstacles (Grahn & Stigsdotter, 2010; Grahn & Stigsdotter, 2003; Stoltz & Grahn, 2021). Other studies have revealed similar importance to 'view', 'openness' and 'open landscapes with a view' (Hedblom et al., 2019; Knez & Eliasson, 2017; Schirpke et al., 2013), although all in mountain landscapes. However, Pouwels et al. (2020) revealed that distance to roads and openness were the two most important factors predicting visitor densities in parks.

Serene was the second most pronounced PSD at favourite places in our study, followed by *Sheltered*. Both these qualities have been strongly associated with restoration of high stress levels and cognitive fatigue (see e.g. Grahn et al., 2010; Stigsdotter et al., 2017; Pálsdóttir et al., 2018; Figure 1b). Hence, this could indicate a bias in our survey sample towards seeking restorative support in the recreational landscape. There was a tendency for *Natural* to be a more pronounced perceived quality at favourite places than *Cultural*. However, neither dimension was generally perceived as particularly articulated (mean values around 50), suggesting that isolation of either quality along this axis might be less important than perhaps expected (a *Natural* quality is generally considered as the more restorative; ibid.).

Similarly, the *Diverse* PSD, associated with perceived biodiversity and structural variations, also appears as a less important factor at favourite places in the study than initially hypothesised (following, e.g. Marselle et al., 2021). Again however, this might reflect a low support for such a quality in the environment rather than a low general demand, something this study is not able to determine. The opposite, *Cohesive* PSD, appears as generally slightly stronger at favourite places than *Diverse*. This might again indicate a bias in our sample towards selecting restorative settings for recreation, since the *Cohesive* PSD generally is considered the more restorative of the two, although the importance of *Diversity* for restoration seems to increase as stress levels and mental fatigue diminish (see e.g. Grahn et al., 2010; Memari et al., 2017; Figure 1b).

Overall, our results here can be compared to another survey study from the south of Sweden, reporting *Open* ('prospect'), *Serene* and *Cohesive* ('space') as the most commonly perceived PSDs, and *Cultural* and *Social* the least (Qiu & Nielsen, 2015). A result much in line with our findings here, although the latter study did not focus specifically on favourite places but rather on perceived availability of the PSDs in a limited number of preselected urban green spaces.

4.2.2 | Correlations between PSDs at favourite places

The oblique rotation factor analysis that is the basis for the PSD model allows for some correlation between qualities (Stoltz & Grahn, 2021; Figure 1a). At the same time, it is important for the relevance of each factor that they are not too closely related but indeed point towards and assess distinct aspects of the perceived environment. That no correlation coefficient exceeds 0.5 in our study here (Table 4) indicates that this is indeed the case; the PSD model seems to assess eight distinct dimensions of the perceived environment, however with some PSDs being more strongly related than others. The model suggests that correlations between qualities are stronger the closer to each other they appear, with the perceived tension between qualities being at its maximum at the opposite quality. Largely, this is confirmed here by the observed correlations between PSDs reported at favourite places (Table 4); correlations tend to decrease when moving away from a quality in the PSD model and to be the lowest around three to five qualities away, as suggested by the model (ibid; Figure 1).

Both the *Social–Serene* and the *Cultural–Natural* axes follow this pattern. *Serene* and *Social* were negatively correlated ($R^2=-0.11$), as predicted by the model. The same was true for the *Natural* and the *Cultural* qualities ($R^2=-0.28$). *Cultural* and *Serene* also appear far apart in the PSD model and are thus predicted to not be strongly correlated, which was also confirmed by our results here ($R^2=-0.17$). The same was true for the *Social* and *Natural* PSDs ($R^2=-0.28$). The relatively strong correlation between *Cultural* and *Social* ($R^2=0.43$) is also in line with what the model would suggest, where these qualities are adjacent. In addition, *Serene* is strongly associated with *Sheltered* ($R^2=0.48$; Table 4). These two qualities are commonly mentioned as the most restorative in empirical studies (e.g. Grahn et al., 2010; Stigsdotter et al., 2017; Pálsdóttir et al., 2018; Figure 1b). The fact that they often occur together at favourite places in our study could thus suggest a preference for restorative sites in our survey sample. This is further supported by an overall negative association with the *Social* PSD and favourite places in our material, a quality usually considered as the least restorative of the PSDs (ibid.; Figure 1b).

There are, however, also some exceptions to this general pattern that are interesting to highlight. One of these is the weak correlation in our study between perceptions of a *Natural* and a *Sheltered* quality (R^2=0.2). They are suggested as closely related by the PSD model and might thus be expected to often occur more together. In part, this might be due to a slight mistranslation of the English word '*sheltered*' into Swedish '*trygg*', that associates with a more general sense of safety rather than the more immediate physical protection emphasised by the English word. The PSD *Sheltered* is associated with both these aspects, however, usually emphasises possibilities for physical protection and possibilities to 'see without being seen' (Stoltz & Grahn, 2021). Furthermore, PSDs *Cohesive* and *Diverse*, which are suggested as opposing qualities in the PSD model, are quite strongly associated here (R^2=0.42). According to an evolutionary model (Stoltz, 2022; Figure 1b), these two PSDs can be seen as evolutionary closely related, which might explain why the distinction between them often is perceived as less sharp compared to that between, for example, a *Natural* and *Cultural*, or a *Social* and *Serene* quality, which appear further away from each other evolutionary, according to this model. Perhaps is this a reason why people often seek environments where both of these qualities can be perceived simultaneously.

The *Open* PSD seems relatively unaffected by the other qualities in our study, although a weak trend can be seen supporting the general rule of thumb of diminished correlation for qualities more distant in the PSD model. However, the specific correlation between PSDs in this study of course also depends on the overall distribution and supply of the different PSDs in the landscape, which was not controlled for. There is a relatively weak positive correlation between a *Sheltered* and an *Open* quality (R^2=0.1), even though these qualities appear as opposites in the PSD model (Stoltz & Grahn, 2021; Figure 1a) suggesting that attributes in the environment supporting a sense of *Shelter* in general decrease perceptions of *Openness*, and vice versa. Our results, however, might indicate that people actively seek out places where these two qualities can be perceived in close proximity, in a similar way as with the *Diverse* and *Cohesive* qualities discussed above. This could be taken as support for the *prospect-refuge theory* suggested by Appleton (1975), the idea that humans share an affinity for settings providing physical protection combined with a broad overview of the landscape, due to evolutionary causes. Finally, the PSD model (Stoltz & Grahn, 2021; Figure 1a) suggests that the *Open* quality is closely related to the *Cohesive* PSD, a sense of spatial and structural unity. In our study here, these two qualities appear as moderately associated (R^2=0.2).

4.2.3 | Cluster analysis of PSDs at favourite places

To further investigate the existence of typical landscape types, defined as combinations of certain PSDs, a cluster analysis was performed. This suggested two clusters in our material. Group 1 is defined by the relative strength of *Cultural*, *Social*, *Diverse*, *Open*, as well as to some degree *Sheltered*. Group 2 is distinguished by a relative strength of *Natural* and *Serene* compared to Group 1. In many ways, these results seem to be in line with what is suggested by the PSD model, where *Social* and *Serene* are suggested as opposites, and also appear in opposite clusters here, the same for *Natural* and *Cultural*. It is thus clear that the separation between the two groups occurs around the *Natural–Cultural* and the *Social–Serene* axes of the PSD model. *Cohesive* and *Sheltered* both grouped with *Social* and *Cultural* (Group 1), rather than with *Natural* and *Serene* (Group 2), in our sample. Considering the suggested relative restorativeness of the PSDs (Stoltz, 2022; Figure 1b), it nevertheless appears as if Group 1 expresses a more outward-directed or activity-oriented recreational experience, whereas Group 2 seems to emphasise a more rest-oriented recreational style, highlighting qualities from the bottom of this gradient.

As mentioned, neighbours in the PSD model are suggested to share associations and supporting attributes, and thus often correlate in the perceived landscape, while opposing qualities in the model might weaken each other and more rarely be strong together (Stoltz & Grahn, 2021; Figure 1a). To a large extent, this seems to be reflected also in how people perceived the PSDs in our study here, as there is a clear gradient for the cluster associations (Figure 7) when moving stepwise in the PSD model (Figure 1a). The *Cohesive* PSD shows barely any difference between the two groups and thus seems to be of equal importance at both main types of favourite places. It thus poses as a potentially more universally relevant PSD, independent of whether the place is perceived as more *Natural* or *Cultural*, *Serene* or *Social*. The *Cohesive* PSD is a quality associated with the capacity to provide the visitor a sense of a united, cohesive whole, a 'world in itself', possible to enter and explore without immediately perceiving its boundaries (ibid.). It thus directly depends on a certain size of the area, that will need to be large enough to support such an experience. However, the overall size of the area indicated as a favourite place was not something this study took into consideration, as each such place was indicated as a point in the map within a 50-m circular buffer. The opposing quality in the PSD model, *Diverse*, is often perceived as more stimulating while the *Cohesive* PSD is emphasised as important for earlier stages of restoration (see e.g. Grahn et al., 2010; Stigsdotter et al., 2017; Memari et al., 2017; Pálsdóttir et al., 2018; Figure 1b). This is also reflected in our results here, where *Diverse* clusters more strongly with the presumably more stimulating, and less restorative, *Social* and *Cultural* qualities (ibid.; Figure 1b).

4.3 | Predicting PSDs with structural landscape characteristics

Our third research question was whether the PSDs can be accurately predicted by objective landscape characteristics independent of individual characteristics. Our PSD machine learning models generally had low explanatory power, showing that the included landscape variables and individual characteristics were largely insufficient to efficiently predict the PSDs. The three strongest models, for PSDs *Natural*, *Cultural* and *Social*, although still having low explanatory power, showed some interesting patterns. The *Natural* PSD had several expected effects, such as being negatively affected by built-up areas and noise. The strongest effect, however, was the degree to which the survey respondent identified as a nature-oriented person, with a strong positive correlation. In previous studies (Gunnarsson et al., 2017), highly nature-oriented persons were shown to perceive more urban greenery-related aesthetics, more greenery-related sounds and greater importance of trees and plants for their perception of bird species in urban greenery compared to less nature-oriented persons. Thus, there seems to potentially be a stronger link between the way people define themselves and the perception of the environment compared to landscape characteristics such as composition of land cover, or type of forest.

The *Cultural* and *Social* models had many commonalities, as expected by their adjacency in the PSD model, and high degree of covariance in the data set. Both were positively correlated with the degree to which the respondent identified as urban-oriented, an increased fraction of built-up area and increased landscape heterogeneity (as measured by Shannon's index). This while increased distance to recreational infrastructure was negatively associated with both qualities. These effects are not surprising, being connected to urban areas or developed recreational areas, which expectedly would score higher on both the *Cultural* and *Social* PSD. The positive correlation with Shannon's index is probably due to the higher heterogeneity of land cover classes within urban and peri-urban areas compared to more natural environments.

It is still possible that each PSD depends reliably on some objective landscape features, only that these were not included in our analysis here. Qiu and Nielsen (2015) suggested that differences regarding factors such as size, location, vegetation structure and management level of green spaces are likely to be the most decisive factors for people's perceptions of the PSDs. They concluded that more diversity of biotopes leads to a greater number of strongly experienced PSDs. This is in line with findings by Plieninger et al. (2013), who concluded that the assignment of perceived landscape values is closely related to biophysical landscape features and spatial properties. Similarly, Björk et al. (2008) and Annerstedt van den Bosch et al. (2015) suggested that the PSDs might be reliably described by similar landscape data as employed in our study here. However, their employed models have not been directly validated against people's perceptions of the PSDs.

Here, we have used similar map data (although with higher spatial resolution) as in the latter studies, together with powerful modelling techniques and with more variables. Our results suggest that the PSDs cannot be easily predicted by such structural parameters alone. Instead, they indicate that the degree to which a PSD is perceived as strong in an environment largely depends on individual factors, such as the degrees to which a person identifies as nature or urban oriented. Other such individual factors might be of a more momentary nature, such as current mood or stress levels, while others might reflect more permanent personality traits. Neither respondents' gender nor educational level, however, significantly influenced the strength of the models in our study, in line with the findings presented by Qiu and Nielsen (2015). This warrants further research into how individual characteristics might shape perceptions of the PSDs. Overall, our results highlight the relevance of an ecological approach to perception (Chemero, 2009; Gibson, 1979) when interpreting perceived qualities such as the PSDs, that is, to regard them as perceived qualities highly dependent on the needs, abilities and perceptual framework of the individual and not on structural landscape characteristics alone.

According to Leslie et al. (2010), a general lack of agreement between objective and perceived measures is not surprising, since the two kinds of measures highlight different aspects of the world. There is, however, a commonly expressed need among various societal actors to translate key perceived qualities into quantifiable factors to create generally applicable design and planning guidelines and reliable tools for environmental evaluations. For such endeavours, results such as ours here present a challenge, as they suggest the need for finer levels of analysis when determining the strengths of PSDs for users in environmental planning. They arguably also put into question the validity of some past claims surrounding the PSDs, where these have been assumed to describe a more objective or universal truth about the landscape, presumably relevant for all users. Even if general connections between objective landscape features and people's perceptions of qualities such as the PSDs could be established, the influence of individual characteristics on such experiences is still likely significant.

4.4 | Strengths and weaknesses of the study

There are some caveats to our presented analyses. In the survey, we chose to count all untouched sliders as having been left in the middle deliberately (counted as 50) as long as any other slider had been interacted with. Likely, some of these sliders were left untouched because the respondent did not understand the statement, or felt that it could not be answered in a meaningful way for their favourite place and should thus have been removed from the analysis. We made the judgement, however, that the respondents leaving them in the middle due to such reasons was still less likely than the alternative, that they were left there on purpose. Qiu and Nielsen (2015) utilised a Yes/No/Don't know structure to their survey of PSDs and had only 8% 'Don't know' answers, showing that in general people can be expected to understand descriptions of these qualities. Furthermore, which qualities that are perceived at favourite places might not only reflect people's preferences but could also depend on the overall supply of qualities in the landscape. We did not ask the survey participants about how they experience the supply situation for each PSD in the available recreational landscape, and thus have no baseline to compare the favourite places to. Our survey also did not offer the opportunity for participants to enter additional information regarding perceived qualities other than the eight PSDs measured through the 0–100 sliders. We thus do not know whether these eight PSDs offer a sufficient basis for covering the main perceived qualities at favourite places in our study.

The response rate of our survey was rather low, with 20% starting the survey and 9% filling it out in full. Decreasing response rates to surveys is a trend (Stedman et al., 2019), especially for web-based surveys (Daikeler et al., 2020). Surveys with a strong local connection, as here, usually have higher response rates (Stedman et al., 2019). We believe the main issue here was technical: To reach the survey, the respondent had to either enter a URL by hand or scan a QR code. The survey was functional on mobile devices, but it was slightly more difficult to fill out; during data collection, we received several emails and phone calls from survey respondents who experienced difficulties. The data used for the analysis here stemmed from the second part of the survey, so was also subject to respondent attrition.

Due to the relatively small sample size, spatial and cultural delimitation etc., the generalisability of our findings to other cohorts or geocultural conditions could be questioned. More research is needed to determine the general validity of our results, and to further investigate the relations between structural landscape characteristics, perceived qualities and people's recreational needs. Methodologically the study might be interesting to replicate with a higher number of participants, across different geocultural conditions. Our predictive models suggest a strong influence on individual characteristics in shaping perceptions of the PSDs, emphasising the need for an ecological approach to perception when analysing such qualities, that is, to also consider how individual abilities and needs shape associated perceptions. The nature and extent of such individual factors are interesting for future studies to investigate further, since only a limited set was employed here. Future studies could also remedy our study's limitations regarding sample size and geocultural extension and include even more detailed landscape data to potentially identify stronger and more fine-tuned recreation indicators.

5 | CONCLUSIONS

Regarding physical landscape characteristics, our study suggests a general importance of forest and water for people when choosing a favourite site for recreation. It also largely confirms the overall relationships between perceived qualities suggested by previous research, while also indicating a division between two basic recreational attitudes. One seems more oriented towards outward-directed activities, with an emphasis on social experiences in cultivated or human-influenced settings. The other seems more rest-oriented, focused on experiences of serenity and freedom from disturbances in landscapes perceived as natural and free from human influence. Moreover, people commonly associate their favourite places with experiences related to vistas and openness, often while simultaneously being provided a sense of safety and shelter.

Our results also suggest that readily available landscape data might be insufficient to provide general predictions of the PSDs, possibly due to the importance of still largely unknown individual factors in shaping such perceptions. This might indicate broader limitations in how perceived qualities such as the PSDs can be represented in, for example, mapping or modelling scenarios. It presents a challenge for various aspects of, for example, landscape architecture, urban planning, rural development etc., where there is a wish to include such qualities side by side with other landscape measures to account for health and well-being effects. Further research is needed to increase the understanding of population- level relationships between structural landscape features, individual characteristics and perceived qualities of potential importance to support health and well-being. However, our results here might indicate a standing need for dialogue and engagement with local users as a complement to structural analyses when planning landscapes for recreational outcomes.

AUTHOR CONTRIBUTIONS

Jonathan Stoltz, Carl Lehto and Marcus Hedblom conceived the research idea and designed the methodology; Carl Lehto collected the data and ran the statistical analyses; Jonathan Stoltz, Carl Lehto and Marcus Hedblom interpreted the results. Jonathan Stoltz led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

A part of this research was funded by the FORMAS Research Council; the project 'Sustainable outdoor living environments—systematic interdisciplinary studies of health effects and impact on social inequalities' (D-nr 2019-01916). The authors would like to thank the Umeå Municipal government for supplying map data.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data and code used in the analysis are accessible at https://zenodo.org/doi/10.5281/zenodo.10060677.

ORCID

Jonathan Stoltz https://orcid.org/0000-0002-7139-2545
Marcus Hedblom https://orcid.org/0000-0003-1594-8665

REFERENCES

Adevi, A., & Grahn, P. (2012). Preferences for landscapes: A matter of cultural determinants or innate reflexes that point to our evolutionary background? *Landscape Research*, *37*, 27–49.

Aerts, R., Honnay, O., & Van Nieuwenhuyse, A. (2018). Biodiversity and human health: Mechanisms and evidence of the positive health effects of diversity in nature and green spaces. *British Medical Bulletin*, *127*(1), 5–22. https://doi.org/10.1093/bmb/ldy021

Annerstedt van den Bosch, M., Östergren, P.-O., Grahn, P., Skärbäck, E., & Währborg, P. (2015). Moving to serene nature may prevent poor mental health—Results from a Swedish longitudinal cohort study. *International Journal of Environmental Research and Public Health*, *12*, 7974–7989.

Appleton, J. (1975). *The experience of landscape*. John Wiley.

Beery, T. K., Jönsson, I., & Elmberg, J. (2015). From environmental connectedness to sustainable futures: Topophilia and human affiliation with nature. *Sustainability*, *7*(7), 8837–8854. https://doi.org/10.3390/su7078837

Björk, J., Albin, M., Grahn, P., Jacobsson, H., Ardö, J., & Wadbro, J. (2008). Recreational values of the natural environment in relation to neighbourhood satisfaction, physical activity, obesity and wellbeing. *Journal of Epidemiology and Community Health*, *62*, e2.

Bratman, G. N., Anderson, C. B., Berman, M. G., Cochran, B., De Vries, S., Flanders, J., Folke, C., Frumkin, H., Gross, J. J., Hartig, T., Kahn, P. H., Jr., Kuo, M., Lawler, J. J., Levin, P. S., Lindahl, T., Meyer-Lindenberg, A., Mitchell, R., Ouyang, Z., Roe, J., … Daily, G. C. (2019). Nature and mental health: An ecosystem service perspective. *Science Advances*, *5*(7), eaax0903. https://doi.org/10.1126/sciadv.aax0903

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, *3*(1), 1–27. https://doi.org/10.1080/03610927408827101

Chemero, A. (2009). *Radical embodied cognitive science*. MIT Press.

Daikeler, J., Bošnjak, M., & Lozar Manfreda, K. (2020). Web versus other survey modes: An updated and extended meta-analysis comparing response rates. *Journal of Survey Statistics and Methodology*, *8*(3), 513–539. https://doi.org/10.1093/jssam/smz008

de Jong, K., Albin, M., Skärback, E., Grahn, P., & Björk, J. (2012). Perceived green qualities were associated with neighborhood satisfaction, physical activity, and general health: Results from a cross-sectional study in suburban and rural Scania, southern Sweden. *Health & Place*, *2012*(18), 1374–1380.

de Jong, K., Albin, M., Skärback, E., Grahn, P., Wadbro, J., Merlo, J., & Björk, J. (2011). Area-aggregated assessments of perceived environmental attributes may overcome single-source bias in studies of green environments and health: Results from a cross-sectional survey in southern Sweden. *Environmental Health*, *2011*(10), 4.

Egorov, A. I., Mudu, P., Braubach, M., & Martuzzi, M. (Eds.). (2016). *Urban green spaces and health. A review of evidence*. World Health Organization, European Centre for Environment and Health.

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*(4), 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5). https://doi.org/10.1214/aos/1013203451

Gao, T., Song, R., Zhu, L., & Qiu, L. (2019). What characteristics of urban green spaces and recreational activities do self-reported stressed individuals like? A case study of Baoji, China. *International Journal of Environmental Research and Public Health*, *16*, 1348.

Gibson, J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.

Grahn, P., & Stigsdotter, U. A. (2003). Landscape planning and stress. *Urban Forestry & Urban Greening*, *2*(1), 1–18.

Grahn, P., & Stigsdotter, U. K. (2010). The relation between perceived sensory dimensions of urban green space and stress restoration. *Landscape and Urban Planning*, *94*, 264–275. https://doi.org/10.1016/j.landurbplan.2009.10.012

Grahn, P., Tenngart Ivarsson, C., Stigsdotter, U. K., & Bengtsson, I. L. (2010). Using affordances as a health-promoting tool in a therapeutic garden. In C. Ward Thompson, S. Bell, & P. Aspinall (Eds.), *Innovative approaches to researching landscape and health* (pp. 116–154). Taylor & Francis. https://doi.org/10.4324/9780203853252

Greenwell, B., Boehmke, B., & Cunningham, J. (2020). *gbm: Generalized boosted regression models* (Version 2.1.8). https://CRAN.R-project.org/package=gbm

Gunnarsson, B., Knez, I., Hedblom, M., & Ode, S. Å. (2017). Effects of biodiversity and environment-related attitude on perception of urban green space. *Urban Ecosystem*, *20*, 37–49. https://doi.org/10.1007/s11252-016-0581-x

Hartig, T., Mitchell, R., de Vries, S., & Frumkin, H. (2014). Nature and health. *Annual Review of Public Health*, *35*, 207–228.

Hartigan, J. A. (1975). *Clustering algorithms*. Wiley.

Hedblom, M., Hedenås, H., Blicharska, M., Adler, S., Knez, I., Mikusiński, G., Svensson, J., Sandström, S., Sandström, P., & Wardle, D. A. (2019). *Landscape perception: Linking physical monitoring data to perceived landscape properties*. Landscape Research.

Kaplan, S. (1995). The restorative benefits of nature: toward an integrative framework. *Journal of Environmental Psychology*, *15*, 169–182. https://doi.org/10.1016/0272-4944(95)90001-2

Kaplan, S., & Berman, M. (2010). Directed attention as a common resource for executive functioning and self-regulation. *Perspectives in Psychological Science*, *5*, 43–57. https://doi.org/10.1177/1745691609356784

Knez, I., & Eliasson, I. (2017). Relationships between personal and collective place identity and well-being in mountain communities. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00079

Leslie, E., Sugiyama, T., Ierodiaconou, D., & Kremer, P. (2010). Perceived and objectively measured greenness of neighbourhoods: Are they measuring the same thing? *Landscape and Urban Planning*, *2010*(95), 28–33.

Lindholst, A. C., Caspersen, O. H., & Konijnendijk van den Bosch, C. (2015). Methods for mapping recreational and social values in urban green spaces in the nordic countries and their comparative merits for urban planning. *Journal of Outdoor Recreation and Tourism*, *12*, 71–81.

Lockwood, A. (2017). *Balancing perceived sensory dimensions and biotopes in urban green space design*. The University of Guelph.

Maikov, K. (2013). Landscape characteristics in Tartu City parks: User influences through design. *WIT Transactions on Ecology and the Environment*, *179*, 353–364. https://doi.org/10.2495/SC130301

Mansor, M., Ghani, N., Harun, N. Z., & Zakariya, K. (2017). Conceptual models of greenspace and health. *Advanced Science Letters*, *23*, 6326–6330. https://doi.org/10.1166/asl.2017.9262

Marselle, M. R., Hartig, T., Cox, D. T., De Bell, S., Knapp, S., Lindley, S., Triguero-Mas, M., Böhning-Gaese, K., Braubach, M., Cook, P. A., de Vries, S., Heintz-Buschart, A., Hofmann, M., Irvine, K. N., Kabisch, N., Kolek, F., Kraemer, R., Markevych, I., Martens, D., … Bonn, A. (2021). Pathways linking biodiversity to human health: A conceptual framework. *Environment International*, *150*, 106420. https://doi.org/10.1016/j.envint.2021.106420

McMahan, E. A., & Estes, D. (2015). The effect of contact with natural environments on positive and negative affect: A meta-analysis. *The Journal of Positive Psychology*, *10*(6), 507–519. https://doi.org/10.1080/17439760.2014.994224

Memari, S., Pazhouhanfar, M., & Nourtaghani, A. (2017). Relationship between perceived sensory dimensions and stress restoration in care settings. *Urban Forestry & Urban Greening*, *26*, 104–113. https://doi.org/10.1016/j.ufug.2017.06.003

Molnar, C. (2018). iml: An R package for interpretable machine learning. *Journal of Open Source Software*, *3*(26), 786. https://doi.org/10.21105/joss.00786

OpenStreetMap Foundation. (n.d.). *OpenStreetMap*. https://www.openstreetmap.org/copyright

Pálsdóttir, A. M., Stigsdotter, U. K., Perqsson, D., Thorpert, P., & Grahn, P. (2018). The qualities of natural environments that support the rehabilitation process of individuals with stress-related mental disorder in nature-based rehabilitation. *Urban Forestry & Urban Greening*, *29*, 312–321. doi:10.1016/j.ufug.2017.11.016

Plieninger, T., Dijks, S., Oteros-Rozas, E., & Bieling, C. (2013). Assessing, mapping, and quantifying cultural ecosystem services at community level. *Land Use Policy*, *33*, 118–129. doi:10.1016/j.landusepol.2012.12.013

Pouwels, R., van Eupen, M., Walvoort, D. J. J., & Jochem, R. (2020). Using GPS tracking to understand the impact of management intervention on visitor densities and bird populations. *Applied Geography*, *116*(2020), 102154.

Qiu & Nielsen. (2015). Are perceived sensory dimensions a reliable tool for urban planning. *Landscape Research*, *40*(7), 1–21. https://doi.org/10.1080/01426397.2015.1029445

R Core Team. (2020). *R: A language and environment for statistical computing*. https://www.R-project.org/

Rousseeuw, P. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Schirpke, U., Holzler, S., Leitinger, G., Bacher, M., Tappeiner, U., & Tasser, E. (2013). Can we model the scenic beauty of an alpine landscape? *Sustainability*, *5*(3), 1080–1094.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423; 623–656.

Skärbäck, E., Björk, J., Stoltz, J., Rydell-Andersson, K., & Grahn, P. (2014). Green perception for well-being in dense urban areas—A tool for socioeconomic integration. *Nordic Journal of Architectural Research*, *26*, 179–206.

Skärbäck, E., Wadbro, J., & Grahn, P. (2009). GIS-analys på regional nivå av rekreationsresurser. *Svensk Geografisk Årsbok*, *85*, 67–96.

SLU. (2015). *The SLU forest map*. https://www.slu.se/en/Collaborative-Centres-and-Projects/the-swedish-national-forest-inventory/foreststatistics/slu-forest-map

Statistics Sweden. (2021). *Befolkningens utbildning Befolkning efter region, ålder, utbildningsnivå och år*. https://www.scb.se/hitta-statistik/sverige-i-siffror/utbildning-jobb-och-pengar/utbildningsnivan-i-sverige/

Stedman, R. C., Connelly, N. A., Heberlein, T. A., Decker, D. J., & Allred, S. B. (2019). The end of the (research) world As we know it? Understanding and coping with declining response rates to mail surveys. *Society & Natural Resources*, *32*(10), 1139–1154. https://doi.org/10.1080/08941920.2019.1587127

Stigsdotter, U. K., Corazon, S. S., Sidenius, U., Refshauge, A. D., & Grahn, P. (2017). Forest design for mental health promotion—Using perceived sensory dimensions to elicit restorative responses. *Landscape and Urban Planning*, *160*, 1–15.

Stoltz, J. (2022). Layered habitats: An evolutionary model for present-day recreational needs. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.914294

Stoltz, J., & Grahn, P. (2021). Perceived sensory dimensions: An evidence-based approach to greenspace aesthetics. *Urban Forestry & Urban Greening*, *59*(4), 126989. https://doi.org/10.1016/j.ufug.2021.126989

Stoltz, J., Lundell, Y., Skärbäck, E., van den Bosch, M. A., Grahn, P., Nordström, E. M., & Dolling, A. (2016). Planning for restorative forests: Describing stress-reducing qualities of forest stands using available forest stand data. *European Journal of Forest Research*, *2016*(135), 803–813.

Stoltz, J., & Schaffer, C. (2018). Salutogenic affordances and sustainability: Multiple benefits with edible forest gardens in urban green spaces. *Frontiers in Psychology*, *9*, 2344. https://doi.org/10.3389/fpsyg.2018.02344

Swedish Environmental Protection Agency (SEPA). (2018). *National groundcover data*. (In Swedish). https://www.naturvardsverket.se/verktyg-och-tjanster/kartor-och-karttjanster/nationella-marktackedata

Timonen, J., Siitonen, J., Gustafsson, L., Kotiaho, J. S., Stokland, J. N., Sverdrup-Thygeson, A., & Mönkkönen, M. (2010). Woodland key habitats in northern Europe: Concepts, inventory and protection. *Scandinavian Journal of Forest Research*, *25*(4), 309–324. https://doi.org/10.1080/02827581.2010.497160

Ulrich, R., Simons, R., Losito, B., Fiorito, E., Miles, M., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology*, *11*, 201–230.

UN, United Nations, Department of Economic and Social Affairs, Population Division. (2019). *World urbanization prospects: The 2018 revision (ST/ESA/SER.A/420)*. United Nations.

Vujcic, M., & Tomicevic-Dubljevic, J. (2017). Urban nature as a response to stress of urban population. *Lesson of Forestry Faculty*, *115*, 185–198.

WHO. (2021). *Non-communicable diseases*. https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Supplementary materials S1: Reclassification of CadasterENV land cover classes.

Supplementary materials S2: Lists of all influential predictors in boosted regression tree models.

How to cite this article: Stoltz, J., Lehto, C., & Hedblom, M. (2024). Favourite places for outdoor recreation: Weak correlations between perceived qualities and structural landscape characteristics in Swedish PPGIS study. *People and Nature*, *6*, 269–285. https://doi.org/10.1002/pan3.10574

APPENDIX A

Original survey statements, in Swedish

1. Platsen inger en känsla av vild och orörd natur
2. Platsen inger en känsla av att vara formad av människans hand
3. Platsen inger en känsla av öppenhet och ger möjlighet till utsikt och vyer
4. Platsen är en social yta som ger möjligheter att interagera med andra människor
5. Platsen inger en känsla av en helhet, av att vara en värld i sig själv
6. Platsen inger en känsla av mångfald och variation
7. Platsen inger en känsla av trygghet
8. Platsen inger en känsla av rofylldhet

Acta Universitatis Agriculturae Sueciae

Doctoral Thesis No. 2024:30

What makes a landscape attractive for outdoor recreation? This thesis explores the question through two surveys on Swedish recreationists and a review of scientific literature. Movement patterns of recreationists is analysed with novel methods, including machine learning and viewshed analysis. The results reveal that recreational infrastructure, proximity to water, and absence of urban noise are important factors. The thesis further discusses how such results can be used to create an index for assessing the recreational potential of landscapes.

Carl Lehto received his PhD education at the department of Ecology, SLU, Uppsala. He received his MSc in Ecology and Conservation at Uppsala University, where he also studied for his BSc in biology.

Acta Universitatis Agriculturae Sueciae presents doctoral theses from the Swedish University of Agricultural Sciences (SLU).

SLU generates knowledge for the sustainable use of biological natural resources. Research, education, extension, as well as environmental monitoring and assessment are used to achieve this goal.

ISSN 1652-6880
ISBN (print version) 978-91-8046-324-9
ISBN (electronic version) 978-91-8046-325-6