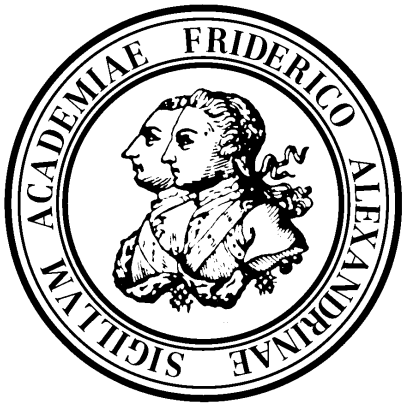
Investigating associations between multiple greenspace exposures and symptoms of depression

Results from the Prospective Urban and Rural Epidemiology (PURE) Study

Bachelorarbeit im Bachelor-Studiengang

Physische Geographie

Friedrich-Alexander-Universität Erlangen-Nürnberg
Institut für Geographie



Sebastian Brinkmann

Matrikel-Nr.: 22368762

Betreuer(in): Prof. Dr. Blake Walker

Erlangen, 2022

Contents

Figures

Tables

Abbreviations

AHEI	Alternative Healthy Eating Index
BART	Bayesian Additive Regression Trees
CGEI	Composite Greenspace Exposure Index
CI	Confidence Intervals
DA	Dissemination Area
DSM	Digital Surface Model
DTM	Digital Terrain Model
GACI	Greenspace Accessibility Exposure Index
GAVI	Greenspace Availability Exposure Index
GIS	Geographic Information System
IDW	Inverse Distance Weighting
KMO	Kaiser-Meyer-Olkin Index
LAI	Leaf Area Index
LOS	Line of Sight
LULC	Land Use Land Cover
MAUP	Modifiable Areal Unit Problem
MDE	Major Depressive Episodes
MSA	Measure of Sampling Adequacy
NDVI	Normalised Difference Vegetation Index
OR	Odds Ratio
OSM	Open Street Map
PCA	Principal Components Analysis
PDP	Partial Dependence Plot
SES	Socioeconomic Status
SMOTE	Synthetic Minority Oversampling Technique
SV	Street View
VGVI	Greenspace Visibility Exposure Index
VSDI	Vancouver Socioeconomic Deprivation Index

1 Introduction

Mental health is a fundamental component of health and well-being (La Placa et al., 2013; Organization, 1948) but over the last 30 years, mental illnesses have become among the highest burdens worldwide (Friedrich, 2017; Liu et al., 2020). This change may, at least partially, be explained by rapid urbanisation (Penkalla & Kohler, 2014). As of 2018 over half of the world's population lives in urban settlements and by 2030 this proportion is estimated to increase to 60% (United Nations, 2018). Five decades ago, only three megacities (>10 million inhabitants) existed: Tokyo, Osaka and New York. As of 2018, there were 33 megacities and by 2030 there will be 43 (United Nations, 2018).

Even though urban areas have the potential to support health (Lennard, 2018; Milgram, 1970), studies suggest that these areas undermine mental health by promoting disconnectedness from nature (Hartig & Kahn, 2016; Soga & Gaston, 2016; Turner et al., 2004). Increased air and noise pollution, heat islands, lack of green space and sedentary behaviour are a few of the many issues created by urban environments. Furthermore, there is considerable variation in the levels of such environmental exposures within and between cities (Nieuwenhuijsen, 2016). With rapid urbanisation it seems important to improve our understanding about the effects of urban environments on mental health, to support mental health as an ecosystem service (Bratman et al., 2019).

The connection between mental health and the natural environment is widely recognised (Kondo et al., 2018; Moore et al., 2018; Moreira et al., 2021; White et al., 2021; Zhang et al., 2021). Several reviews have shown that residential greenness (i.e. overall vegetation) and green spaces such as parks, community gardens or forests (henceforth both types of urban nature are generalised as "greenspace"; Taylor and Hochuli, 2017) may have positive effects on a variety of health-related measures (e.g. Dadvand et al., 2016; Dzhambov et al., 2014; Markevych et al., 2017; Rautio et al., 2017). Previous studies have identified neighbourhood greenspace exposure with lower levels of symptomatology for depression, anxiety and stress (Beyer et al., 2014), improved sleep (Grigsby-Toussaint et al., 2015) and improved mental wellbeing (Houlden et al., 2019). However, some studies also found mixed (Dzhambov et al., 2018; White et al., 2017), weak or insignificant results (Huynh et al., 2013; White et al., 2013). The inconsistencies in the literature can, at least partially, be explained due to a lack of consensus regarding spatial methods for measuring and modelling greenspace exposure and their pathways (Davis et al., 2021; Dzhambov et al., 2020; Giles-Corti et al., 2018; Labib et al., 2020a). Furthermore, studies often use different types of greenspace exposures (e.g. greenspace visibility or access to parks) that might have different effects on mental health. A deeper understanding of these factors is needed for providing spe-

cific recommendations for decision makers and urban designers (Davis et al., 2021; Koohsari et al., 2015).

In a recent study by Labib et al. (2021b), the authors introduced a multiple greenspace exposure framework to address the aforementioned concerns. In this framework three types of greenspace (i.e. availability, visibility, and accessibility of greenspace) were joined in a Composite Greenspace Exposure Index (CGEI). This novel index differs from traditional single exposure metrics in that it integrates multiple types of greenspace to capture contact to nature holistically and objectively. However, as of now there has been no study to apply this novel index on individual-level data. Furthermore, there are no recommendations on how to parameterise the different single greenspace layers of the CGEI in the context of mental health. Therefore, this study aims to calculate the CGEI for Vancouver, Canada, to investigate associations between contact to nature and depressive symptoms.

2 Background

2.1 Potential underlying pathways linking greenspace to health

In the context of this study, the term pathway describes underlying biopsychosocial mechanisms that can lead to health benefits. As there are a manifold of potential pathways (e.g. Dadvand and Nieuwenhuijsen, 2019; Hartig et al., 2014) Markevych et al. (2017) provide an organisational approach (Fig. 1), which emphasises three general mechanisms of greenspace: reducing harm, restoring capacities and building capacities. In the following I provide a brief overview of pathways within each domain.

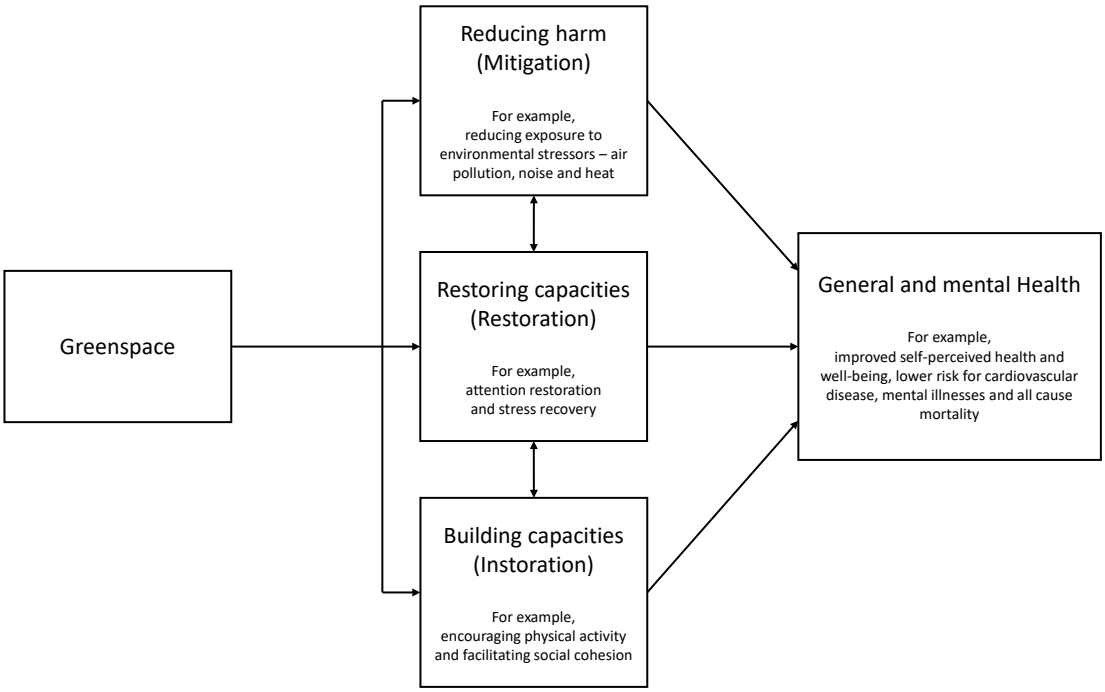


Figure 1 Conceptualisation of pathways linking multiple greenspace exposures to positive health outcomes. The arrows represent hypothetical relationships between greenspace, specific pathways and health outcomes (Figure adapted from Fig 1. in Markevych et al., 2017).

2.1.1 Reducing harm (Mitigation)

Urban environments provide potentially harming effects like air pollution, high air temperature and noise through traffic (Dadvand & Nieuwenhuijsen, 2019). Researchers have found that air pollutant concentrations are lower near greenspace (e.g. Nowak et al., 2014). One reason for this is that plants remove pollutants from the air through

stomatal uptake or non-stomata deposition on the plants surfaces (Givoni, 1991; Paoletti et al., 2011). As a consequence, traffic-related air pollution is lower in areas with high levels of greenspace. Furthermore, studies have shown that levels of traffic-related air pollution are lower with proximity to urban parks, and in residential areas and schools with higher surrounding greenspace (Dadvand et al., 2012; Dadvand et al., 2015; Su et al., 2011). Studies also demonstrate the impact of air pollution on mental disorders such as depression (Ali & Khoja, 2019) and cognitive development in children (Dadvand et al., 2015).

Urban features like high-rise buildings, industry, traffic or the use of materials such as concrete lead to an increase in air-temperature. This effect is called urban heat island (Heaviside et al., 2017). It is typically attributed to the replacement of vegetation with such man-made features, which absorb more sunlight, causing a local enhancement in the urban heat budget, as well as reducing wind speed (Phelan et al., 2015; Voogt & Oke, 2003). Exposure to heat is associated with an increase in hospitalisation rates and death (Basu, 2009) and enhances other causes such as respiratory illness (D'Ippoliti et al., 2010). While extreme heat events are associated with higher rates of emergency department visits due to mental health-related conditions, the link between elevated temperature and specific mental health conditions remains poorly understood (Nori-Sarma et al., 2022).

Research about the effects of greenspace on reducing traffic noise has limited and mixed evidence (Dadvand & Nieuwenhuijsen, 2019). Vegetation builds a physical barrier which has the potential of reducing noise levels by 5-10 dB through absorption or diffraction of sound waves (Van Renterghem et al., 2015). However, in some cases such as street canyons, this barrier may actually increase pedestrian level noise exposure by reflecting the sound waves (Jang et al., 2015). Exposure to noise has been associated with several health effects such as cardiovascular disease, metabolic disorders, or cognitive impairment (Basner et al., 2014). While there are some studies that analyse the effects of residential greenspace on noise and sleep quality (e.g. Xie et al., 2020), further research explicitly analysing how noise mediates the associations between greenspace and mental health is needed (Markevych et al., 2017).

2.1.2 Restoring capacities (Restoration)

Numerous experiments and observations show that exposure to green spaces can significantly lower stress and restore cognitive function. The stress reduction theory suggests that natural environments promote reduction in stress. In the context of the theory of evolution, landscapes with water or vegetation have been seen as safe havens. Therefore, seeing or being present in such environments produces affective reactions that lead to stress reducing effects (Ulrich, 1981; Ulrich et al., 1991).

The attention restoration theory proposes that exposure to nature has the potential to restore the mind's ability to actively focus. In the context of this theory there are two types of attention: directed attention is required to (actively) focus on cognitive stimuli. After prolonged use, we suffer from "attention fatigue". Involuntary attention on the other hand describes indirect (i.e. effortless) attention utilising cognitive faculties not normally used, therefore allowing the neural mechanisms underlying directed attention to restore. There are four essential components an environment must contain for it to promote involuntary attention: (1) *being away* which describes opportunities to gain distance from thoughts and concerns. (2) *soft fascination* refers to stimuli that hold human attention effortlessly. The (3) *extent* describes the scope of the experience and the possibility of feeling immersed within it, and (4) *compatibility* of an environment so it may provide a "match" between the individual needs and desires and the environment. Natural environments usually contain all components simultaneously, therefore providing the possibility of directed attention restoration (Kaplan & Kaplan, 1989). These pathways have been shown to act as beneficial influences of greenspace on mental health (Dadvand et al., 2016).

2.1.3 Building capacities (Instoration)

Greenspace is likely to build capacities for example through facilitating social cohesion and encouraging physical activity. Social cohesion within a neighbourhood describes a feeling of belonging, respect of each other and safety. As residential greenspace such as parks provide opportunities for social contact it is likely to increase social cohesion (Holtan et al., 2014; Weinstein et al., 2015). Research shows that strengthening social cohesion within the neighbourhood could reduce poor mental health and increase wellbeing (Fone et al., 2014; Williams et al., 2020), reduce negative effects on adolescent mental health following stressful life events (Kingsbury et al., 2019) and improve general health (Dadvand et al., 2016). However, not all greenspace has the same effect on social cohesion. In this context, parks for example with opportunities for social contact are different to an isolated forest path. Therefore, more refined greenspace measures are required (Markevych et al., 2017).

In addition to increased social cohesion, greenspace is also likely to increase and improve the benefits of physical activity (Duncan et al., 2014; Lachowycz & Jones, 2011; McGrath et al., 2015). Previous studies suggest that physical activity may play an important role in the management of mental health diseases and improve wellbeing (Biddle et al., 2003; Paluska & Schwenk, 2000). In accord with these findings, studies show that "green exercise" has greater psychological benefits compared to the same exercise settings with relatively little greenspace (Mitchell, 2013; Pretty et al., 2005). However, as it is the case for social cohesion, not all greenspace has the same effects on physical activity and mental health. The sole availability of greenspace (e.g. vege-

tation in backyards) is likely to be a poor indicator for opportunities of "green exercise". Therefore, more refined characteristics such as size and good access to public parks are necessary to analyse associations of greenspace with physical activity such as recreational walking (Giles-Corti et al., 2005).

2.2 Multiple Greenspace Exposure Framework

Although most individual-level studies have focused on one particular pathway from nature to health, it is clear that multiple pathways are likely to be engaged simultaneously and affect one another (Hartig et al., 2014; Markevych et al., 2017). For example, people entering a park for social cohesion purposes generally need to engage in some form of physical activity to do so. Previous studies often consider only one of multiple possible greenspace exposures (e.g. overall greenness or distance to parks) when analysing the connection between health and the natural environment (Hoffimann et al., 2017; Schüle et al., 2017).

As contact with nature is invariably integrated into a complex web of health determinants, accounting for multiple exposures addresses this complexity (Frumkin et al., 2017). For that purpose in this study a multiple greenspace exposure framework has been used accounting for availability, visibility, and accessibility of greenspace (Fig. 2) (Dadvand & Nieuwenhuijsen, 2019; Labib et al., 2021b) representing the most commonly used spatial measures of greenspace exposure (Labib et al., 2020a).

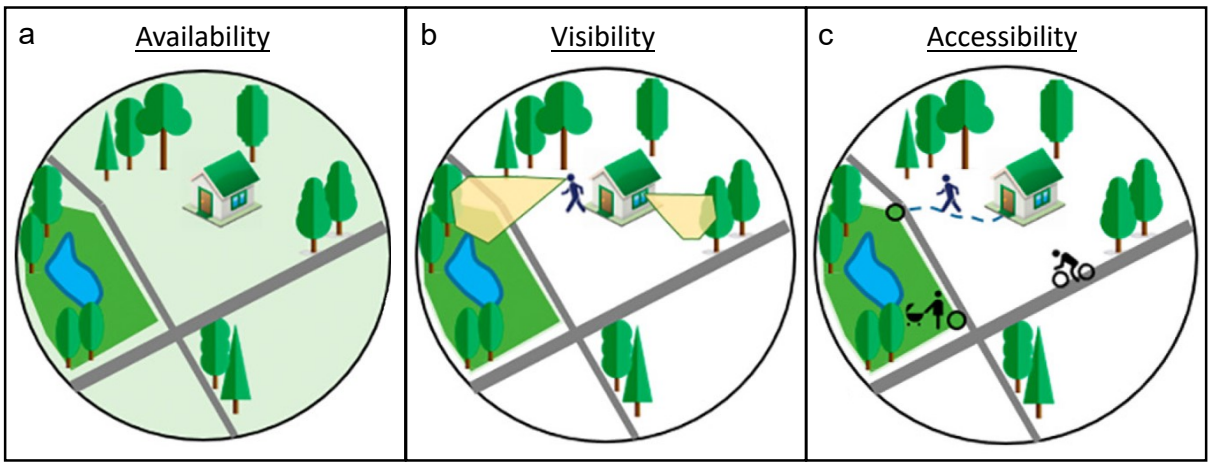


Figure 2 A schematic illustration of the multiple greenspace exposure framework (Figure adapted from Fig 1. in Labib et al., 2021b).

2.2.1 Availability

Availability of greenspace (Fig. 2a) describes the physical amount of vegetation. Typically it is estimated by simple map-bound 2D buffer analysis around the home or work place using remote sensing data. Indices such as the Normalised Difference

Vegetation Index (NDVI) to measure the photosynthetically active vegetation or Land Use Land Cover (LULC) maps to quantify the percentage of green land cover are commonly used (Labib et al., 2020a). To account for seasonal effects and different qualities of greenspace which can be measured using different vegetation indices, previous studies suggest using multiple remote sensing products at different scales when measuring greenspace availability (Labib et al., 2020b; Markevych et al., 2017). The sole availability of residential greenspace is relevant for all three domains explained in 2.1. However, it plays an especially important role when accounting for mitigation pathways (reducing harm) (Markevych et al., 2017).

2.2.2 Visibility

Visibility of greenspace (Fig. 2b) describes the visual amount of greenspace that can be seen from an eye-level perspective of a study participant. One methodology to estimate the exposure to visible greenspace describes using Street View (SV) images (e.g., Google Street View, Baidu Street View). While this method is most commonly used in the literature it has several shortcomings, such that SV images are typically limited to roads accessible by car and seasonal inconsistency between SV images (Li et al., 2015). Furthermore, it is still difficult to accurately classify vegetation from SV images due to many factors, such as shadows and confusion between human made green features and vegetation (Li et al., 2015). Geographic Information System (GIS) based viewshed analysis describes an alternative method for calculating eye-level greenness visibility. Recent studies have demonstrated the use of city-wide scale viewshed-based greenness visibility as a highly accurate alternative to SV-based visual exposure (Cimburova & Blumentrath, 2022; Labib et al., 2021a; Tabrizian et al., 2020). Such methodologies are not dependent on the availability of SV-images and recent work has improved computation time so that the analysis of visible greenspace can be scaled to large areas with little effort (Brinkmann et al., 2022). Greenspace visibility measurements are relevant when accounting for the restoration pathways (restoring capacities) (Markevych et al., 2017).

2.2.3 Accessibility

Accessibility of greenspace (Fig. 2c) describes the physical and legal access to greenspaces such as parks, community gardens or forests. To assess proximity to such greenspace researchers could (a) include questions about the walking distance to the nearest park in a questionnaire or (b) calculate the distance to the closest park from a participant's home. For that either the euclidean distance or a road network analysis can be used. However, structural or legal access as non-disclosed or non-spatial information is hard to obtain, especially for larger areas (Labib et al., 2020a).

Greenspace accessibility is relevant when accounting for the instoration (building capacities) and restoration (restoring capacities) pathways (Markevych et al., 2017).

2.3 The role of Socioeconomic Status

The role as a significant covariate and effect modifier of individual and neighbourhood level Socioeconomic Status (SES) in the context of health is widely recognized (Hajat et al., 2021; Markevych et al., 2017; Scarpone et al., 2020; Walker et al., 2022; Walker et al., 2019). Studies have shown that populations with low SES often live in areas with higher pollution and lower greenspace exposures, and have worse health status (Hoffimann et al., 2017; Schüle et al., 2017). Furthermore, groups with higher SES are more likely to engage with greenspace farther away from their homes because of higher mobility. In contrast, those with lower SES are generally less mobile and spend more time in proximity to their homes, increasing the importance of their immediate greenspace (Hoffimann et al., 2017; Maas et al., 2009). As a consequence the beneficial effects of greenspace on health are strongest for those living in neighbourhoods with low SES and those with low individual-level SES. Therefore, it has been suggested to include SES as a covariate when analysing environmental effects on health (Markevych et al., 2017).

3 Methods

Data preparation and analysis were conducted using R (v.4.2) and the C++ programming languages on a Linux system (AMD Ryzen 9 3900x CPU, 64GB DDR4). Code and detailed documentation for all processing and modelling steps are available at https://github.com/STBrinkmann/Bachelorthesis.

3.1 Participant Data

The PURE study is an ongoing global epidemiological study in 27 countries. For the purpose of this study a subset of 1,607 participants recruited between 2006 and 2009 in Vancouver, Canada has been used. All participants gave their informed consent to be part of the study and the study was approved by the local institutional review board (Teo et al., 2009).

Self-reported participant variables comprised age, sex, household income range, tobacco and daily alcohol consumption. In addition, the dietary choices and intake of nutrients of each participant were assessed using the Alternative Healthy Eating Index (AHEI), a nine-component index designed to assess dietary-based disease risk (McCullough et al., 2002). Symptomatology for Major Depressive Episodes (MDE) was assessed using an adapted Short-Form Composite International Diagnostic Interview (CIDI-SF) for MDE (Gigantesco & Morosini, 2008; Kessler & Üstün, 2004). Participants have been asked if they had felt depressed, sad, or blue for two weeks or more in a row in the past 12 months. If so, they have been further asked whether they experienced loss of interest, felt tiredness, had trouble sleeping or concentrating, have gained or lost weight, or if they had thoughts about death. For each participant the sum of positive answers has been taken and risk of MDE was classified if a participant had a score of three or more. All participant variables are listed in Table 2. Furthermore, for each participant a robust residential address has been provided which has been geolocated using the Open Street Map (OSM) Nominatim API (OpenStreetMap contributors, 2017).

3.2 Environmental Data

Following the recommendation of previous research of using multiple greenspace metrics at different scales, this study utilised four different greenspace metrics commonly used in the existing literature, (i) NDVI, (ii) Leaf Area Index (LAI), (iii) LULC, and (iv) public parks (Fig. 3) (Labib et al., 2020b; Markevych et al., 2017). As this study focuses on the effects of greenspaces, bluespaces like lakes, rivers or the sea have been removed from all greenspace metrics using a water mask derived from the high resolution LULC raster. Both NDVI and LAI were derived from a cloud-free Sentinel-2 L1C satellite image (04.10.2015) which has been acquired through the EO Browser platform. The L1C level (Top-Of-Atmosphere) already includes radiometric and geometric corrections but no atmospheric correction (Drusch et al., 2012). Therefore,

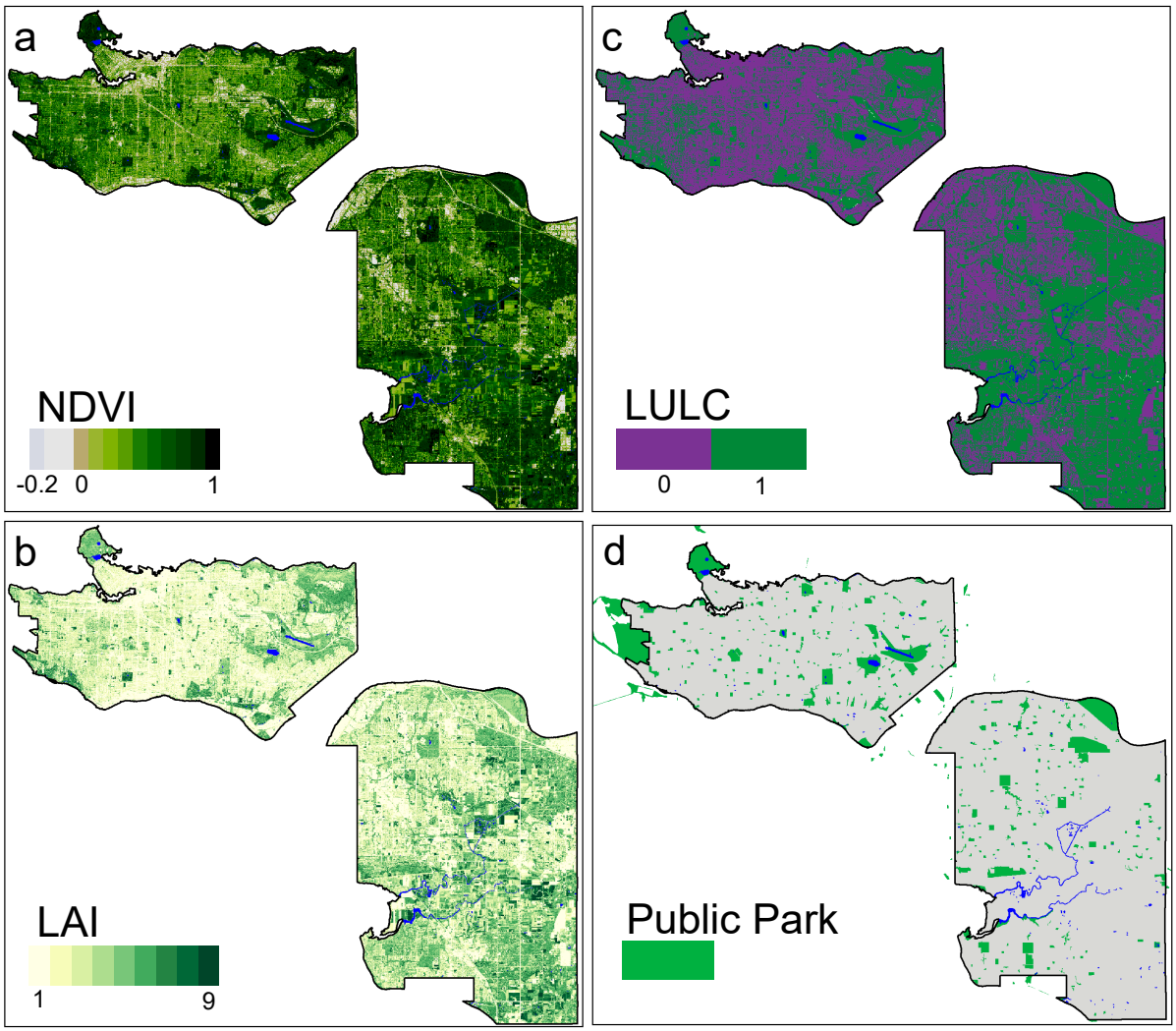


Figure 3 Greenspace metrics used in the study area. Blue areas represent terrestrial water.

as part of the preprocessing, the Sentinel-2 image has been corrected to L2A level (Bottom-Of-Atmosphere) using the Sen2Cor processor algorithm (Main-Knorn et al., 2017) provided in the ESA SNAP toolbox.

The NDVI (Fig. 3a) characterises greenspace density and was calculated at 10 m spatial resolution using the standard equation (Eq. 3.1):

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{3.1}$$

where NIR refers to the near-infrared band and RED refers to the visible red wavelengths (Drusch et al., 2012). The LAI (Fig. 3b) characterises greenspace volume. It measures the amount of photosynthetic area and levels of transpiration of the vegetation. These are key parameters for regulating ecosystem functions such as air temperature regulation and air pollution reduction (Lin & Lin, 2010; Nowak et al., 2014). The LAI has been calculated at 10 m spatial resolution using the biophysical processor algorithm (Weiss et al., 2020) included in the ESA SNAP toolbox.

Publicly available LULC data from 2014 has been acquired by Metro Vancouver (2019) at 2 m resolution and was reclassified to a binary greenspace raster, where tree canopy, shrub and grass represent green (value of 1) and other classes - including non-photosynthetic vegetation - indicate not-green (value of 0). The binary LULC layer (Fig. 3c) represents the overall presence or absence of greenspace. The LULC map was created using RapidEye 5 m multispectral satellite imagery and LiDAR data.

An official, local dataset of parks was publicly available, however only for a small sub-region of Vancouver. Therefore, polygon shapefiles of public parks of the complete study area (Fig. 3d) have been acquired by OSM (OpenStreetMap contributors, 2017). To account for edge effects data has been downloaded using a buffer of 1 km around the study area. Small features (< 1 ha) and those without the term “park” in their name have been removed, resulting in 456 parks with a mean (sd) size of 13.35 ha (58.42 ha).

To analyse the visibility of greenspace, a LiDAR derived Digital Surface Model (DSM) and Digital Terrain Model (DTM) at 1 m resolution has been used for its proven ability to represent the above-ground elements and its accuracy in estimating surface visibility (Van Berkel et al., 2018). The DSM has been used to account for ground surface objects like trees or buildings, and the DTM to represent the ground terrain. Both elevation models are publicly available (Natural Resources Canada, 2019). In order to compute network-based distance metrics, streets data has been acquired from OSM using the R-package osmdata (Padgham et al., 2017). Road types not suitable for walking were removed (i.e., motorways, trunks, and raceways). The network data was topologically corrected and split into 20 metre-long segments using the R package nngeo (Dorman, 2022).

To account for neighbourhood SES, data from the 2006 Canadian census has been acquired from Statistics Canada at the regional level of Dissemination Area (DA) (the smallest available census area with an average population of 400-800 residents). Following the recommendations of previous research, eleven SES variables have been included in this analysis and are listed in Table 2 (Hajat et al., 2021; Walker et al., 2022).

3.3 Greenspace Exposure Modelling

In this study, I combine three previously developed individual metrics of greenspace exposure: availability, visibility, and accessibility (Section 2.2). Each of these exposure metrics used various geographic data following a particular methodology. A raster map of each metric has been calculated for the complete study area at 10 m resolution (3975 $\times$ 3479 cells). The following sections summarise the approaches taken for calculating these three different types of greenspace exposure metrics.

3.3.1 Availability

Availability of greenspace is often measured by applying a simple 2D buffer around the home location and aggregating the values within this area (e.g. mean NDVI of a 300 m buffer). When applied on a raster map, the neighbourhood of all cells will be summarised using a moving window with a certain distance threshold. The Greenspace Availability Exposure Index (GAVI) (Labib et al., 2020b) used in this study combines three commonly used greenspace metrics (i.e. NDVI, LAI, and LULC) at five spatial scales (i.e., 50, 100, 200, 300, and 400 m buffer distance) as a multi-scale, multi-metric map. The three greenspace metrics represent different characteristics of photosynthetically active vegetation (Section 3.2). Multiple spatial scales have been used as they (i) represent different ecosystem functions, and (ii) account for scale dependent statistical inference.

Multiple scales and the modifiable areal unit problem

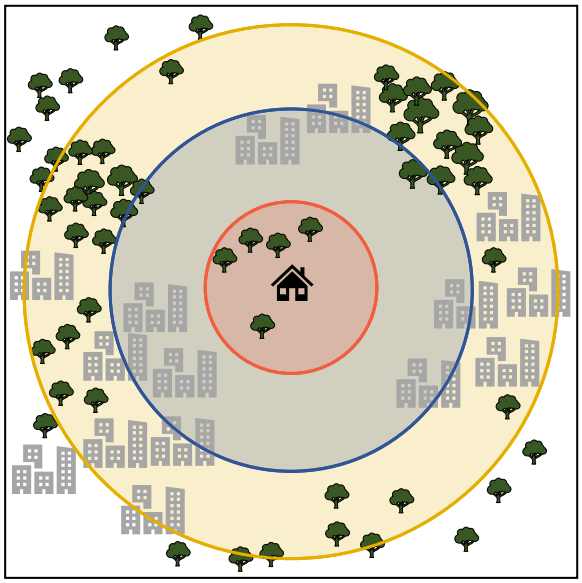


Figure 4 Schematic illustration of the effects of multiple scales as differing buffer size result in different aggregated values.

As of now, there is no consensus what exact buffer size is suitable when analysing mental illnesses. In a review by Labib et al. (2020a), the researchers found that when capturing the effects of greenspace on physical health studies used values of 300 - 600 m and 601 - 1000 m as the most and second most used buffer distances, respectively. In the context of mental health, relatively small values of less than 300 m and 300 - 600 m as the most and second most used buffer distances, respectively, have been used. These mixed results in regards to an “optimal” buffer distance seem intuitive as the selection of a buffer size may represent different ecosystem functions. For example, a smaller buffer (< 100 m) may represent physical vegetation barriers (e.g. scrub and trees) along the roadways, which may mitigate the effects of traffic related air and noise pollution, whereas a larger buffer might better capture greenspace that is being used for (recreational) physical activity (Markevych et al., 2017). As a consequence, recent studies applied multiple buffer thresholds to analyse the effects at distinct distance levels or even combined them in a multi-scale map (e.g. Browning et al., 2019; Labib et al., 2020b; Su et al., 2019).

The second argument for using multiple spatial scales is that they account for scale dependent statistical inference. In geographical statistics, the Modifiable Areal Unit Problem (MAUP) refers to the problem that analytical results are sensitive to the area for which data are collected (Fotheringham & Wong, 1991; Openshaw, 1984). As a consequence, variations in scale lead to differences in model outcomes, and as such, different process understanding (Comber et al., 2022). As visualised in Fig. 4 using different euclidean buffer distances from a home location has different outcomes. While a small buffer (red circle) only captures the greenspace close to the home location, a medium size buffer (blue circle) additionally captures the urban structures visualised as grey buildings. When using a large buffer distance (orange circle) the ratio between urban features and greenspace gets approximately even. Furthermore, when comparing multiple locations, the variance of the aggregated buffers decreases with increasing buffer distance (Comber et al., 2022).

Lacunarity

In a recent paper Labib et al. (2020b) addressed the MAUP by introducing the mathematical concept of *lacunarity*. Lacunarity literally refers to the *gappiness* or *heterogeneity* of a fractal or non-fractal image. When aggregating all cells in a raster map the resulting raster will become increasingly homogeneous as the buffer distance increases, and at a certain level all cells will have the same value. As a scale dependent measure of heterogeneity, Dong (2000) introduced lacunarity for spatial heterogeneity measurements in GIS. Based on the research of Plotnick et al. (1993), the author implemented a lacunarity algorithm for binary raster images. Hoechstetter et al. (2011) further improved lacunarity measurements for continuous images. In a recent paper, Labib et al. (2020b) applied lacunarity on binary (i.e. LULC) and continuous (i.e. NDVI or LAI) raster images to derive scale sensitive weights used in a buffer analysis to build a multiple-exposure greenspace availability metric.

To compute lacunarity a sliding box algorithm that is based on previous works has been used (Hoechstetter et al., 2011; Labib et al., 2020b; Plotnick et al., 1993). To calculate lacunarity $\Lambda(r)$ ("lambda") for a spatial raster M (representing NDVI, LAI, or LULC) with width W and length L, a square box of size $r \times r$ (initial value $r = 2$) is placed upon the top-left corner of M. The box represents the neighbourhood of a certain raster cell and r the distance value at which lacunarity is being calculated. After that the box is shifted one cell to the right, so that the new box overlaps with the previous one. First, the total number of boxes $N[r]$ can be described as (Eq. 3.2):

$$N[r] = (W - r + 1)(L - r + 1) \tag{3.2}$$

Next, the box mass $S(r)$ for each box of size r is being calculated by taking the box sum, or the range of all box-values, for binary or continuous raster, respectively. For *binary* (i.e. LULC) images the number of boxes of size r containing box mass S are

counted as $n[S,r]$, and converted into a probability distribution $Q(S,r)$ by dividing by $N[r]$ (Eq. 3.3):

$$Q(S,r) = \frac{n[S,r]}{N[r]} \tag{3.3}$$

Next, the first and second moments of this probability distribution are estimated using Eq. (3.4), (3.5), respectively:

$$Z(1) = \sum S(r) \times Q(S,r) \tag{3.4}$$

$$Z(2) = \sum S^2(r) \times Q(S,r) \tag{3.5}$$

Finally, lacunarity $\Lambda(r)$ can be calculated as (Eq. 3.6):

$$\Lambda(r) = \frac{Z(2)}{Z^2(1)} \tag{3.6}$$

If the raster is *continuous* (i.e. NDVI or LAI) the first and second moments can be expressed as the mean $E[S(r)]$ and variance $Var[S(r)]$ of the box mass values, respectively (Plotnick et al., 1993). Lacunarity $\Lambda(r)$ can now be described as (Eq. 3.7):

$$\Lambda(r) = 1 + \frac{Var[S(r)]}{E^2[S(r)]} \tag{3.7}$$

Table 1 Box sizes and weights at five scales for the three greenspace metrics of both sub-regions used for calculating the Greenspace Availability Exposure Index (GAVI).

Box size radius (in metres)	**Scale [i]**	**Weighting- [w] (Lacunarity value)**		
		NDVI	LAI	LULC
Vancouver				
50	1	0.086	0.216	0.541
100	2	0.053	0.160	0.488
200	3	0.040	0.114	0.443
300	4	0.040	0.094	0.421
400	5	0.043	0.083	0.407
Surrey				
50	1	0.144	0.185	0.297
100	2	0.071	0.119	0.260
200	3	0.037	0.078	0.231
300	4	0.031	0.062	0.220
400	5	0.030	0.054	0.216

An established procedural method for interpreting lacunarity of an image is plotting the natural logarithms of both $\Lambda(r)$ and r (Hoechstetter et al., 2011). Fig. 5 visualises three simulated elevation models with an extent of 100 $\times$ 100 pixels and values ranging from 0 (green) to 1 (white). These three images describe different landforms. Fig. 5a and 5b both show a continuous landscape without gaps, however, the hills in 5a have a larger diameter (25 pixels) compared to 5b (10 pixels). Fig. 5c shows a simulated landscape with large hills (25 pixels) but contains two gaps. As a consequence of these holes, lacunarity decreases relatively slowly representing the high gappiness of this image. Due to the small radius of the features in Fig. 5b, increasing the box size r quickly leads to a homogeneous pattern of values and consequently to low lacunarity.

Multi-scale and multi-metric exposure

To combine the multiple greenspace metrics from multiple scales, first lacunarity has been calculated for each metric and scale. The five scales 50, 100, 200, 300, and 400 m have been used as suggested by the literature (see Section 3.3.1). As visualised in Fig. 3, there are large-scale variations in the landscape patterns within the study area. Therefore, lacunarity has been calculated independently for both sub-regions - Vancouver in the north-west, and Surrey in the south-east. The resulting lacunarity values are listed in Table 1 and have been used as objective weights in a spatial decay function, so that the reduced variance with increasing spatial scale is accounted for by the corresponding lacunarity values. Lacunarity has been calculated using the spatLac R package (Brinkmann, 2021) following the above mentioned algorithm. To aggregate the raster images at the specific scales a standard focal function from the terra R package (Hijmans, 2022) has been used.

Following the procedure of Labib et al. (2020b), combined multi-scale exposure maps CM_m for each greenspace metric m with $m = NDVI, LAI, LULC$ were computed as (Eq. 3.8):

$$CM_m = \frac{\sum_{i=1}^{L}(M_i \times w_i)}{\sum_{i=1}^{L} w_i} \tag{3.8}$$

L describes the number of scales (in this study $L = 5$), M_i the aggregated greenspace metric m at scale i, and w_i is the corresponding weight of the lacunarity analysis at scale i (Table 1). The resulting maps are being called CM_{NDVI}, CM_{LAI}, and CM_{LULC}. Each of these combined multi-scale exposure metrics were reclassified into nine categories using the Jenks natural breaks classification method (Jenks, 1977), where the lowest and highest exposure is represented as 1 and 9, respectively. Finally, the multi-scale, multi-metric GAVI is calculated using (Eq. 3.9):

$$GAVI = \frac{CM_{NDVI} + CM_{LAI} + CM_{LULC}}{3} \tag{3.9}$$

In a last step, the GAVI of both sub-regions has been combined and again reclassified using the Jenks natural breaks classification method, so that extremely low and

extremely high levels of greenspace availability are represented as 1 and 9, respectively. The final GAVI map has been masked with the study area and a water shapefile.

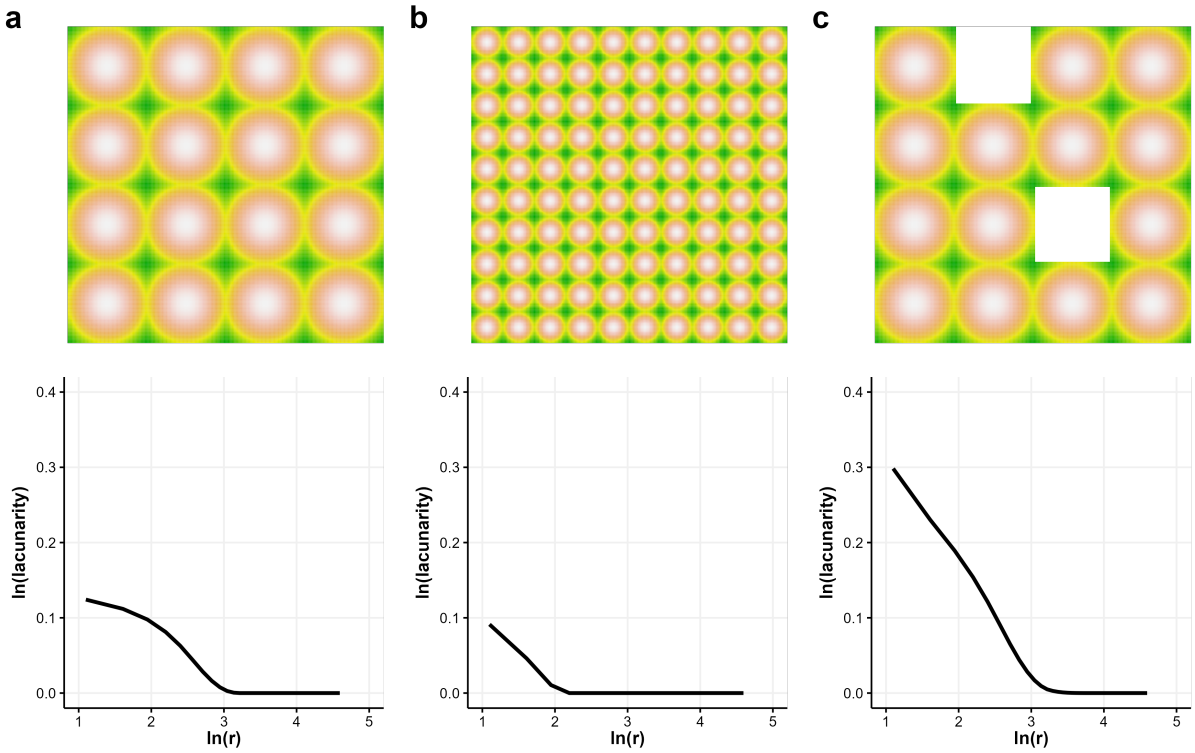


Figure 5 Simulated elevation models of different regularity and variability and their corresponding lacunarity diagrams. All elevation models possess the same extent of 100 $\times$ 100 pixels and values ranging from 0 (green) to 1 (white) (Figure adapted from Fig 3. in Hoechstetter et al., 2011).

3.3.2 Visibility

In this study, the Greenspace Visibility Exposure Index (VGVI) (Labib et al., 2021a) has been implemented to represent eye-level visibility of greenspaces. The VGVI expresses the ratio of visible greenspace to the total visible area an observer can see at a specific location. For this metric, greenspace is represented using the LULC derived binary greenspace layer and visibility is calculated from elevation data using viewshed analysis. A point based viewshed describes a binary, circular raster with a fixed viewing distance (e.g. 300 m) representing visible and non-visible areas. The height of the observer in its centre is derived using ground-level elevation from a DTM and the observer height offset (e.g. 1.8 m). Height within the viewshed is evaluated using a DSM to account for obstacles such as buildings. For each cell within the viewshed visibility is calculated using a Line of Sight (LOS) algorithm. As visualised in Fig. 6, dichotomous visibility (e.g., visible/not visible) along a LOS is evaluated with simple geometry. Information about greenness is extracted from the binary greenspace raster. Finally, all cells within the viewshed are weighted using a distance decay function to account

for the reduced visual prominence of an object in space with increasing distance from the observer (Labib et al., 2021a). The final VGVI expresses the proportion of visible green cells to all visible cells with values ranging from 0 to 1, where 0 = no green cells are visible, and 1 = all of the visible cells are green.

The VGVI has been calculated for the complete study area on a regular point-based grid with 5 m intervals, except for when the point represents buildings or water. The total number of observer locations was 17,329,345 and information about the land use was extracted from the original LULC raster. After computation, the point grid has been aggregated to a continuous raster with 10 m spatial resolution using an Inverse Distance Weighting (IDW) interpolation algorithm. IDW is one of the most frequently used methods in spatial interpolation. Its general idea is that the interpolated value of any location is based on the values of known surrounding locations, assuming closer values are more related than further values. IDW can be calculated as (Eq. 3.10)

$$Z = \frac{\sum_i^n \frac{z_i}{d_i^\beta}}{\sum_i^n \frac{1}{d_i^\beta}} \tag{3.10}$$

where z is the value to be interpolated, n the number of nearest observer locations, and z_i and d_i their corresponding value and distance, respectively. β describes the distance power that determines the degree to which nearer points are preferred over more

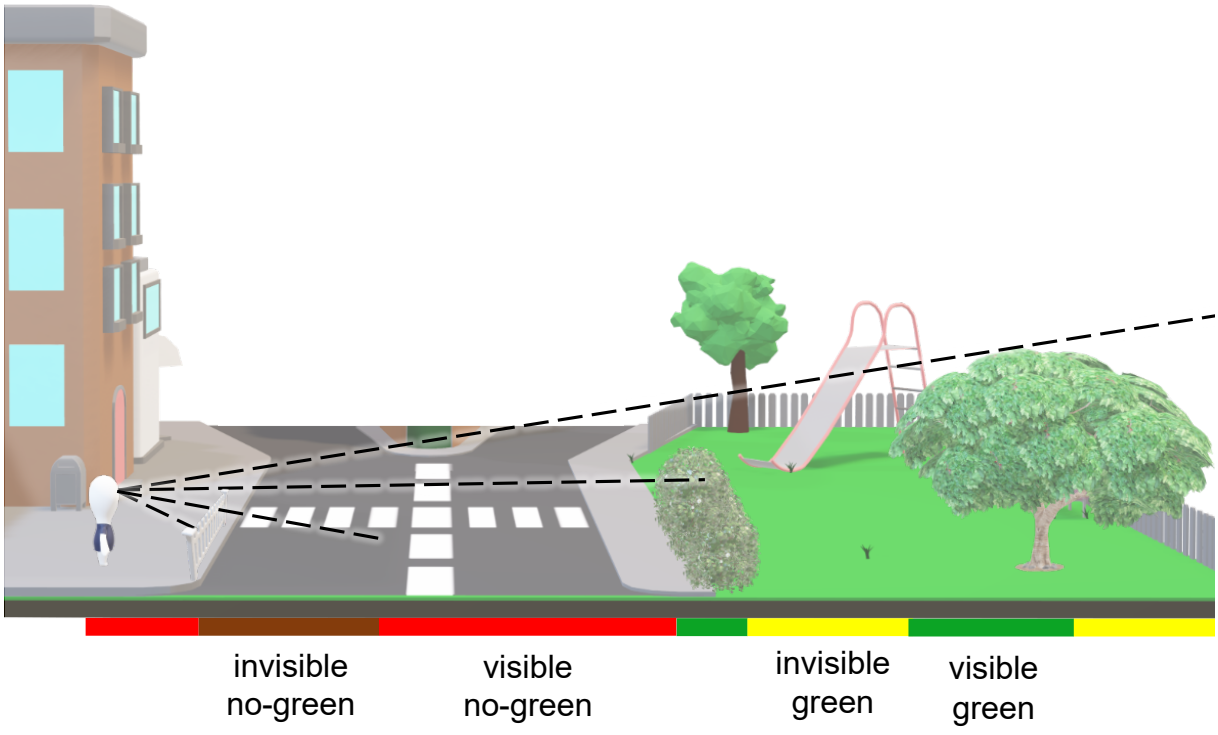


Figure 6 Conceptualisation of the greenness visibility estimation. Dichotomous visibility (e.g. visible, not visible) and intersected greenness values (e.g. green, no-green) are evaluated along the line of sight (Brinkmann et al., 2022).

distant points (Hartmann et al., 2018). To reduce computation time typically a maximum distance threshold D is used so that only the n nearest observer locations with $d \leq D$ are taken into account. VGVI and the IDW interpolation have been computed using the GVI R package (Brinkmann & Labib, 2022). The GVI R package uses an implementation of the viewshed algorithm proposed by Brinkmann et al. (2022). The authors improved computation time of the original VGVI algorithm by implementing a novel prototyping approach making use of shared processing steps between multiple observer locations. Furthermore, the GVI R package supports parallel computation on multiple CPU cores. Parameters for the IDW have been configured through sensitivity analysis with $\beta = 2$, $n = 10$, and $D = 600m$.

As a last step, the VGVI has been reclassified using the Jenks natural breaks classification method, so that extremely low and extremely high levels of greenspace visibility are represented as 1 and 9, respectively, and masked with the study area and a water shapefile.

3.3.3 Accessibility

The Greenspace Accessibility Exposure Index (GACI) (Labib et al., 2021b) has been used to measure the access to public parks. It has been computed following a three-step method: First, access points to public parks have been identified. Second, for the complete study area walking distance has been calculated through network analysis, accounting not only for the distance, but also the size of surrounding parks. Finally, all values have been normalised to generate the GACI map. These steps are explained in the following.

As public parks in urban environments are usually entered through streets or paths, these access points have been identified as the first step. For that, the boundary of the OSM based polygon shapefiles of public parks have been intersected with an OSM road network. The geometric intersections of these layers indicate points, where parks can be accessed through streets and paths. In the case that no street or path intersects with a park boundary, the centroid of a park has been calculated, too. Both point layers have been combined to build the access points (n = 4,559).

To determine the walking distance from a given location to a park, a network analysis has been applied. The network analysis has been conducted using a local instance of the OSRM routing engine (Luxen & Vetter, 2011). OSRM uses a lightweight implementation of the Multi-Level Dijkstra algorithm (Delling et al., 2009) to enable fast computations of large road networks. A regular point-based grid with 10 m intervals has been generated for the complete study area, except for points that were located inside parks (n = 4,839,747). For each location, walking distance to the access points has been calculated. To reduce computation time, only the k nearest access points of a given location have been selected using a k-nearest neighbour algorithm with

$k = \sqrt{4559}$ (Karney 2017). Since not only the distance to a park, but also its size is important for park usage, weights for distance (weight = 1.91) and log-scaled park size (weight = 0.85) have been applied to determine the most attractive park at a given location (Giles-Corti et al., 2005). Walking distance to the most attractive park has been calculated for each location using OSRM.

Finally, the GACI raster map with 10 m spatial resolution has been filled and cells that represent parks have been assigned the lowest walking distance values. To account for extremely high values, all values greater than the 90% decile $Q_{0.9}$ have been assigned to $Q_{0.9}$. As a last step, the GACI has been reclassified using the Jenks natural breaks algorithm, so that 1 represents extremely high, and 9 extremely low values of travel time.

3.3.4 Composite Greenspace Exposure Index

Labib et al. (2021b) suggest combining the three exposure indices as a single Composite Greenspace Exposure Index (CGEI). CGEI has been calculated as (Eq. 3.11):

$$CGEI = \frac{GAVI \times w_{GAVI} + VGVI \times w_{VGVI} + GACI \times w_{GACI}}{w_{GAVI} + w_{VGVI} + w_{GACI}} \tag{3.11}$$

with the greenspace exposure specific weight w. However, no recommendations are given for calculating CGEI in the context of mental health. Therefore, to estimate optimal weights of each exposure metric all combinations of w_{GAVI}, w_{VGVI}, and w_{GACI} were calculated on a regular grid from 0 to 1 within an interval of 0.025. This resulted in a total of 64,000 CGEI combinations. Each CGEI has been entered in a multivariate logistic regression model (Section 3.6; Model 1) and the estimated Odds Ratio (OR), and the three metric specific weights have been stored. To evaluate the combination with the strongest effects on MDE symptoms, I then applied a Bayesian Additive Regression Trees (BART) (Chipman et al., 2010) model with the estimated OR as the dependent variable, and the three weights as independent variables. As demonstrated in previous research (Scarpone et al., 2020), BART has great potential as an exploratory tool when combined with Partial Dependence Plots (PDPs).

BART is an ensemble-of-trees method, such as random forests (Breiman, 2001) and stochastic gradient boosting (Friedman, 2002). Tree-based regression models have an advantage, as they can flexibly fit interactions and non-linearities. A sum-of-trees model - such as BART or random forest - has an even greater ability in understanding these interactions than a single-tree model while also reducing the risk of overfitting. However, BART differs from these two methods, as it uses an underlying Bayesian probability model (Kapelner & Bleich, 2016). One of the advantages of using a Bayesian approach is that it computes Bayesian posterior distributions to approximate the nonparametric model parameters. The priors aim to prevent a single regression from dominating, thus reducing the risk of overfitting (Kapelner & Bleich,

2016; Scarpone et al., 2020). To explore and visualise the marginal effects of each weight on the OR of the CGEI I used PDPs. A PDP is a graphical output that illustrates the marginal effect of an independent variable on the response variable (Friedman, 2002; Scarpone et al., 2020).

Model diagnostics, as well as the R^2 of the BART model have been reported and the final weights have been evaluated from the PDPs by selecting the weight of each metric with the strongest effect on MDE symptoms. Finally, the CGEI has been calculated using Eq. 3.11 with the optimal set of weights.

3.4 Neighbourhood Socioeconomic Status modelling

To account for the effects of neighbourhood SES on mental health, I used a previously developed local SES model. As elaborated in recent analyses (Scarpone et al., 2020; Walker et al., 2019), nearly all previous studies in the literature use either census unit boundaries or simple buffer zones to characterise an individual's local SES (Fuertes et al., 2014; Gong et al., 2014). In a recent study by Walker et al. (2022), the authors presented a distance-weighted, road network-based model for quantifying neighbourhood SES. In order to estimate each participant's potential exposure to local SES, (i) age- and sex-specific walkable zones were mapped around their residential address, and (ii) a negative logit weighting function has been applied, so that the estimated effect of SES decreases as distance from the home increases.

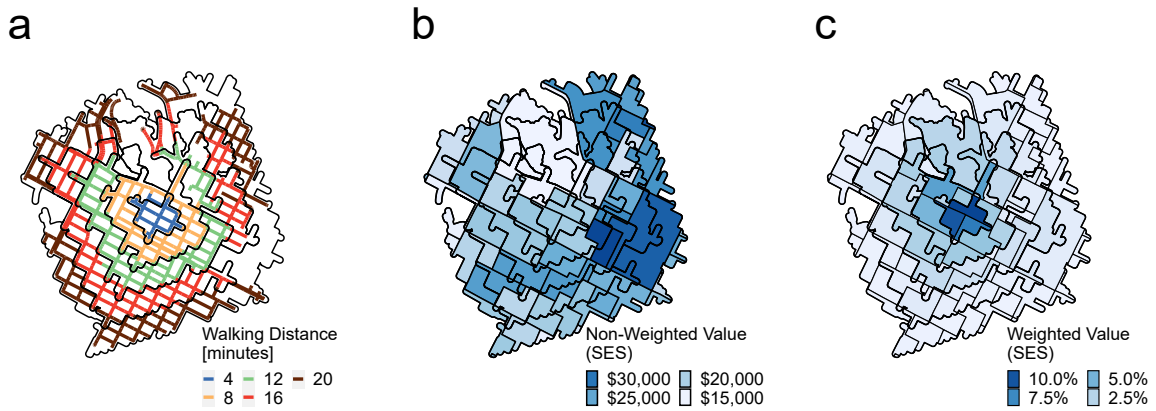


Figure 7 Spatial weighting procedure to account for diminishing effects of distance in which (**a**) road network based isochrones were used, (**b**) a non-weighted variable is mapped over the census dissemination areas and the isochrones, and (**c**) the values are weighted using a distance-decay function (Figure adapted from Fig 1. in Walker et al., 2022).

Using age- and sex-specific walking speeds (average male–female difference = 0.13 km/h; Dewulf et al., 2012) each participant's walkable areas was calculated with a maximum of 20 minutes walking distance, in 2-minute increments. These walkable areas (hereinafter referred to as "isochrones") were computed using the A*-algorithm (Fig. 7a) (Hart et al., 1968). This resulted in each participant having ten concentric

isochrones, the sizes of which are a function of individual walking speed and road network.

In order to account for the diminishing effect of SES as distance increases, a logit function has been applied to weight each incremental isochrone, such that the influence of a variable decreases with increasing distance from the household, i.e., features that are farther away have less influence than nearby features. A logit function was selected as it heuristically approximates a suitable distance-decay function (Bauer & Groneberg, 2016; Jia et al., 2019). The distance weighting of each isochrone has been calculated using (Eq. 3.12):

$$G_t = \begin{cases} \dfrac{\int_0^{r_t} g(r)dr}{\int_0^{r_{tmax}} g(r)dr}, t = 1 \\ \dfrac{\int_{r_{t-1}}^{r_t} g(r)dr}{\int_0^{r_{tmax}} g(r)dr}, t > 1 \end{cases} \tag{3.12}$$

Each isochrone t is assigned a distance weight G_t, calculated as the integral of the logistic distance decay function $g(r)$ (Eq. 3.13)

$$g(r) = \frac{1}{1 + e^{b(r-m)}} \tag{3.13}$$

with $b = 8$ and $m = 0.6$, in the interval between the mean inner radius r_{t-1} and mean outer radius r_t of the isochrone (e.g. 2 to 4 minutes isochrone). This interval has been normalised by the integral from 0 to the outermost isochrone boundary r_{tmax} (i.e. 20 minutes isochrone). As visualised in Fig. 7, an isochrone can cover multiple census areas. Therefore, the proportional weights of the census areas within an isochrone are further defined as (Eq. 3.14)

$$A_{tj} = \frac{A(C_j \cap I_t)}{A(I_t)} \tag{3.14}$$

as the area of the intersection with the census area C_j and the isochrone I_t, divided by the area of the isochrone I_t. The weighted value of the SES variable x_i in the census area j is then defined as (Eq. 3.15):

$$\sum_t \left(G_t \sum_j x_{ij}\ Atj \right) \tag{3.15}$$

Fig. 7b illustrates the unweighted values of a SES variable and Fig. 7c its corresponding proportional, distance adjusted weights. Neighbourhood level SES has been calculated for all participants and for each SES variable listed in Table 2.

3.5 Index Derivation

Previous research often used a single index of socioeconomic deprivation to account for local SES (e.g. Walker et al., 2022). To combine the distance adjusted neighbourhood SES variables I used Principal Components Analysis (PCA) in accordance

with recommendations in the literature (Hajat et al., 2021). PCA is an unsupervised technique for dimensionality reduction in data that focuses on linear combinations of the predictors. The new features, the principal components, attempt to account for as much variation as possible in the original data (Ringnér, 2008). In this study the Vancouver Socioeconomic Deprivation Index (VSDI) has been calculated from the SES variables listed in Table 2. The features were centred and scaled and tested for Measure of Sampling Adequacy (MSA) using the Kaiser-Meyer-Olkin Index (KMO) and for suitability using Bartlett's test. Variables with low KMO were removed, resulting in a final set of 9 SES variables. The predictors were entered into a PCA model using the tidymodels framework (Kuhn & Wickham, 2020). The first principal component was then rescaled to a range of -1 to 1 and used as the VSDI, where low values correspond with socially deprived areas.

3.6 Statistical Modelling and Validation

Using individual-level MDE symptoms as a binary response variable, three sets of logistic regression models were run: (a) bivariate models with the PCA derived VSDI, the three greenspace exposure indices, the CGEI, and all individual level control variables (age, sex, household income range, AHEI score, smoking status and daily alcohol consumption); (b) a fully-adjusted multivariable model using CGEI, VSDI, and all individual level control variables (Model 1); and (c) a semi-adjusted multivariable model using CGEI, VSDI, and a subset of individual level control variables (age, sex, household income range) (Model 2). Furthermore, to analyse whether VSDI is attenuated by individual-level household income, Model 1 was also calculated without household income range.

The data has been split in stratified subsets for training (80%; n = 1,284) and testing (20%; n = 323). The daily alcohol consumption variable contained missing values (n = 240), therefore missing values have been imputed. As the feature contained high outlier values, median imputation has been applied. To account for the class imbalance (i.e., ratio of participants with MDE symptoms < 3 to MDE symptoms ≥ 3), an oversampling approach has been used. When oversampling, the sample size of the minority class is increased by replicating or synthetically generating samples (Hoens & Chawla, 2013). In order to optimise this balance while reducing risk of overfitting, I used the Synthetic Minority Oversampling Technique (SMOTE) algorithm (Liu et al., 2019), which synthetically generates new samples by first identifying k (default $k = 5$) minority-class nearest neighbours of a random minority-class participant. Finally, one of the k neighbours is chosen and a new sample is calculated at a random location along the vector between the two samples (Chawla et al., 2002). The oversampling step has been applied only on the training subset.

All three greenspace indices and the CGEI have been rescaled to a range of -1 to 1, so that effect sizes can be compared to VSDI. OR with 95% Confidence Intervals (CI) and p-values were reported for all models. In-sample statistics (sensitivity, specificity, and accuracy) of the training subset were calculated through 10-fold cross-validation, and compared to the out-of-sample statistics of the testing subset, to test for overfitting and out-of-sample statistics were reported.

4 Results

4.1 Sample Characteristics

Of the total 1,607 participants included in this analysis, 436 (27.1%) had three or more MDE symptoms (Table 2). Participant median age was 52 (IQR 45, 60). Prevalence was higher in women (292 [67.0%] vs 144 [33.0%] in men), and lower at household income of more than CAD $45.000. Alcohol consumption, smoking and healthy diet showed little differences in the study population. Participants with $<$ 3 MDE symptoms showed higher values for all three greenspace exposure metrics, however the differences were small.

Table 2 Summary statistics of the study population.

Variable	**$<$ 3 MDE symptoms (n = 1,171)**	**$\geq$ 3 MDE symptoms (n = 436)**
Baseline characteristics		
Age (years)		
Mean (SD)	52.7 (9.6)	50.7 (8.9)
Median (Q1; Q3)	53.0 (45.0, 60.0)	51.0 (43.0, 57.0)
Sex		
Male	607 (51.8%)	144 (33.0%)
Female	564 (48.2%)	292 (67.0%)
Household income range		
$>$ 90 k	560 (47.8%)	172 (39.4%)
65 k–90 k	244 (20.8%)	82 (18.8%)
45 k–65 k	177 (15.1%)	71 (16.3%)
30 k–45 k	102 (8.7%)	54 (12.4%)
20 k–30 k	50 (4.3%)	28 (6.4%)
$<$ 20 k	38 (3.2%)	29 (6.7%)
AHEI score		
Mean (SD)	40.6 (9.7)	39.6 (9.0)
Median (Q1; Q3)	40.8 (33.7, 46.9)	39.6 (34.0, 45.8)
Current/Former smoker		
No	707 (60.8%)	256 (59.1%)
Yes	455 (39.2%)	177 (40.9%)
Daily alcohol consumption		
Mean (SD)	1.1 (1.5)	1.1 (1.9)
Median (Q1; Q3)	0.6 (0.2, 1.2)	0.5 (0.2, 1.2)
Participants' neighbourhood socioeconomic status		
Lone parent families (%)		
Mean (SD)	14.4 (4.2)	14.4 (4.7)
Median (Q1; Q3)	14.0 (12.1, 17.0)	14.3 (11.6, 17.9)
Private Dwellings - Owned (%)		
Mean (SD)	66.5 (17.2)	64.2 (18.8)
Median (Q1; Q3)	66.3 (55.1, 81.6)	64.9 (50.2, 80.2)
Private Dwellings - Rented (%)		

Continued on next page

Table 2 – continued from previous page

Variable	**$<$ 3 MDE symptoms (n = 1,171)**	**$\geq$ 3 MDE symptoms (n = 436)**
Mean (SD)	32.4 (16.2)	34.1 (17.4)
Median (Q1; Q3)	32.8 (18.2, 42.7)	34.4 (18.8, 46.6)
Labour force participation rate (%)		
Mean (SD)	66.5 (7.0)	66.8 (7.3)
Median (Q1; Q3)	66.6 (62.1, 71.5)	66.9 (62.7, 71.7)
Unemployment rate (%)		
Mean (SD)	5.5 (2.0)	5.4 (1.9)
Median (Q1; Q3)	5.4 (4.0, 6.6)	5.3 (3.9, 6.6)
Individual mean income (CAD/1000)		
Mean (SD)	38.7 (17.0)	37.5 (14.9)
Median (Q1; Q3)	35.0 (30.0, 41.0)	35.0 (30.0, 41.0)
Gov't transfer payments (%)		
Mean (SD)	8.4 (3.5)	8.5 (3.8)
Median (Q1; Q3)	8.1 (5.7, 10.8)	8.0 (5.6, 11.0)
Prevalence of low income (%)		
Mean (SD)	12.2 (6.1)	12.1 (6.0)
Median (Q1; Q3)	11.5 (7.5, 16.0)	11.6 (7.9, 16.3)
Household median income (CAD/1000)		
Mean (SD)	62.2 (17.7)	60.9 (17.0)
Median (Q1; Q3)	59.0 (53.0, 68.0)	58.0 (51.0, 67.0)
Education - No degree (%)		
Mean (SD)	15.2 (5.6)	15.1 (6.1)
Median (Q1; Q3)	15.1 (10.9, 18.8)	15.0 (10.5, 18.8)
Commute Walking/Bicycle (%)		
Mean (SD)	8.4 (7.4)	9.0 (8.3)
Median (Q1; Q3)	5.3 (3.5, 12.3)	5.5 (3.5, 13.8)
Greenspace Exposure metrics		
Availability (GAVI)		
Mean (SD)	4.2 (1.3)	4.1 (1.3)
Median (Q1; Q3)	4.0 (3.0, 5.0)	4.0 (3.0, 5.0)
Visibility (VGVI)		
Mean (SD)	4.8 (2.1)	4.6 (2.1)
Median (Q1; Q3)	5.0 (3.0, 6.0)	5.0 (3.0, 6.0)
Accessibility (GACI)		
Mean (SD)	7.1 (1.3)	7.0 (1.4)
Median (Q1; Q3)	7.0 (6.0, 8.0)	7.0 (6.0, 8.0)

4.2 Greenspace Exposure Metrics

Fig. 8 shows the result of the lacunarity analysis of Vancouver for calculating the GAVI. Results for both sub-regions, Vancouver and Surrey, are reported in Appendix A.1. Fig. 8a displays log-log lacunarity curves for all three greenspace metrics. Lacunarity values were highest for small values in box size, and as the spatial units of the analysis increased, lacunarity approached zero. This implicates that with increasing box size,

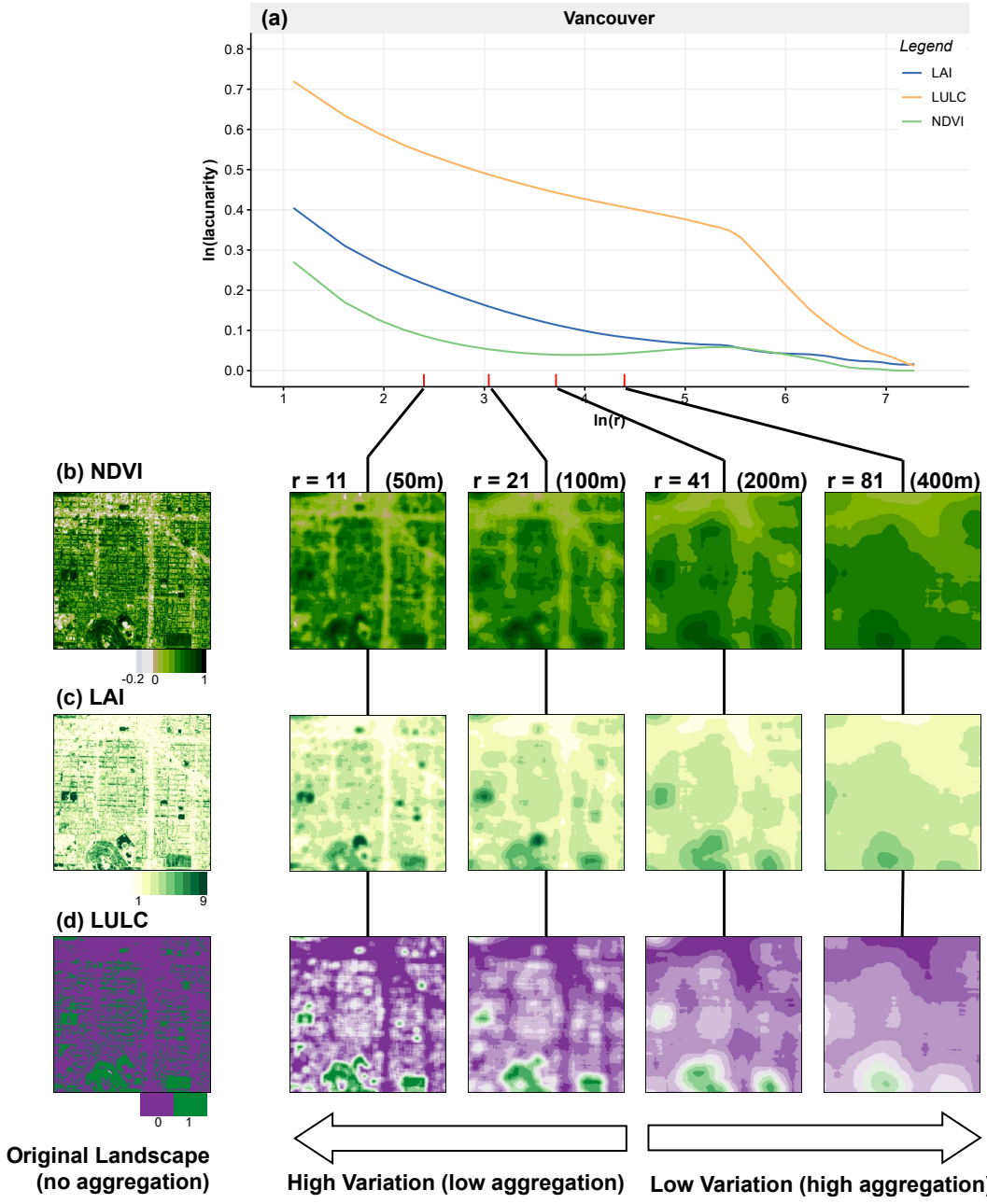


Figure 8 Results of the lacunarity analysis for Vancouver. (**a**) *log-log* lacunarity curves from all three greenspace metrics. For (**b**) NDVI, (**c**) LAI, and (**d**) LULC an example raster has been mapped at original and aggregated scales and compared to their corresponding lacunarity values.

the variance in the resulting greenspace metric decreases. Fig. 8a also shows that the binary LULC raster has the greatest heterogeneity. To better illustrate the changes in lacunarity, the original maps of NDVI (Fig. 8b), LAI (Fig. 8c), and LULC (Fig. 8d) have been compared to their corresponding aggregated maps at different box size values. The figures illustrate more homogeneous maps with lower variance as the values of r increase. The final GAVI map (Fig. A.2a) has been calculated using the scale and metric specific lacunarity values as weights, as described in Section 3.3.1.

Maps of all three greenspace exposure indices are visualised in the Appendix A.2. The visual examination of these indicates that GAVI (Fig. A.2a) and VGVI (Fig. A.2b) have a similar broad spatial pattern. The values are generally lower in densely populated areas and higher in proximity to large bodies of vegetation like parks and agricultural areas. However, VGVI shows greater variance and more details compared to the aggregated GAVI. Furthermore, as buildings and trees block visibility, small differences between the two exposure metrics can be seen, too. The GACI (Fig. A.2c) displays a different spatial pattern, as its values are a function of parks and the road network. GACI values are generally high in densely populated areas, as these promote good

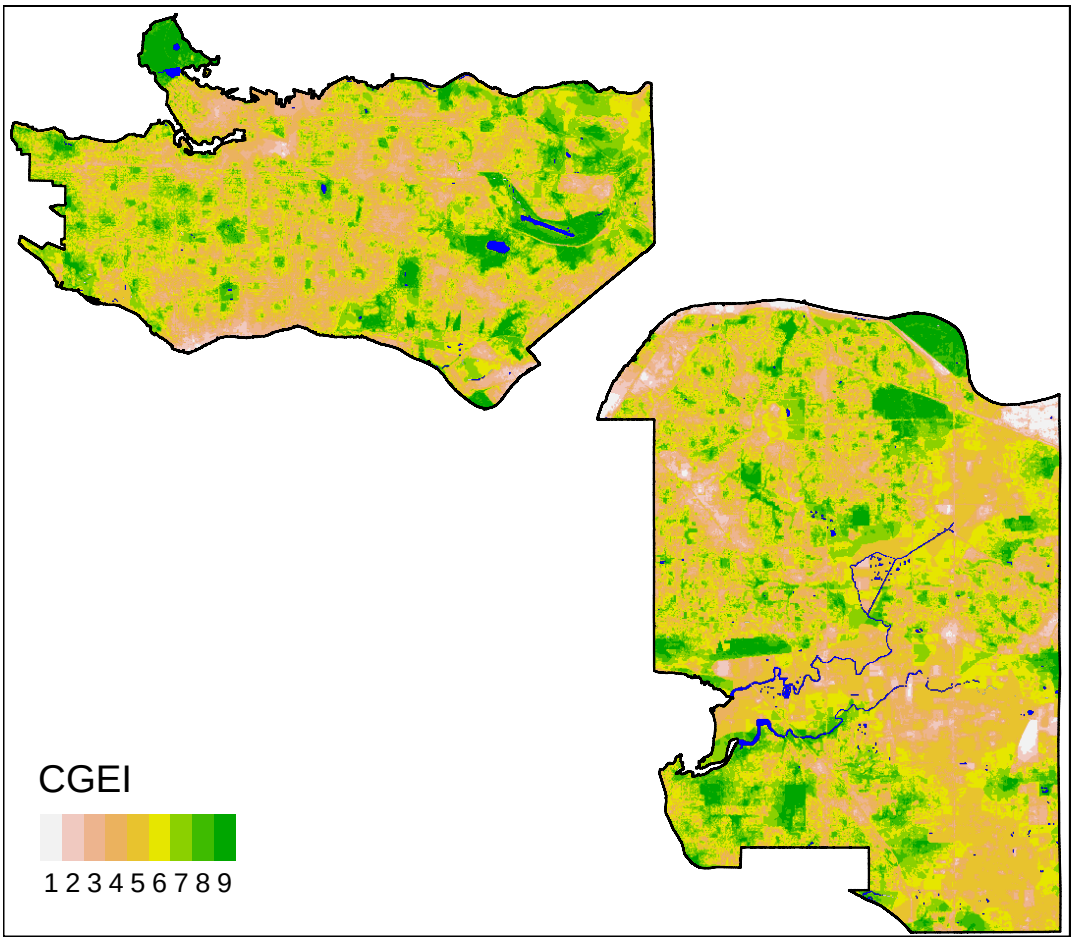


Figure 9 Map of the Composite Greenspace Exposure Index (CGEI). Terrestrial water bodies are visualised in blue.

access to parks. In contrast, the south and east of the study area show large patterns of low access, as these regions are either used for agricultural purposes or connectivity and availability of parks are low.

The final CGEI map has been calculated by combining the GAVI, VGVI and GACI with metric specific weights. The estimated ORs of CGEI on MDE symptoms from all 64,000 possible weight combinations showed a mean (SD) of 0.95 (0.02), ranging from 0.92 to 0.99, where lower ORs indicate a stronger positive effect on MDE symptoms. The BART model of CGEI OR and the three metric specific weights produced an R^2 of 0.99 (p-value < 0.001). PDPs (Fig. A.3) of the three independent variables were visually examined. The optimal set of weights of 0.35, 0.60, and 0.75 has been selected for GAVI, VGVI and GACI, respectively. These absolute weights can be converted to relative weights to report the proportion of each metric on the CGEI. The GACI had the highest contribution to the CGEI with 44%, followed by VGVI (35%) and GAVI (21%). The final CGEI is visualised in Fig. 9.

4.3 Index Derivation

The MSA from the KMO of two SES variables were low with 0.28 and 0.55 for labour force participation rate and private dwellings - rented, respectively. After removing these predictors, the overall MSA was 0.74, and Bartlett's test results confirmed strong non-sphericity ($p < 0.001$). The first principal component explained 40.1% of total variance and was used as the VSDI. The factor loadings are listed in Table 3. Low values of VSDI indicate socially deprived areas and high values describe DAs with high income.

Table 3 Results of the Principal Component Analysis. Factor loadings of the first principal component correspond to variable weights used for calculating the Vancouver Socioeconomic Deprivation Index (VSDI).

Variable	**Factor loadings (PC1)**
Individual mean income	0.33
Household median income	0.26
Commute active	0.24
Private dwellings—owned	0.03
Lone parent families	-0.33
Prevalence of low income	-0.34
Unemployment rate	-0.38
Education—no degree	-0.39
Government transfer payments	-0.48

4.4 Regression Models

The results of the logistic regression models are shown in Table 4 and forest plots are provided in Fig. A.4. When using the three greenspace variables in bivariate models, the crude ORs showed positive effects on MDE symptoms, whereas GACI had the lowest, and GAVI the highest effect size. Combining all greenspace variables in a single index, the CGEI showed a significant and stronger positive association across all models. In the bivariate model, the VSDI showed a non-significant, week negative effect on MDE symptoms. When controlling for the other variables, VSDI exhibited a consistent negative association across both multivariate models. However, the 95% CIs were large, ranging from 0.94 to 1.83 in the fully-adjusted model (Model 1). When not adjusting for individual-level household income, associations between VSDI and MDE symptoms showed lower OR of 1.07 (0.78-1.48, p=0.672). Increased risk of MDE symptoms were measured in women, and for participants with higher daily alcohol consumption, as well as for persons who were smoking. A significant lower risk was observed for higher household income, healthy eating and for increased age. The out-of-sample performance of the fully-adjusted model (Model 1) showed an accuracy of 0.62, sensitivity of 0.61, and specificity of 0.62.

Table 4 Logistic models with Odds Ratio (OR) for Major Depressive Episodes symptoms $\geq$ 3, 95% CI, and p-values. Composite Greenspace Exposure Index (CGEI), Vancouver Socioeconomic Deprivation Index (VSDI), Alternative Healthy Eating Index (AHEI), Greenspace Availability Exposure Index (GAVI), Greenspace Visibility Exposure Index (VGVI), Greenspace Accessibility Exposure Index (GACI).

		Model 1	**Model 2**
Parameter	**OR (bivariate)**	**OR (multivariate)**	**OR (multivariate)**
CGEI	0.92 (0.86-0.98, p=0.010)	0.92 (0.86-0.99, p=0.032)	0.93 (0.87-1.00, p=0.040)
VSDI	1.05 (0.77-1.42, p=0.757)	1.31 (0.94-1.83, p=0.113)	1.29 (0.93-1.79, p=0.135)
Age (5 year-interval)	0.87 (0.83-0.91, p<0.001)	0.83 (0.79-0.87, p<0.001)	0.83 (0.79-0.88, p<0.001)
Sex: female	2.56 (2.15-3.06, p<0.001)	2.71 (2.25-3.27, p<0.001)	2.51 (2.09-3.01, p<0.001)
Household income range	0.87 (0.82-0.92, p<0.001)	0.84 (0.79-0.90, p<0.001)	0.84 (0.79-0.90, p<0.001)
AHEI Score	0.99 (0.98-1.00, p=0.013)	0.99 (0.98-1.00, p=0.004)	
Current/Former smoker: yes	1.05 (0.88-1.24, p=0.618)	1.13 (0.94-1.36, p=0.207)	

Continued on next page

Table 4 – continued from previous page

		Model 1	**Model 2**
Parameter	**OR (bivariate)**	**OR (multivariate)**	**OR (multivariate)**
Alcohol: drinks per day	1.01 (0.95-1.06, p=0.808)	1.06 (1.00-1.13, p=0.044)	
GAVI	0.91 (0.86-0.97, p=0.006)		
VGVI	0.95 (0.92-0.99, p=0.008)		
GACI	0.99 (0.93-1.06, p=0.831)		

5 Discussion

5.1 Evaluation of Results

In this study I calculated the CGEI to estimate the associations between contact to nature and depressive symptoms. The CGEI combines multiple greenspace metrics (i.e. availability, visibility, and accessibility of greenspace) in a single combined exposure index to account for multiple inter-twined pathways to health. As of now, no recommendations about metric specific weights in the context of mental health were provided. Therefore, I adopted an analytical approach to identify the relative importance of availability, visibility, and accessibility exposures in the context of mental health. I used easily interpretable PDPs to estimate optimal weights of each greenspace metric. The results of the analysis indicate that access to parks (GACI) had the strongest contribution to the CGEI with 44%. Visibility of greenspace (VGVI) had the second highest contribution with 35%, followed by overall availability of greenspace (GAVI; 21%). The GACI and VGVI are both relevant when accounting for the restoration (restoring capacities) and instoration (building capacities) pathways, the GAVI plays an important role when accounting for mitigation pathways (reducing harm). The beneficial influences of the restoration pathway on mental health are well understood and intuitive, as these mechanisms have the potential to directly lower stress and restore cognitive function. The instoration pathway describes mechanisms of greenspace with indirect positive effects on mental health through facilitating social cohesion and encouraging physical activity. The mitigation pathway has mixed and week effects on mental health and describes mechanisms that have the potential to reduce harming effects like air pollution, high air temperature and noise through traffic. As GACI and VGVI had the strongest contributions when calculating the CGEI, these findings emphasise that the results are in line with the theoretical background.

Logistic regression models were used to analyse the effects of CGEI on risk of MDE symptoms, while controlling for multiple individual-level variables, and neighbourhood SES. The results indicate that using the CGEI exhibited a more robust model and higher effect sizes compared to the single greenspace exposure metrics. This demonstrates the great potential of a combined greenspace metric to account for the multiple inter-twined pathways from nature to health. Furthermore, in line with the hypothesis that contact to nature is associated with improved health and wellbeing, the results of this study demonstrate a significantly reduced risk of MDE symptoms with increased greenspace. The values of the CGEI range from 1 to 9, where 9 represents extremely high levels of greenspace exposure. The OR for CGEI of 0.92 (0.86-0.99, p=0.032) indicates that increasing the CGEI by 1 reduces the risk of MDE symptoms by 8%. Moreover, an increase of 2 may have effects as strong as increasing the household income range by one category (e.g. increasing the household income range from $30.000 - $45.000 to $45.000 - $65.000).

The SES has been estimated for each participant using a previously developed and tested methodology. The results of the logistic regression models do not fit with the theory that high neighbourhood SES has a positive effect on mental health. A possible explanation might be that the relationship between neighbourhood socioeconomic deprivation and risk of MDE symptoms may be attenuated by individual-level household income (Neally et al., 2022). When not accounting for individual-level household income, the VSDI showed lower negative effects on MDE symptoms and the ORs were similar to the crude OR and showed a large range of 95% CIs. Furthermore, these results might be explained by recent findings that other physical and social neighbourhood environment factors that may relate to depressive symptoms are needed when accounting for neighbourhood SES (Neally et al., 2022).

5.2 Limitations

This study has several limitations. First and foremost, the CGEI has been calculated by combining the three single exposure metrics GAVI, VGVI, and GACI with metric specific weights. Therefore, the interpretation of the OR of CGEI may be non intuitive. However, using the metric specific relative weights may propose an intuitive approach for communicating the relative importance of each single exposure metric for mental health.

Second, this study used an adapted Short-Form Composite International Diagnostic Interview (CIDI-SF) for MDE, consisting of seven questions to assess self-reported MDE symptoms. While previous studies have validated the performance of this questionnaire (Gigantesco & Morosini, 2008), more refined questionnaires might improve the reliability of the results.

Third, the results of this study were constrained by temporal differences between the PURE study and environmental data. In the PURE study participants were recruited between 2006 and 2009, however environmental data was not available for this time period. The LULC raster is based on LiDAR data from 2014, and the NDVI and LAI were calculated from a Sentinel 2 satellite image from 2015. These temporal differences may lead to different measurements of environmental exposures compared to the baseline date. However, it can be expected that the urban environment of Vancouver did not change drastically between 2006 and 2015.

Forth, when estimating access to parks, Giles-Corti et al. (2005) suggest to account for distance, size, and attractiveness of parks. While this study allowed for the first two qualities and therefore improves the estimation of GAVI by Labib et al. (2021b), attractiveness of parks has not been taken into account. While attractiveness of parks is typically estimated through local inspections, future studies might alternatively include a diversity score of remote sensing derived land cover.

Finally, the generalisability of the results is limited by the geographic boundaries of the study, as the analysis focuses on the urban and peri-urban regions of Vancouver, Canada. However, the PURE study is a large-scale epidemiological study in 17 low-, middle-, and high-income countries around the world (Teo et al., 2009). Therefore, future research might expand the demonstrated methodology to investigating associations between multiple greenspace exposures and symptoms of depression for multiple communities and countries around the world.

6 Conclusion

To the authors best knowledge, this is the first study to apply the CGEI on individual-level data. I used this novel index to estimate associations between contact to nature and depressive symptoms. As contact with nature is invariably integrated into a complex web of health determinants, it is clear that multiple pathways are likely to be engaged simultaneously and affect one another. The multiple greenspace exposure framework differs from traditional single exposure metrics in that it can objectively and holistically combine multiple inter-twined greenspace exposure types. As such it is more likely to account for the complexity of multiple pathways. In this study, I combined three exposure metrics, (i) GAVI, (ii) VGVI, and (iii) GACI. However, as of now, no recommendations regarding the metric specific weights to parameterise the CGEI in the context of mental health have been given. Therefore, I provide a methodology for estimating these metric specific weights by using easily interpretable PDPs. The results implicate that access to parks has the greatest contribution on CGEI, followed by greenspace visibility and overall residential greenspace availability. These results support findings that the restoration (restoring capacities) and instoration (building capacities) pathways are important in the context of mental health. Furthermore, logistic regression models were used to investigate associations between GACI and symptoms of depression, while controlling for multiple individual-level variables, and neighbourhood SES. The results of this study demonstrate significantly reduced risk of depressive symptoms with increased greenspace. Furthermore, combining the three exposure metrics in a single index, the CGEI showed a stronger positive association with reduced risk of MDE symptoms. Future research should incorporate the CGEI as it has great potential to account for the complexity of multiple pathways and to further investigate specific effect sizes of the different exposure metrics.

References

Ali, N. A., & Khoja, A. (2019). Growing evidence for the impact of air pollution on depression. *Ochsner Journal*, *19*(1), 4–4. https://doi.org/10.31486/toj.19.0011

Basner, M., Babisch, W., Davis, A., Brink, M., Clark, C., Janssen, S., & Stansfeld, S. (2014). Auditory and non-auditory effects of noise on health. *The Lancet*, *383*(9925), 1325–1332. https://doi.org/10.1016/s0140-6736(13)61613-x

Basu, R. (2009). High ambient temperature and mortality: A review of epidemiologic studies from 2001 to 2008. *Environmental Health*, *8*(1). https://doi.org/10.1186/1476-069x-8-40

Bauer, J., & Groneberg, D. A. (2016). Measuring spatial accessibility of health care providers – introduction of a variable distance decay function within the floating catchment area (FCA) method (K. Deribe, Ed.). *PLOS ONE*, *11*(7), e0159148. https://doi.org/10.1371/journal.pone.0159148

Beyer, K., Kaltenbach, A., Szabo, A., Bogar, S., Nieto, F., & Malecki, K. (2014). Exposure to neighborhood green space and mental health: Evidence from the survey of the health of wisconsin. *International Journal of Environmental Research and Public Health*, *11*(3), 3453–3472. https://doi.org/10.3390/ijerph110303453

Biddle, S. J., Fox, K., & Boutcher, S. (Eds.). (2003). *Physical activity and psychological well-being*. Routledge. https://doi.org/10.4324/9780203468326

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., 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). https://doi.org/10.1126/sciadv.aax0903

Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, *16*(3). https://doi.org/10.1214/ss/1009213726

Brinkmann, S. T. (2021). *Spatlac: Spatial lacunarity*. Zenodo. https://doi.org/10.5281/ZENODO.5786547

Brinkmann, S. T., Kremer, D., & Walker, B. B. (2022). Modelling eye-level visibility of urban green space: Optimising city-wide point-based viewshed computations through prototyping. *AGILE: GIScience Series*, *3*, 1–7. https://doi.org/10.5194/agile-giss-3-27-2022

Brinkmann, S. T., & Labib, S. (2022). *GVI: Greenness visibility index r package*. Zenodo. https://doi.org/10.5281/ZENODO.7057132

Browning, M. H., Lee, K., & Wolf, K. L. (2019). Tree cover shows an inverse relationship with depressive symptoms in elderly residents living in u.s. nursing homes. *Urban Forestry & Urban Greening*, *41*, 23–32. https://doi.org/10.1016/j.ufug.2019.03.002

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953

Chipman, H. A., George, E. I., & McCulloch, R. E. (2010). BART: Bayesian additive regression trees. *The Annals of Applied Statistics*, *4*(1). https://doi.org/10.1214/09-aoas285

Cimburova, Z., & Blumentrath, S. (2022). Viewshed-based modelling of visual exposure to urban greenery – an efficient gis tool for practical planning applications. *Landscape and Urban Planning*, *222*, 104395. https://doi.org/10.1016/j.landurbplan.2022.104395

Comber, A., Harris, P., Bratkova, K., Phe, H. H., Kieu, M., Bui, Q. T., Nguyen, T. T. H., Wanjau, E., & Malleson, N. (2022). Handling the MAUP: Methods for identifying appropriate scales of aggregation based on measures on spatial and non-spatial variance. *AGILE: GIScience Series*, *3*, 1–5. https://doi.org/10.5194/agile-giss-3-30-2022

Dadvand, P., Bartoll, X., Basagaña, X., Dalmau-Bueno, A., Martinez, D., Ambros, A., Cirach, M., Triguero-Mas, M., Gascon, M., Borrell, C., & Nieuwenhuijsen, M. J. (2016). Green spaces and general health: Roles of mental health status, social support, and physical activity. *Environment International*, *91*, 161–167. https://doi.org/10.1016/j.envint.2016.02.029

Dadvand, P., de Nazelle, A., Triguero-Mas, M., Schembari, A., Cirach, M., Amoly, E., Figueras, F., Basagaña, X., Ostro, B., & Nieuwenhuijsen, M. (2012). Surrounding greenness and exposure to air pollution during pregnancy: An analysis of personal monitoring data. *Environmental Health Perspectives*, *120*(9), 1286–1290. https://doi.org/10.1289/ehp.1104609

Dadvand, P., & Nieuwenhuijsen, M. (2019). Green space and health. Springer International Publishing. https://doi.org/10.1007/978-3-319-74983-9_20

Dadvand, P., Nieuwenhuijsen, M. J., Esnaola, M., Forns, J., Basagaña, X., Alvarez-Pedrerol, M., Rivas, I., López-Vicente, M., De Castro Pascual, M., Su, J., Jerrett, M., Querol, X., & Sunyer, J. (2015). Green spaces and cognitive development in primary schoolchildren. *Proceedings of the National Academy of Sciences*, *112*(26), 7937–7942. https://doi.org/10.1073/pnas.1503402112

Davis, Z., Guhn, M., Jarvis, I., Jerrett, M., Nesbitt, L., Oberlander, T., Sbihi, H., Su, J., & van den Bosch, M. (2021). The association between natural environments and childhood mental health and development: A systematic review and assessment of different exposure measurements. *International Journal of Hygiene and Environmental Health*, *235*, 113767. https://doi.org/10.1016/j.ijheh.2021.113767

Delling, D., Sanders, P., Schultes, D., & Wagner, D. (2009). Engineering route planning algorithms. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-02094-0_7

Dewulf, B., Neutens, T., Van Dyck, D., de Bourdeaudhuij, I., & Van de Weghe, N. (2012). Correspondence between objective and perceived walking times to urban destinations: Influence of physical activity, neighbourhood walkability, and socio-demographics. *International Journal of Health Geographics*, *11*(1), 43. https://doi.org/10.1186/1476-072x-11-43

D'Ippoliti, D., Michelozzi, P., Marino, C., de'Donato, F., Menne, B., Katsouyanni, K., Kirchmayer, U., Analitis, A., Medina-Ramón, M., Paldy, A., Atkinson, R., Kovats, S., Bisanti, L., Schneider, A., Lefranc, A., Iñiguez, C., & Perucci, C. A. (2010). The impact of heat waves on mortality in 9 european cities: Results from the EuroHEAT project. *Environmental Health*, *9*(1). https://doi.org/10.1186/1476-069x-9-37

Dong, P. (2000). Lacunarity for spatial heterogeneity measurement in GIS. *Annals of GIS*, *6*(1), 20–26. https://doi.org/10.1080/10824000009480530

Dorman, M. (2022). *Nngeo: K-nearest neighbor join for spatial data* [R package version 0.4.6]. https://CRAN.R-project.org/package=nngeo

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for gmes operational services. *Remote Sensing of Environment*, *120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

Duncan, M., Clarke, N., Birch, S., Tallis, J., Hankey, J., Bryant, E., & Eyre, E. (2014). The effect of green exercise on blood pressure, heart rate and mood state in primary school children. *International Journal of Environmental Research and Public Health*, *11*(4), 3678–3688. https://doi.org/10.3390/ijerph110403678

Dzhambov, A., Hartig, T., Markevych, I., Tilov, B., & Dimitrova, D. (2018). Urban residential greenspace and mental health in youth: Different approaches to testing multiple pathways yield different conclusions. *Environmental Research*, *160*, 47–59. https://doi.org/10.1016/j.envres.2017.09.015

Dzhambov, A. M., Browning, M. H., Markevych, I., Hartig, T., & Lercher, P. (2020). Analytical approaches to testing pathways linking greenspace to health: A scoping review of the empirical literature. *Environmental Research*, *186*, 109613. https://doi.org/10.1016/j.envres.2020.109613

Dzhambov, A. M., Dimitrova, D. D., & Dimitrakova, E. D. (2014). Association between residential greenness and birth weight: Systematic review and meta-analysis. *Urban Forestry & Urban Greening*, *13*(4), 621–629. https://doi.org/10.1016/j.ufug.2014.09.004

Fone, D., White, J., Farewell, D., Kelly, M., John, G., Lloyd, K., Williams, G., & Dunstan, F. (2014). Effect of neighbourhood deprivation and social cohesion on mental health inequality: A multilevel population-based longitudinal study. *Psychological Medicine*, *44*(11), 2449–2460. https://doi.org/10.1017/s0033291713003255

Fotheringham, A. S., & Wong, D. W. S. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A: Economy and Space*, *23*(7), 1025–1044. https://doi.org/10.1068/a231025

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, *38*(4), 367–378. https://doi.org/10.1016/s0167-9473(01)00065-2

Friedrich, M. (2017). Depression is the leading cause of disability around the world. *JAMA*, *317*(15), 1517. https://doi.org/10.1001/jama.2017.3826

Frumkin, H., Bratman, G. N., Breslow, S. J., Cochran, B., Kahn Jr, P. H., Lawler, J. J., Levin, P. S., Tandon, P. S., Varanasi, U., Wolf, K. L., & Wood, S. A. (2017). Nature contact and human health: A research agenda. *Environmental Health Perspectives*, *125*(7), 075001. https://doi.org/10.1289/ehp1663

Fuertes, E., Markevych, I., von Berg, A., Bauer, C.-P., Berdel, D., Koletzko, S., Sugiri, D., & Heinrich, J. (2014). Greenness and allergies: Evidence of differential associations in two areas in germany. *Journal of Epidemiology and Community Health*, *68*(8), 787–790. https://doi.org/10.1136/jech-2014-203903

Gigantesco, A., & Morosini, P. (2008). Development, reliability and factor analysis of a self-administered questionnaire which originates from the world health organization's composite international diagnostic interview - short form (CIDI-SF) for assessing mental disorders. *Clin. Pract. Epidemiol. Ment. Health*, *4*(1), 8.

Giles-Corti, B., Broomhall, M. H., Knuiman, M., Collins, C., Douglas, K., Ng, K., Lange, A., & Donovan, R. J. (2005). Increasing walking. *American Journal of Preventive Medicine*, *28*(2), 169–176. https://doi.org/10.1016/j.amepre.2004.10.018

Giles-Corti, B., Gunn, L., Hooper, P., Boulange, C., Diomedi, B., Pettit, C., & Foster, S. (2018). Built environment and physical activity. Springer International Publishing. https://doi.org/10.1007/978-3-319-74983-9_18

Givoni, B. (1991). Impact of planted areas on urban environmental quality: A review. *Atmospheric Environment. Part B. Urban Atmosphere*, *25*(3), 289–299. https://doi.org/10.1016/0957-1272(91)90001-u

Gong, Y., Gallacher, J., Palmer, S., & Fone, D. (2014). Neighbourhood green space, physical function and participation in physical activities among elderly men: The caerphilly prospective study. *International Journal of Behavioral Nutrition and Physical Activity*, *11*(1), 40. https://doi.org/10.1186/1479-5868-11-40

Grigsby-Toussaint, D. S., Turi, K. N., Krupa, M., Williams, N. J., Pandi-Perumal, S. R., & Jean-Louis, G. (2015). Sleep insufficiency and the natural environment: Results from the us behavioral risk factor surveillance system survey. *Preventive Medicine*, *78*, 78–84. https://doi.org/10.1016/j.ypmed.2015.07.011

Hajat, A., MacLehose, R. F., Rosofsky, A., Walker, K. D., & Clougherty, J. E. (2021). Confounding by socioeconomic status in epidemiological studies of air pollution and health: Challenges and opportunities. *Environmental Health Perspectives*, *129*(6), 065001. https://doi.org/10.1289/ehp7980

Hart, P., Nilsson, N., & Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics*, *4*(2), 100–107. https://doi.org/10.1109/tssc.1968.300136

Hartig, T., & Kahn, P. H. (2016). Living in cities, naturally. *Science*, *352*(6288), 938–940. https://doi.org/10.1126/science.aaf3759

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

Hartmann, K., Krois, J., & Waske, B. (2018). E-learning project soga: Statistics and geospatial data analysis. *Department of Geographiy, University of Kansas Occasional Paper*. https://www.geo.fu-berlin.de/en/v/soga/Geodata-analysis/geostatistics/index.html

Heaviside, C., Macintyre, H., & Vardoulakis, S. (2017). The urban heat island: Implications for health in a changing environment. *Current Environmental Health Reports*, *4*(3), 296–305. https://doi.org/10.1007/s40572-017-0150-3

Hijmans, R. J. (2022). *Terra: Spatial data analysis* [R package version 1.6-3]. https://CRAN.R-project.org/package=terra

Hoechstetter, S., Walz, U., & Thinh, N. X. (2011). Adapting lacunarity techniques for gradient-based analyses of landscape surfaces. *Ecological Complexity*, *8*(3), 229–238. https://doi.org/10.1016/j.ecocom.2011.01.001

Hoens, T. R., & Chawla, N. V. (2013). Imbalanced datasets: From sampling to classifiers. Wiley, John; Sons, Inc. https://doi.org/10.1002/9781118646106.ch3

Hoffimann, E., Barros, H., & Ribeiro, A. (2017). Socioeconomic inequalities in green space quality and accessibility—evidence from a southern european city. *International Journal of Environmental Research and Public Health*, *14*(8), 916. https://doi.org/10.3390/ijerph14080916

Holtan, M. T., Dieterlen, S. L., & Sullivan, W. C. (2014). Social life under cover. *Environment and Behavior*, *47*(5), 502–525. https://doi.org/10.1177/0013916513518064

Houlden, V., Porto de Albuquerque, J., Weich, S., & Jarvis, S. (2019). A spatial analysis of proximate greenspace and mental wellbeing in london. *Applied Geography*, *109*, 102036. https://doi.org/10.1016/j.apgeog.2019.102036

Huynh, Q., Craig, W., Janssen, I., & Pickett, W. (2013). Exposure to public natural space as a protective factor for emotional well-being among young people in canada. *BMC Public Health*, *13*(1). https://doi.org/10.1186/1471-2458-13-407

Jang, H. S., Lee, S. C., Jeon, J. Y., & Kang, J. (2015). Evaluation of road traffic noise abatement by vegetation treatment in a 1:10 urban scale model. *The Journal of the Acoustical Society of America*, *138*(6), 3884–3895. https://doi.org/10.1121/1.4937769

Jenks, G. F. (1977). Optimal data classification for choropleth maps. *Department of Geographiy, University of Kansas Occasional Paper*.

Jia, P., Wang, F., & Xierali, I. M. (2019). Differential effects of distance decay on hospital inpatient visits among subpopulations in florida, usa. *Environmental Monitoring and Assessment*, *191*(S2). https://doi.org/10.1007/s10661-019-7468-2

Kapelner, A., & Bleich, J. (2016). bartMachine: Machine learning with bayesian additive regression trees. *Journal of Statistical Software*, *70*(4). https://doi.org/10.18637/jss.v070.i04

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

Kessler, R. C., & Üstün, T. B. (2004). The world mental health (WMH) survey initiative version of the world health organization (WHO) composite international diagnostic interview (CIDI). *International Journal of Methods in Psychiatric Research*, *13*(2), 93–121. https://doi.org/10.1002/mpr.168

Kingsbury, M., Clayborne, Z., Colman, I., & Kirkbride, J. B. (2019). The protective effect of neighbourhood social cohesion on adolescent mental health following stressful life events. *Psychological Medicine*, *50*(8), 1292–1299. https://doi.org/10.1017/s0033291719001235

Kondo, M. C., Jacoby, S. F., & South, E. C. (2018). Does spending time outdoors reduce stress? a review of real-time stress response to outdoor environments. *Health & Place*, *51*, 136–150. https://doi.org/10.1016/j.healthplace.2018.03.001

Koohsari, M. J., Mavoa, S., Villanueva, K., Sugiyama, T., Badland, H., Kaczynski, A. T., Owen, N., & Giles-Corti, B. (2015). Public open space, physical activity, urban design and public health: Concepts, methods and research agenda. *Health & Place*, *33*, 75–82. https://doi.org/10.1016/j.healthplace.2015.02.009

Kuhn, M., & Wickham, H. (2020). *Tidymodels: A collection of packages for modeling and machine learning using tidyverse principles.* https://www.tidymodels.org

La Placa, V., McNaught, A., & Knight, A. (2013). Discourse on wellbeing in research and practice. *International Journal of Wellbeing*, *3*(1), 116–125. https://doi.org/10.5502/ijw.v3i1.7

Labib, S., Huck, J. J., & Lindley, S. (2021a). Modelling and mapping eye-level greenness visibility exposure using multi-source data at high spatial resolutions. *Science of The Total Environment*, *755*, 143050. https://doi.org/10.1016/j.scitotenv.2020.143050

Labib, S., Lindley, S., & Huck, J. J. (2020a). Spatial dimensions of the influence of urban green-blue spaces on human health: A systematic review. *Environmental Research*, *180*, 108869. https://doi.org/10.1016/j.envres.2019.108869

Labib, S., Lindley, S., & Huck, J. J. (2020b). Scale effects in remotely sensed greenspace metrics and how to mitigate them for environmental health exposure assessment. *Computers, Environment and Urban Systems*, *82*, 101501. https://doi.org/10.1016/j.compenvurbsys.2020.101501

Labib, S., Lindley, S., & Huck, J. J. (2021b). Estimating multiple greenspace exposure types and their associations with neighbourhood premature mortality: A socioecological study. *Science of The Total Environment*, *789*, 147919. https://doi.org/10.1016/j.scitotenv.2021.147919

Lachowycz, K., & Jones, A. P. (2011). Greenspace and obesity: A systematic review of the evidence. *Obesity Reviews*, *12*(5), e183–e189. https://doi.org/10.1111/j.1467-789x.2010.00827.x

Lennard, S. H. C. (2018). Livable cities: Concepts and role in improving health. Springer International Publishing. https://doi.org/10.1007/978-3-319-74983-9_4

Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., & Zhang, W. (2015). Assessing street-level urban greenery using google street view and a modified green view index. *Urban Forestry & Urban Greening*, *14*(3), 675–685. https://doi.org/10.1016/j.ufug.2015.06.006

Lin, B.-S., & Lin, Y.-J. (2010). Cooling effect of shade trees with different characteristics in a subtropical urban park. *HortScience*, *45*(1), 83–86. https://doi.org/10.21273/hortsci.45.1.83

Liu, Q., He, H., Yang, J., Feng, X., Zhao, F., & Lyu, J. (2020). Changes in the global burden of depression from 1990 to 2017: Findings from the global burden of disease study. *Journal of Psychiatric Research*, *126*, 134–140. https://doi.org/10.1016/j.jpsychires.2019.08.002

Liu, S., Lun Ong, M., Kin Mun, K., Yao, J., & Motani, M. (2019). Early prediction of sepsis via SMOTE upsampling and mutual information based downsampling. *2019 Computing in Cardiology Conference (CinC)*. https://doi.org/10.22489/cinc.2019.239

Luxen, D., & Vetter, C. (2011). Real-time routing with openstreetmap data. *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 513–516. https://doi.org/10.1145/2093973.2094062

Maas, J., et al. (2009). *Vitamin g: Green environments-healthy environments*. Nivel.

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., & Gascon, F. (2017). Sen2cor for sentinel-2 (L. Bruzzone, F. Bovolo, & J. A. Benediktsson, Eds.). *Image and Signal Processing for Remote Sensing XXIII*. https://doi.org/10.1117/12.2278218

Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A. M., de Vries, S., Triguero-Mas, M., Brauer, M., Nieuwenhuijsen, M. J., Lupp, G., Richardson, E. A., Astell-Burt, T., Dimitrova, D., Feng, X., Sadeh, M., Standl, M., Heinrich, J., & Fuertes, E. (2017). Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environmental Research*, *158*, 301–317. https://doi.org/10.1016/j.envres.2017.06.028

McCullough, M. L., Feskanich, D., Stampfer, M. J., Giovannucci, E. L., Rimm, E. B., Hu, F. B., Spiegelman, D., Hunter, D. J., Colditz, G. A., & Willett, W. C. (2002).

Diet quality and major chronic disease risk in men and women: Moving toward improved dietary guidance. *The American Journal of Clinical Nutrition*, *76*(6), 1261–1271. https://doi.org/10.1093/ajcn/76.6.1261

McGrath, L. J., Hopkins, W. G., & Hinckson, E. A. (2015). Associations of objectively measured built-environment attributes with youth moderate–vigorous physical activity: A systematic review and meta-analysis. *Sports Medicine*, *45*(6), 841–865. https://doi.org/10.1007/s40279-015-0301-3

Metro Vancouver. (2019). Land cover classification 2014 - 2m lidar (raster) [data retrieved from Metro Vancouver on 02.08.2022, http://www.metrovancouver.org/data].

Milgram, S. (1970). The experience of living in cities. *Science*, *167*(3924), 1461–1468. https://doi.org/10.1126/science.167.3924.1461

Mitchell, R. (2013). Is physical activity in natural environments better for mental health than physical activity in other environments? *Social Science & Medicine*, *91*, 130–134. https://doi.org/10.1016/j.socscimed.2012.04.012

Moore, T., Kesten, J., López-López, J., Ijaz, S., McAleenan, A., Richards, A., Gray, S., Savović, J., & Audrey, S. (2018). The effects of changes to the built environment on the mental health and well-being of adults: Systematic review. *Health & Place*, *53*, 237–257. https://doi.org/10.1016/j.healthplace.2018.07.012

Moreira, T. C. L., Polize, J. L., Brito, M., da Silva Filho, D. F., Chiavegato Filho, A. D. P., Viana, M. C., Andrade, L. H., & Mauad, T. (2021). Assessing the impact of urban environment and green infrastructure on mental health: Results from the são paulo megacity mental health survey. *Journal of Exposure Science & Environmental Epidemiology*, *32*(2), 205–212. https://doi.org/10.1038/s41370-021-00349-x

Natural Resources Canada. (2019). Igh resolution digital elevation model (hrdem) - canelevation serie [data retrieved from Natural Resources Canada on 01.08.2022, https://open.canada.ca/data/en/dataset/957782bf-847c-4644-a757-e383c0057995].

Neally, S. J., Tamura, K., Langerman, S. D., Claudel, S. E., Farmer, N., Vijayakumar, N. P., Curlin, K., Andrews, M. R., Ceasar, J. N., Baumer, Y., & Powell-Wiley, T. M. (2022). Associations between neighborhood socioeconomic deprivation and severity of depression: Data from the national health and nutrition examination survey, 2011–2014. *SSM - Population Health*, *18*, 101111. https://doi.org/https://doi.org/10.1016/j.ssmph.2022.101111

Nieuwenhuijsen, M. J. (2016). Urban and transport planning, environmental exposures and health-new concepts, methods and tools to improve health in cities. *Environmental Health*, *15*(S1). https://doi.org/10.1186/s12940-016-0108-1

Nori-Sarma, A., Sun, S., Sun, Y., Spangler, K. R., Oblath, R., Galea, S., Gradus, J. L., & Wellenius, G. A. (2022). Association between ambient heat and risk of emergency department visits for mental health among us adults, 2010 to 2019. *JAMA Psychiatry*, *79*(4), 341. https://doi.org/10.1001/jamapsychiatry.2021.4369

Nowak, D. J., Hirabayashi, S., Bodine, A., & Greenfield, E. (2014). Tree and forest effects on air quality and human health in the united states. *Environmental Pollution*, *193*, 119–129. https://doi.org/10.1016/j.envpol.2014.05.028

Openshaw, S. (1984). The modifiable areal unit problem. *CATMOG*, *38*.

OpenStreetMap contributors. (2017). Planet dump retrieved from https://planet.osm.org.

Organization, W. H. (1948). Constitution of the who. *Geneva*.

Padgham, M., Rudis, B., Lovelace, R., & Salmon, M. (2017). Osmdata. *The Journal of Open Source Software*, *2*(14). https://doi.org/10.21105/joss.00305

Paluska, S. A., & Schwenk, T. L. (2000). Physical activity and mental health. *Sports Medicine*, *29*(3), 167–180. https://doi.org/10.2165/00007256-200029030-00003

Paoletti, E., Bardelli, T., Giovannini, G., & Pecchioli, L. (2011). Air quality impact of an urban park over time. *Procedia Environmental Sciences*, *4*, 10–16. https://doi.org/10.1016/j.proenv.2011.03.002

Penkalla, A. M., & Kohler, S. (2014). Urbanicity and mental health in europe: A systematic review. *European Journal of Mental Health*, *9*(2), 163–177. https://doi.org/10.5708/ejmh.9.2014.2.2

Phelan, P. E., Kaloush, K., Miner, M., Golden, J., Phelan, B., Silva, H., & Taylor, R. A. (2015). Urban heat island: Mechanisms, implications, and possible remedies. *Annual Review of Environment and Resources*, *40*(1), 285–307. https://doi.org/10.1146/annurev-environ-102014-021155

Plotnick, R. E., Gardner, R. H., & O'Neill, R. V. (1993). Lacunarity indices as measures of landscape texture. *Landscape Ecology*, *8*(3), 201–211. https://doi.org/10.1007/bf00125351

Pretty, J., Peacock, J., Sellens, M., & Griffin, M. (2005). The mental and physical health outcomes of green exercise. *International Journal of Environmental Health Research*, *15*(5), 319–337. https://doi.org/10.1080/09603120500155963

Rautio, N., Filatova, S., Lehtiniemi, H., & Miettunen, J. (2017). Living environment and its relationship to depressive mood: A systematic review. *International Journal of Social Psychiatry*, *64*(1), 92–103. https://doi.org/10.1177/0020764017744582

Ringnér, M. (2008). What is principal component analysis? *Nature Biotechnology*, *26*(3), 303–304. https://doi.org/10.1038/nbt0308-303

Scarpone, C., Brinkmann, S. T., Große, T., Sonnenwald, D., Fuchs, M., & Walker, B. B. (2020). A multimethod approach for county-scale geospatial analysis of emerging infectious diseases: A cross-sectional case study of COVID-19 incidence in germany. *International Journal of Health Geographics*, *19*(1). https://doi.org/10.1186/s12942-020-00225-1

Schüle, S. A., Gabriel, K. M., & Bolte, G. (2017). Relationship between neighbourhood socioeconomic position and neighbourhood public green space availability: An environmental inequality analysis in a large german city applying generalized linear models. *International Journal of Hygiene and Environmental Health*, *220*(4), 711–718. https://doi.org/10.1016/j.ijheh.2017.02.006

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. https://doi.org/10.1002/fee.1225

Su, J. G., Dadvand, P., Nieuwenhuijsen, M. J., Bartoll, X., & Jerrett, M. (2019). Associations of green space metrics with health and behavior outcomes at different buffer sizes and remote sensing sensor resolutions. *Environment International*, *126*, 162–170. https://doi.org/10.1016/j.envint.2019.02.008

Su, J. G., Jerrett, M., de Nazelle, A., & Wolch, J. (2011). Does exposure to air pollution in urban parks have socioeconomic, racial or ethnic gradients? *Environmental Research*, *111*(3), 319–328. https://doi.org/10.1016/j.envres.2011.01.002

Tabrizian, P., Baran, P. K., Van Berkel, D., Mitasova, H., & Meentemeyer, R. (2020). Modeling restorative potential of urban environments by coupling viewscape

analysis of lidar data with experiments in immersive virtual environments. *Landscape and Urban Planning*, *195*, 103704. https://doi.org/10.1016/j.landurbplan.2019.103704

Taylor, L., & Hochuli, D. F. (2017). Defining greenspace: Multiple uses across multiple disciplines. *Landscape and Urban Planning*, *158*, 25–38. https://doi.org/10.1016/j.landurbplan.2016.09.024

Teo, K., Chow, C. K., Vaz, M., Rangarajan, S., & Yusuf, S. (2009). The prospective urban rural epidemiology (PURE) study: Examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries. *American Heart Journal*, *158*(1), 1–7.e1. https://doi.org/10.1016/j.ahj.2009.04.019

Turner, W. R., Nakamura, T., & Dinetti, M. (2004). Global urbanization and the separation of humans from nature. *BioScience*, *54*(6), 585. https://doi.org/10.1641/0006-3568(2004)054[0585:guatso]2.0.co;2

Ulrich, R. S. (1981). Natural versus urban scenes. *Environment and Behavior*, *13*(5), 523–556. https://doi.org/10.1177/0013916581135001

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology*, *11*(3), 201–230. https://doi.org/10.1016/s0272-4944(05)80184-7

United Nations. (2018). The world's cities in 2018—data booklet. *Department of Economic and Social Affairs, Population Division*. https://population.un.org/wup/Publications/

Van Berkel, D. B., Tabrizian, P., Dorning, M. A., Smart, L., Newcomb, D., Mehaffey, M., Neale, A., & Meentemeyer, R. K. (2018). Quantifying the visual-sensory landscape qualities that contribute to cultural ecosystem services using social media and lidar. *Ecosystem Services*, *31*, 326–335. https://doi.org/10.1016/j.ecoser.2018.03.022

Van Renterghem, T., Forssén, J., Attenborough, K., Jean, P., Defrance, J., Hornikx, M., & Kang, J. (2015). Using natural means to reduce surface transport noise during propagation outdoors. *Applied Acoustics*, *92*, 86–101. https://doi.org/10.1016/j.apacoust.2015.01.004

Voogt, J., & Oke, T. (2003). Thermal remote sensing of urban climates. *Remote Sensing of Environment*, *86*(3), 370–384. https://doi.org/10.1016/s0034-4257(03)00079-8

Walker, B. B., Brinkmann, S. T., Große, T., Kremer, D., Schuurman, N., Hystad, P., Rangarajan, S., Teo, K., Yusuf, S., & Lear, S. A. (2022). Neighborhood greenspace and socioeconomic risk are associated with diabetes risk at the sub-neighborhood scale: Results from the prospective urban and rural epidemiology (PURE) study. *Journal of Urban Health*, *99*(3), 506–518. https://doi.org/10.1007/s11524-022-00630-w

Walker, B. B., Shashank, A., Gasevic, D., Schuurman, N., Poirier, P., Teo, K., Rangarajan, S., Yusuf, S., & Lear, S. A. (2019). The local food environment and obesity: Evidence from three cities. *Obesity*, *28*(1), 40–45. https://doi.org/10.1002/oby.22614

Weinstein, N., Balmford, A., DeHaan, C. R., Gladwell, V., Bradbury, R. B., & Amano, T. (2015). Seeing community for the trees: The links among contact with natu-

ral environments, community cohesion, and crime. *BioScience*, *65*(12), 1141–1153. https://doi.org/10.1093/biosci/biv151

Weiss, M., Baret, F., & Jay, S. (2020). *S2toolbox level 2 products lai, fapar, fcover* (Research Report). EMMAH-CAPTE, INRAe Avignon. https://hal.inrae.fr/hal-03584016

White, M. P., Alcock, I., Wheeler, B. W., & Depledge, M. H. (2013). Would you be happier living in a greener urban area? a fixed-effects analysis of panel data. *Psychological Science*, *24*(6), 920–928. https://doi.org/10.1177/0956797612464659

White, M. P., Elliott, L. R., Grellier, J., Economou, T., Bell, S., Bratman, G. N., Cirach, M., Gascon, M., Lima, M. L., Lõhmus, M., Nieuwenhuijsen, M., Ojala, A., Roiko, A., Schultz, P. W., van den Bosch, M., & Fleming, L. E. (2021). Associations between green/blue spaces and mental health across 18 countries. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-87675-0

White, M. P., Pahl, S., Wheeler, B. W., Depledge, M. H., & Fleming, L. E. (2017). Natural environments and subjective wellbeing: Different types of exposure are associated with different aspects of wellbeing. *Health & Place*, *45*, 77–84. https://doi.org/10.1016/j.healthplace.2017.03.008

Williams, A. J., Maguire, K., Morrissey, K., Taylor, T., & Wyatt, K. (2020). Social cohesion, mental wellbeing and health-related quality of life among a cohort of social housing residents in cornwall: A cross sectional study. *BMC Public Health*, *20*(1). https://doi.org/10.1186/s12889-020-09078-6

Xie, Y., Xiang, H., Di, N., Mao, Z., Hou, J., Liu, X., Huo, W., Yang, B., Dong, G., Wang, C., Chen, G., Guo, Y., & Li, S. (2020). Association between residential greenness and sleep quality in chinese rural population. *Environment International*, *145*, 106100. https://doi.org/10.1016/j.envint.2020.106100

Zhang, R., Zhang, C.-Q., & Rhodes, R. E. (2021). The pathways linking objectively-measured greenspace exposure and mental health: A systematic review of observational studies. *Environmental Research*, *198*, 111233. https://doi.org/10.1016/j.envres.2021.111233

Appendices

A Figures

A.1 Lacunarity Analysis

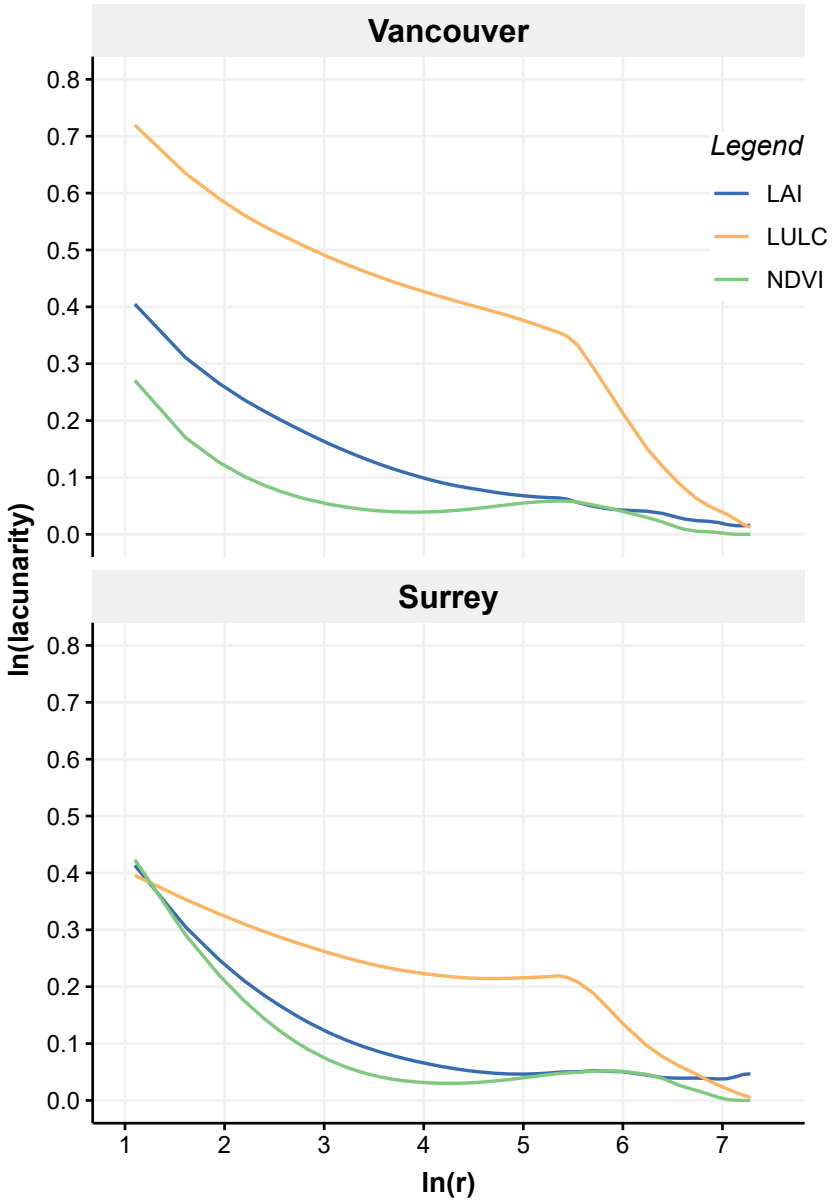


Figure A.1 Lacunarity diagrams for both sub-regions of the study area - Vancouver in the north-west, and Surrey in the south-east.

Results of the lacunarity analysis from Section 3.3.1 for both sub-regions, Vancouver and Surrey. Fig. A.1 displays log-log lacunarity curves for all three greenspace metrics. The broad pattern is similar in both regions and lacunarity values were highest

for LULC. However, Surrey generally showed lower values indicating lower variance, which might be explained by more homogeneous large-scale patterns (e.g. agricultural areas in the south-east) and a suburban character compared to the urban centre of Vancouver.

A.2 Greenspace Exposure Metrics

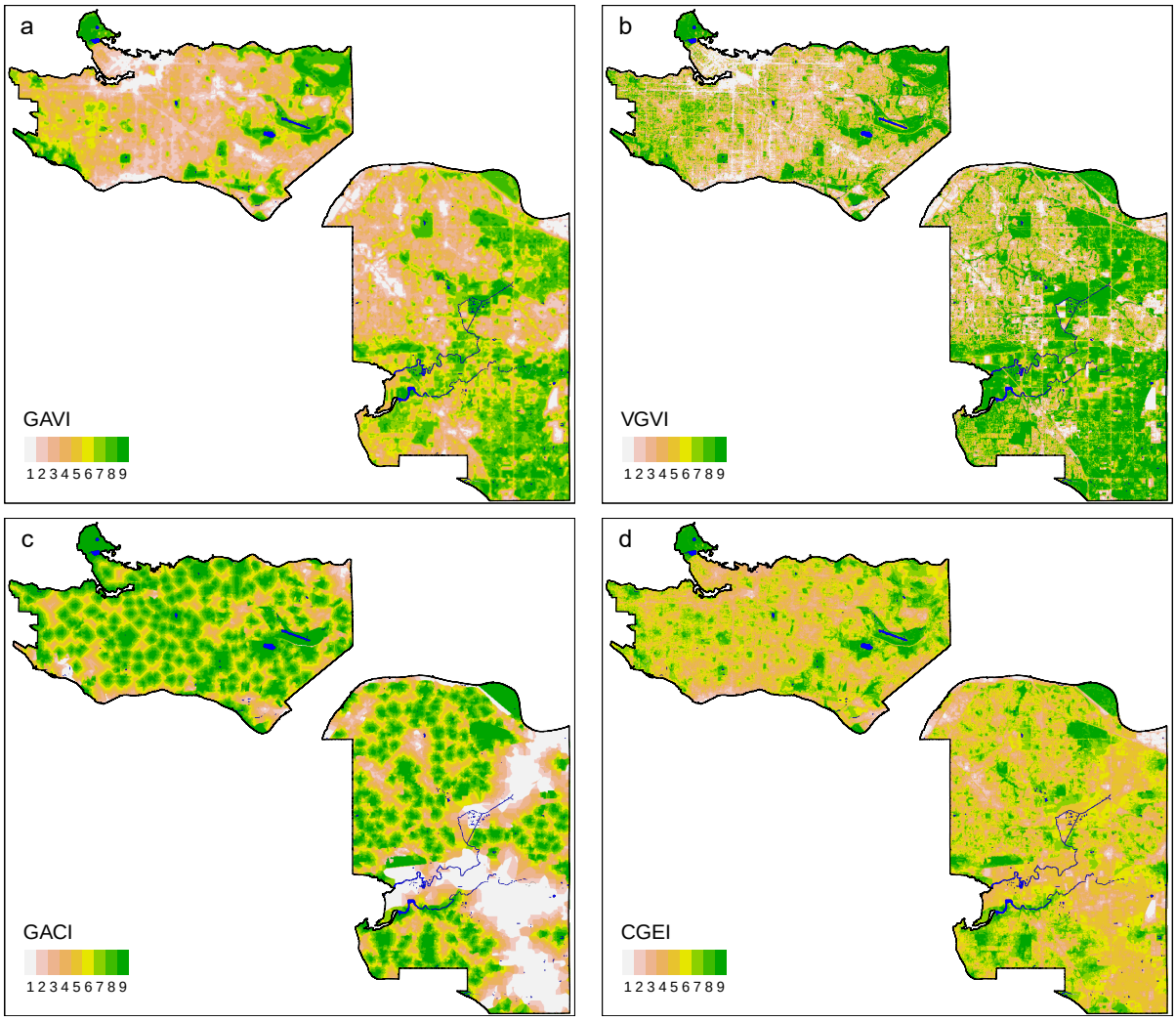


Figure A.2 Maps of the greenspace (**a**) availability, (**b**) visibility, and (**c**) accessibility exposure, and the Composite Greenspace Exposure Index (CGEI) (**d**). Terrestrial water bodies are visualised in blue.

A.3 Partial Dependence Plots

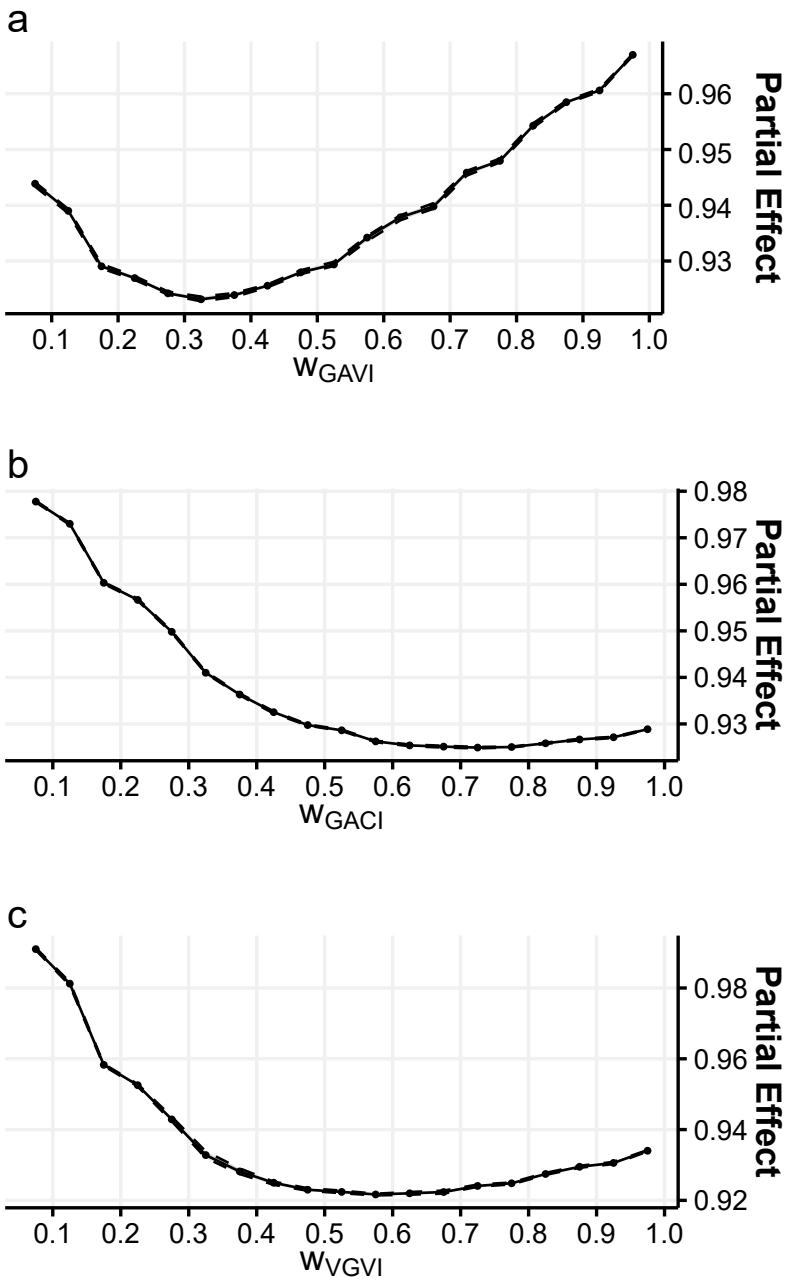


Figure A.3 Partial Dependence Plots (PDPs) of the three metric specific weights (**a**) GAVI, (**b**) GACI, and (**c**) GACI. The x-axis shows the value of the weight variable and the y-axis its partial effect on the estimated MDE symptoms OR. Lower values of partial effect indicate a stronger positive effect on MDE symptoms.

A.4 Regression Models

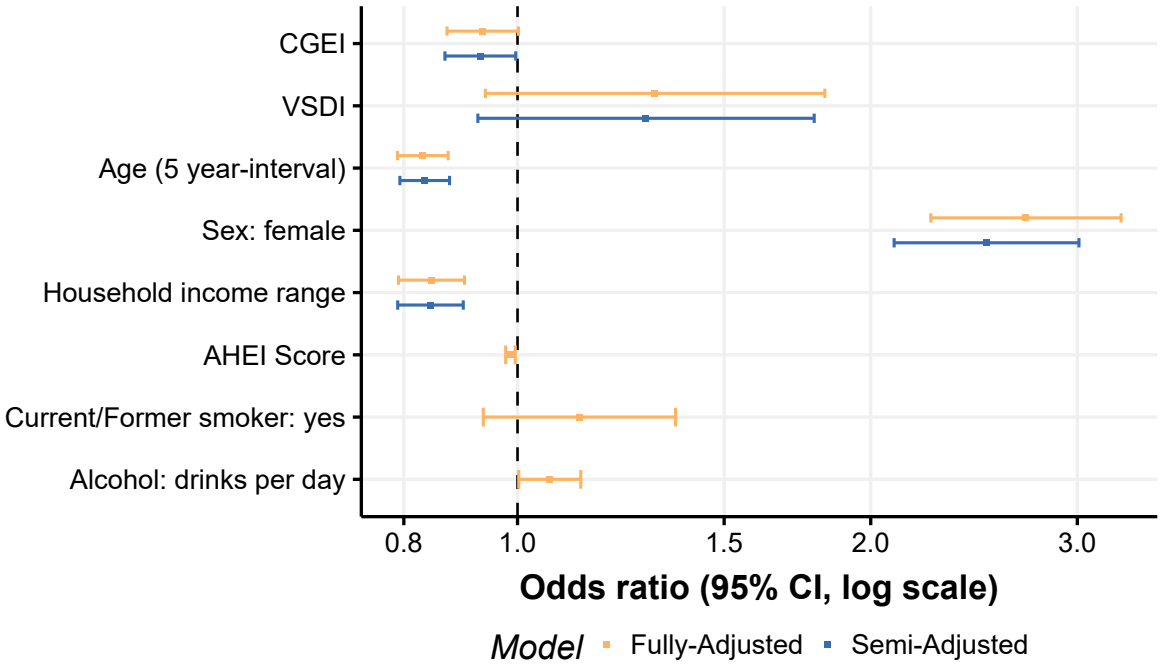
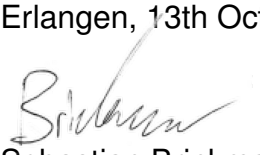


Figure A.4 Forest plots showing significant effects for both semi-adjusted and fully adjusted multivariable logistic models. Composite Greenspace Exposure Index (CGEI), Vancouver Socioeconomic Deprivation Index (VSDI), Alternative Healthy Eating Index (AHEI).

Declaration of Academic Integrity

Hiermit erkläre ich, dass die vorliegende Bachelorarbeit von mir selbstständig verfasst wurde und dass keine anderen als die angegebenen Hilfsmittel benutzt wurden. Die Stellen der Arbeit, die anderen Werken dem Wortlaut oder Sinn nach entnommen sind, sind in jedem einzelnen Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht. Diese Erklärung erstreckt sich auch auf etwa in der Arbeit enthaltene Zeichnungen, Karten(skizzen) und bildliche Darstellungen.

Erlangen, 13th October 2022



Sebastian Brinkmann