Relationship of perceived emotional response to the soundscape and urban green space based on a deep learning approach

by

Xiaohao Yang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Master of Landscape Architecture
Master of Science / Geospatial Data Sciences
(Environment and Sustainability)
in the University of Michigan
2022

Thesis Committee:

Professor Mark Lindquist, Chair
Professor Derek Van Berkel

Acknowledgments

I would like to thank those people who have provided me with support and inspiration. And this thesis could be completed to this stage is the effort of everyone in the lab.

My deepest gratitude goes foremost to my advisors Dr. Mark Lindquist and Dr. Derek Van Berkel. The thesis could not have been possible without the summer research internship in the lab that allowed me to do research about my favorite subject: perceived emotion from the soundscape. They've had more faith in me than I've had myself, and I cannot fully express my gratitude for their constant encouragement, accurate comments, and guidance. The idea for this thesis was shaped by their discussions with me. They have led me through all the processes of writing this thesis.

Secondly, I would like to thank Ph.D. student David Grace and the postdoc Dr.Nathan Fox in the lab for their precious assistance and review. The writing of this thesis would not reach this level without their priceless comments.

Lastly, I greatly appreciate the psychological support and accompany of my parents and girlfriend. It was their kindness and concern that encouraged me to overcome the difficulty through graduate study and the hardest moment of my life.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

While urban greenspace is widely recognized as important for human health and well-being, research on this topic of urgent importance is regularly not scaled for landscape planning and design applications. Here we propose a spatially explicit, deep-learning-based method to assess auditory stimuli within an important area of healthcare supply and we predict and map emotional responses to its soundscapes. Decomposing soundscape emotion by its dimensions of pleasantness and eventfulness, we find that both pleasantness and eventfulness are significant correlated with greenspace. Pleasantness is positively associated with greenspace whereas eventfulness is negatively associated with greenspace. The direction of emotional response to urban greenspace comports with current understandings of restorative landscapes. Our findings indicates that restorative soundscapes for hospitals may be insufficient due to potential impacts of urban contexts. Our spatially explicit framework helps to inform understandings of landscape restorativeness at the landscape planning scale and can be replicated to help ensure landscape design reaches its restorative potential, particularly in critical urban applications where populations and public health needs are increasingly concentrated.

Chapter 1 Introduction

Urban environments can have numerous negative health and well-being consequences owing to a variety of issues including inadequate access to greenspace (Wolch et al. 2014), increased proximity to urban pollutants (Hanna-Attisha, 2016; Duzgorin-Aydin, 2007), and the impact of urban noise (Stansfeld et al. 2010). Moreover, due to the population density of cities, public health impacts can be amplified in urban environments as the number of impacted persons harmed by local pollutants increases with population density in affected areas (WHO, 2010). Urban greenspaces contribute to mental and emotional well-being and are increasingly recognized for their importance as places where non-pharmaceutical therapy can occur (Hunter et al. 2019). Urban designs with greenspace likely ameliorate some of the negative health issues of urban environments and benefit residents' physical and psychological health by limiting exposure to pollutants (Lei et al. 2021), increasing the potential for relaxation and exercise (Richardson et al. 2013), and reducing urban noise or possibly altering subjective responses to noise (Koprowska et al. 2018). For example, while the restorative visual benefits of nearby greenspace might be reduced by urban noise, the distribution of greenspace might attenuate negative effects generated from the urban context. Visual exposure to natural landscapes has been found to have many psychological benefits including reduced stress (Hunter et al., 2019), stress recovery (Zhu et al., 2020), attention restoration (Kaplan & Kaplan, 1989), as well as improved cognitive and emotional states (Kellert and Wilson, 1995, Rosley et al., 2014, Ulrich, 1993). Other senses, particularly auditory perception of the environment, can also provide significant psychological benefits for mental wellness and restoration (Ratcliffe, 2021).

Soundscape - the “acoustic environment as perceived or experienced and/or understood by a person or people, in context” (ISO 12913-1 2014) - is an important factor contributing to perceptual and psychological responses to the environment. While there has been increased focus over the past two decades on the study of the soundscape as the previously neglected auditory dimensions of landscape perception, the visual qualities of the landscape are still most commonly researched. Recent advances in analyzing perception of urban greenspace as a multisensory phenomenon (Hedblom et al. 2019, Lindquist et al. 2020) is a positive direction for evaluating the impact of visual and auditory stimuli on restorativeness. Yet, research across heterogenous environmental contexts is needed to clarify which findings linking soundscape to wellbeing permit generalization and to identify the extent soundscape health impacts are a function of spatial relationships, environmental quality, landscape composition, and culture. To encourage further actionable knowledge creation in this domain, we apply machine learning tools to multisensory soundscape analysis at a landscape planning and design relevant scale. This approach has substantial policy relevance for identifying critical urban locations and characteristics that would benefit urban populations and contribute to public health broadly.

Previous soundscape research has decomposed the complex dimensions of soundscape perception into orthogonal axes of pleasantness and eventfulness, which were identified as the first two principal components explaining 50% and 18% of soundscape response variance respectively in a PCA (Axelsson et al., 2010); the theoretical underpinnings for the naming of these axes derives from Russel’s (1980) circumplex model of affect whose principal axes have subsequently been described as valence and arousal (Erfanian et al. 2021). The space modeled by the soundscape dimensions of pleasantness and eventfulness can be used to characterize subjective emotional responses to soundscape dimensions (Fig 1).

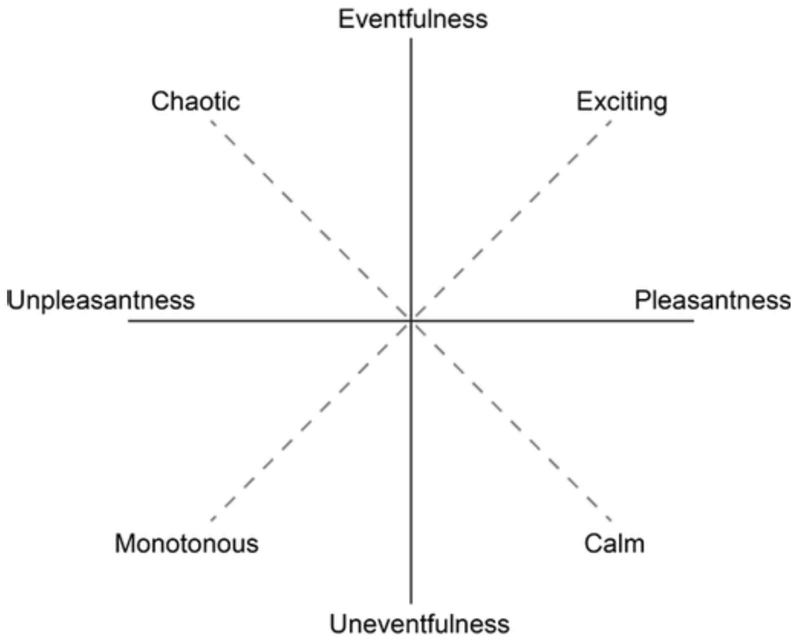


Figure 1.1: Soundscape emotion dimension (adapted from Axelsson et al. 2010, fig. 4)

The pleasantness-unpleasantness dimension is a measure of the degree of pleasantness appraised in the acoustic environment (Russell et al., 1989). The eventfulness-uneventfulness dimension describes the richness of events in the soundscape and represents the perceived activity intensity of the stimulus. The principal axes of pleasantness and eventfulness delineate quadrants that suggest emotional states: excitement, calm, chaos, and monotony. The lines through the origin bisecting the quadrants suggest a continuum of the magnitude within a qualitative emotional state. Likewise, distance away from the 45-degree line in a circular rotation at a given magnitude suggests shift along a continuum between emotional states.

Research characterizing the relationship between soundscape dimensions and affective, the emotional response has suggested positive and negative emotional responses to particular acoustic aspects of the environment (Ulrich et al., 1991; Yang et al., 2005; Irvine et al., 2009). A key finding of this research is that natural sounds - such as bird sounds and water - are often perceived to be pleasant whereas non-natural urban noise - such as traffic - is perceived to be unpleasant (Ricciardi et al., 2015; Shan Shu and Hui Ma, 2020). Yet, perception of soundscape quality can be highly subjective with differences between demographic groups, such as the finding that women and older participants found natural sounds more calming than other participants (Hedblom et al., 2017). However, other evidence suggests demographic differences within one city may explain considerably less than that attributed to differences in location (Erfanian et al. 2021). Since the mix of natural and non-natural sounds demographic groups encounter in urban greenspace is spatially dependent, spatially explicit methods can help create actionable knowledge for planning and design applications that consider environmental and population distribution characteristics at the same time. Earlier findings suggest this line of research has relevance to understanding the components of human wellbeing. For instance, natural or positively perceived sounds can contribute to improved mental health (Medvedev et al., 2015). However, limits to generalizing existing findings given ex-situ and non-spatially explicit study designs should be noted. Further, scalable methods are necessary to replicate and extend findings to landscape planning and design scales.

The choice to study the soundscape impacts using an in-situ or ex-situ approach has unique limitations. Aletta et al. (2019) compare results from the two in-situ soundscape data collection protocols in Annex C of the ISO/TS 12913-2:2018, but do not compare with the ex-situ protocol in that annex. In-situ surveys evaluating the site-specific features of landscape appreciation have been commonly used (e.g. Yang and Kang, 2005) but can be challenging due to the difficulty and time necessary to recruit a large group of people to annotate sound source collections. Ex-situ studies have been based on small samples of the population in experimental survey designs (Hong, 2016, Hedblom et al. 2019; Lindquist et al. 2020). Scaling such information for landscape planning/design purposes requires methods that can relate people's perception to spatial attributes of the landscape from site to urban scale (Collins et al. 2020). However, these methods are time and labor-intensive and have restricted soundscape studies from being scaled to larger spatial extents. Promising work on general models of human responses to sound qualities has increased over the last decade relating acoustic environments to acoustic signals features. For example, Eigenfeldy and Bizzocchi (2015) used surveys of sound appreciation and analysis of sound spectrograms to develop models that are likely to bring broader context to sound quality at scale.

Addressing the limits of scale, space, heterogeneity, and multiple sense perception in soundscape research will be key to urban greenspace planning and design. Natural sounds in cities are usually supported by greenspace, particularly design supporting the ecological quality of urban green space and mediating mechanical and human noise (Irvine et al., 2009). By extension, landscape quality could mediate emotional response to urban sound perception. Liu et al. (2013) found that visual landscape has significant effects on the perception of soundscapes, especially natural sounds. But far less is known about how the distribution of green space is related to sound emotional responses on dimensions of pleasantness and eventfulness. To our knowledge, existing studies on soundscape perception prediction did not use a validated dataset of the ground truth data with affective annotations, which is important to large-scale soundscape mapping.

Chapter 2 Literature Review

2.1 Soundscape Emotion

Research on soundscapes has increased in popularity over the past three decades and researchers continue to develop approaches to study soundscape characteristics. One particular area of research focus is soundscape perception, including research exploring how human emotion is affected by the acoustic environment. The emotional response to soundscape has been primarily measured by self-reported methods (i.e., verbal and non-verbal report) (Bradley et al., 1994) and physiological reaction (i.e., skin conductance, respiration rate, and heart rate) (Hume et al. 2008 and 2013).

More generally there has been proposed the concept of emotion dimensions (Wundt, 1906), and since then many researchers have developed the theory of semantic dimensions that is the foundation of verbal and non-verbal approaches. Mehrabian and Russell (1974) found the main independent factors pleasure, arousal, and dominance that can be used to explain the variance in emotional meaning in all situations. Based on this work, Russell (1980) built a circumplex model of affect, proposing that affective states in all situations are combinations of varying degrees of valence and arousal. Mehrabian and Russell’s work was a starting point, and soundscape researchers adopted the circumplex model of basic emotions. For example, Axelsson et al. (2010) proposed a Pleasantness–Eventfulness model for soundscapes that capture the circumplex concept of Russell. Later, ISO TS 12913:3 (2019) provides a similar soundscape model of pleasantness and eventfulness. Since emotions are psychological states, conscious emotions can be studied. Ciuk et al. (2018) concluded that emotion self-reports and physiological reactions converge to some extent. Even if emotions can be different according to behavioral and physiological responses and biased estimates of emotions can be caused by individual experience or preference, self report is still essential for studying perceived soundscape emotions (Fiebig et al.,2020).

Although the variables of general dimensions vary from study to study, pleasantness and arousal have been commonly investigated in both soundscape and multisensory studies in terms of noise exposure (Jiang and Kang, 2017), urban sound mapping (Kang et al., 2018), and automatic prediction of soundscape emotion (Thorogood and Pasquier, 2013, Lopes et al., 2017). Table 1 shows the papers that used pleasantness and arousal or similar soundscape descriptors as emotion dimensions. Both verbal and non-verbal report approaches based on the two-dimension system (Pleasantness–Eventfulness model) have been frequently used to collect rated data from a set of audio stimuli in the field of studying human emotional reactions to acoustic environments (Fiebig et al.,2020).

Different from the in-situ method including physiological laboratory experiments (Annerstedt et al., 2013, Hedblom et al., 2019) and on-site experiments, the ex-situ self-report method can be conducted easier without temporal limitations. Previous researchers have frequently used this method to get perception ratings. For instance, to study socio-cultural differences in soundscape perception, Serebrennikova et al. (2019) collected soundscape evaluations by requiring participants to rate sound sources from different counties using a set of semantic sound attributes. In the last decade, researchers have utilized ex-situ self-report with different media to generate large datasets for soundscape prediction. For example, Giannakopoulos et al. (2019) built a soundscape perception dataset by allowing users to annotate the perceived quality of the recorded soundscape using a smartphone application. Besides, Fan et al. (2017) has released an annotated soundscape emotion dataset created by a valence-arousal based dimensional approach for the purpose of automatic soundscape affect recognition (Fan et al., 2016).

Author	Emotion dimensions	Applied method	Study object/purpose
Bradley et al., 2000	Valence*, arousal**	Non-verbal report	Emotional reactions to natural sounds
Irwin et al., 2011	Pleasantness*, vibrancy**	Verbal report	Emotional responses to urban soundscapes
Hall et al., 2013	Pleasantness*, vibrancy**, calmness, comfort, informational content, intrusiveness	Verbal report	Perceptual and spectro-temporal properties of urban soundscapes
Fan et al., 2016	Pleasantness*, eventfulness**	Verbal report	Predicting model for valence and arousal of soundscapes
Stevens et al., 2016	Valence*, arousal**, and dominance	Non-verbal report	Relationship between psychological and physiological responses to soundscapes
Stevens et al., 2017	Valence*, arousal**	Non-verbal report	Relationship between soundscape categorization and subjective evaluation.
Zhang and Kang, 2020	Enjoyment*, excitement**, desolation, tension, familiarity	Verbal report	Emotions constituting the perceptions and feelings of urban soundscapes
Erfanian et al., 2021	Pleasantness*, eventfulness**	Verbal report	Influence of psychological well-being and demographic factors on sound perception
Frescura and Lee, 2022	Valence*, arousal**	Non-verbal report	Emotions and physiological responses elicited by neighbor sound

Table 2.1: The dimensions are marked with an asterisk if they are similar to Mehrabian and Russell's pleasure (*) and arousal (**) dimensions.

2.2 Health Benefit of Soundscape And Landscape

Soundscape, including urban noise and natural sounds, has either positive or negative effects on health and quality of life (Brown et al., 2016). A previous review on positive health-related effects and soundscapes found that the associations between positive soundscape perceptual constructs and health benefits were statistically significant (Aletta et al., 2018). Medvedev et al. (2015) argued that acoustic environments can impact people both psychologically and physiologically. Exposure to pleasant natural sounds can make stress recovery faster than less pleasant noise (Alvarsson et al., 2010). However long-term exposure to traffic noise can lead to many health problems including cardiovascular diseases, cognitive impairment, sleep disturbance, hypertension and annoyance, potentially leading to premature death (World Health Organization, 2018). It is the negative impact of urban sounds that makes the health benefits of soundscapes get more and more attention in research. The evidence for restorative outcomes of acoustic experiences of nature has been well studied and documented by researchers through subjective and objective measurements (Goel and Etwaroo, 2006, Jahncke et al., 2011, Gould van Praag et al., 2017).

The visual landscape, especially in nature, also has been proved to be beneficial to improving restoration-related effects by evoking positive perceptions such as aesthetic appeal, fascination, and sense of safety. (Ulrich, 1986, Ulrich et al.,1991). The visible green vegetation has been confirmed to be able to reduce noise annoyance (Van Renterghem and Botteldooren, 2016). The percentage of green in street view provides measures of visible green space for health research (Larkin and Hystad, 2019). Recently, the positive effect of exposure to eye-level urban greenness or street view green on the mental health of elderly people has been studied using deep learning for image segmentation. More recently, Wang and Zhao (2020) found significant effects of evergreen plants on psychological restoration in spring. Tabrizian et al. (2020) predicted the urban park landscape's restorative potential based on studying the relationship between self-rate restorativeness and spatial structures and features of landscapes. However, the landscape's restorative benefits have not been spatially assessed in the urban-scale acoustic environment.

The soundscape and visual landscape have been considered as a total environment to be studied (Cassidy, 2013). The effects of audio-visual interactions on soundscape assessment have been well documented in past decades (Liu et al., 2014, Liu et al. 2019, Li and Lau, 2020). At the same time, there are investigations on the perception of the visual landscape influenced by audio stimuli (Anderson et al., 1983, Hetherington et al., 1993, Benfield et al., 2010), whereas many have focused on the whole audio-visual environment (e.g., only audio-visual condition) to study the perception of soundscape or landscape or overall perception (Sanchez et al., 2017, Pheasant and Watts, 2015, Zhao et al., 2018). The health effects of soundscape and landscape can be interrelated. By conducting two comparative experiments on visual stimuli and auditory-visual stimuli, Li Deng, et al. (2020) found that nature-related audio-visual stimuli had significant restorative attributes and better health potential than visual stimuli alone. However, the relationship between soundscape perception and landscape has not been well uncovered in spatial research.

Since Attention Restoration Theory was proposed by Kaplan and Kaplan (1989), previous studies of environmental perception made efforts to quantitatively evaluate the restorative effects of landscape and soundscape leveraging various methods and found plenty of evidence to support the health benefits of soundscape and landscape. Researchers have used a variety of measures to assess health-related attributes of soundscape and landscape. Some of these measures are interchangeable with those for evaluating soundscape emotions. Table 2.2 shows the papers that used emotion-related measures in health-related soundscape and landscape research. In conclusion, the health-related effects of acoustic and visual environments are significantly associated, and the measures of perceived emotion have been widely used to assess health benefits. What’s more, the interactions between human and audio-visual environments and the meanings of the environmental experiences have been intensively studied. To move forward, the study in the field of soundscape benefit may be worth focusing on uncovering the pathway, in which different environmental settings can support psychological well-being (Ratcliffe, 2021). There is a need to understand the spatial association between perceived emotion from soundscapes and landscapes.

Author	Measures	Study object/purpose
Alvarsson et al., 2010	Pleasantness, Eventfulness, Familiarity	Nature sounds facilitate recovery after a psychological stressor.
Hume and Ahtamad, 2013	Pleasantness, Arousal	Pleasant soundscapes facilitate faster recovery from stress compared to unpleasant soundscapes.
Medvedev et al., 2015	Pleasantness, Eventfulness, Familiarity, Arousal, Dominance	Experience of unpleasant soundscapes at rest produces greater stress than pleasant soundscapes.
Maehr et al., 2015	Valence, Arousal	The emotional response to the visual impact on the landscape

Table 2.2: Papers that used emotion-related measures in health-related soundscape and landscape research

2.3 Soundscape Mapping and Machine Learning

The first attempts of representing soundscapes using cartography can be traced back to the book Pure geography (Granö, 1997) by Johannes Gabriel Granö (Radicchi, 2013). The spatial study of soundscape is meaningful to the understanding of the distribution and variation of acoustic attributes and relationships between acoustic perception and urban space structures. Job et al. (2016) mapped sound pressure levels in three terrestrial habitats to understand how habitats vary acoustically over space and time using locally recorded sounds.

For urban planning and policy decision making, mapping impacts or attributes of soundscape in the urban environment are important to the goal of achieving noise management (de Kluijver and Stoter, 2003, Murphy and King, 2014), land use management (Keyel et al., 2017), health city (Radicchi and Grant, 2021, Schulte-Fortkamp, 2021), and environmental equity (Engel et al., 2019). For example, Wang and Kang (2011) simulated traffic noise based on the transportation systems, land use, and building environment to understand the effects of urban morphology on noise distribution. The findings of these studies would finally guide or inspire the optimization of urban configurations for the improvement of health benefits (Weber, 2014). With the development of spatial analysis techniques and data resources, soundscape researchers have been established frameworks for the spatial visualization and analysis of soundscapes based on the self-report data with geo-locations or data-driven approaches. For instance, Hong and Jeon (2014) mapped self-report soundscape perceptions in different urban contexts using GIS techniques.

In recent two decades, researchers have taken advantage of machine learning and computation algorithms to effectively map soundscape attributes and experiences. For developing predictive soundscape models, Aletta et al. (2016) proposed a conceptual framework that contains three steps: collecting sound data; characterizing the acoustic environment; creating a model that relates the soundscape perception to the physical properties of the acoustic environment. Mennitt et al. (2013) built geospatial sound models leveraging the random forest algorithm to predict the sound pressure level across the contiguous United States. They trained the regressional model that related the acoustic metrics from a set of natural and anthropogenic sound sources to geospatial data including biogeophysical, climatic, and anthropogenic variables. The sounds in the training set were locally recoded from 244 geographically unique locations. With the models, they spatially visualized the impacts of anthropogenic noise and activity on natural areas and protected lands. Interestingly, with the growth of crowdsourced data, social media has also been used as a dataset for sound mapping at city scale. For example, Aiello et al. (2016) mapped the perceived emotion of soundscape on streets based on tagging information of georeferenced photos from Flickr and sound-related words from Freesound. With tagging information from social media, Aiello et al. classified sound types from sound-related words on georeferenced social media content so that they classified the types of streets (i.e., human, indoor, music, mechanical, etc.). Similarly, potential semantic emotions (i.e., joy, trust, fear, surprise, etc.) were computed from tags. And they studied the associations between street types and emotions to match the potential emotion to each type of street. They conducted soundwalks on streets across areas in Europe to have participants classify sound (street) types and rate their emotional experiences (ie. chaotic, monotonous, calm, and exciting) from that. Based on the relationship between sound types and people's acoustic experience, Aiello et al. finally mapped the chaotic, monotonous, calm, and exciting streets. Although social media makes it possible to capture soundscapes at large scale, the framework of the transformation from social media to perceived sound emotion is still complicated. On the other hand, the method highly depends on the density and amount of a social media dataset. Therefore, it would be challenging to apply the method to those non-tourism cities, which don't have such dense social media data.

Most recently, neural networks (deep learning method) have been widely used in sound features recognition, sound classification, and sound perception assessment. Interestingly, the deep learning method and machine learning method can be used together for sound perception mapping. For example, Verma et al. (2020) trained custom deep learning models using locally collected data from streets to predict audio-visual variables from sound clips and street views. They created datasets of subjective audio-visual experience annotated by experts through images and sound clips. Finally, spatiotemporal visual and auditory perception maps were created based on the random forest algorithm model that related subjective audio-visual perception to audio-visual attributes. Through the comparison of the regression model and artificial neural network, Brocolini et al. (2012) pointed out that the predictive point of view in the two approaches is similar but the advantage of the artificial neural network is that it may highlight the relative influence of each variable on the perception for a specific location.

As a subset of machine learning, deep learning is essentially a neural network with three or more convolutional layers. Generally, there are two parts in a dataset for training networks: images and ground truth labels. The ground truth label corresponds to each image and the labels could be a series of class names or metric numbers. In the forward stage, the input image is transformed into a new representation using filtering and thresholding operations (weights and bias) and the ground truth label and prediction output are computed in the loss function to get the loss cost. In the backward stage, the gradient of each weight is computed based on the loss cost. Then, based on the gradients, the initial weights are updated and are used for the next iteration until the iterations are sufficient. There are several keys to making a neural network a 'black box'. The weights can be changed depending on loss functions, the weights initialized at the start of training, the order in which the images are shown to the network, and training iterations (Del Campo et al., 2021).

The deep learning approach has been proven to be promising and beneficial to large-scale sound mapping research due to the establishment of public ground-truth datasets. Verma et al. (2019) leveraged AudioSet dataset (Gemmeke et al., 2017) to train a deep learning model to classify sound sources that were manually collected from local streets. They compared the sound sources classified by models with that classified by experts and found the result of prediction was compatible with that of experts. Based on the classified sound recordings, they mapped the noise level of different types of sound including human-based, biotic, and anthropogenic sounds using spatial interpolation techniques of GIS. Multiple datasets of overhead images and geo-tagged sounds have been trained to develop sound predictive models for mapping purposes. Salem et al. (2018) predicted semantic sound type at a certain location using overhead images based on a neural network that related sounds to overhead images at the same locations. They finally mapped the distribution of soundscapes using the hierarchical clustering method to cluster predicted sounds. Table 2.3 shows the papers that adopted environmental sound datasets to sound mapping.

The machine learning method has reduced the cost of the spatial research of soundscape and made the process automatic. Currently, most researchers have been mapped the semantic soundscapes based on automatic sound classification. Machine learning especially the deep learning approach is still rarely applied to quantitatively map sound emotions.

Author	Sound Dataset	Predictions	Method
Mennitt et al., 2013	Locally recorded sound sources	Sound pressure level	Random forest algorithm
Salem et al., 2018	Freesound	Sematic emotion	Deep learning
Verma et al., 2019	AudioSet	Sematic emotion	Deep learning
Verma et al., 2020	Locally recorded sound sources	Sematic emotion	Deep learning

Table 2.3: Papers that adopted environmental sound datasets to sound mapping

Chapter 3 Methods

This study aims to explore spatial relationships between green space and sound emotion responses on an urban scale leveraging deep learning to model emotion responses and LiDAR data to model urban green space. A public soundscape dataset, Emo-Soundscapes, is used to train the emotion prediction model. Pleasantness and eventfulness are used as soundscape emotion descriptors[1]. We analyze an area around a major hospital system in Ann Arbor, Michigan to evaluate possible health and well-being benefits of soundscape for healthcare workers, patients, and other visitors. Our assessment used 200 sound recordings randomly sampled within a 15-minute walking distance of hospitals. Spearman's rank correlation is used to examine the relationships between green space and sound emotional responses.

3.1 Study Area

For this study we examine the city of Ann Arbor, Michigan, USA, focusing on areas surrounding the University of Michigan hospital system. The University of Michigan Hospital system was founded in 1848 and is now a large employer and healthcare provider. As highlighted during the COVID-19 pandemic, healthcare workers navigate highly stressful situations on a daily basis (Greenberg, 2020) and access to nature is being recognized as a critical component of health (Naomi, 2020). As of 2021, the U-M hospital system employs more than 24,417 people, has 1,107 licensed beds, and comprises 5 main hospitals and centers (U-M health). In addition to its 94,785 emergency visitors in 2021, the hospital provides long-term care

[1] The Fan et al. (2017) study uses valence and arousal as equivalent descriptors for pleasantness and eventfulness respectively (p.198-9).

such as through the C.S. Mott Children's Hospital and the Rogel Cancer Center. In 2021 there were 2,645,178 patient clinic visits who were affected by the hospital soundscape, as well as its healthcare providers (Michigan Medicine, 2021). Following recognition of nature's benefits to both healthy people and those suffering illnesses (White et al. 2019), including better job and life satisfaction for health care providers (Irvine and Warber, 2002) and particularly the health benefits of group walks in urban greenspaces (Marselle et al. 2013, 2015), our study focuses on the restorative role of soundscape within accessible walks outside the hospitals; we randomly sampled 200 locations along accessible paths which we defined as a 15-minute walk from building entrances based on sidewalk locations. The samples cover gradients of urban impervious and natural landscapes ranging from city core (i.e., The downtown area) to major urban parks (i.e., river walkways, the Arboretum) suggesting differences in soundscape due to landscape characteristics (Fig. 3.1, Fig. 3.2).

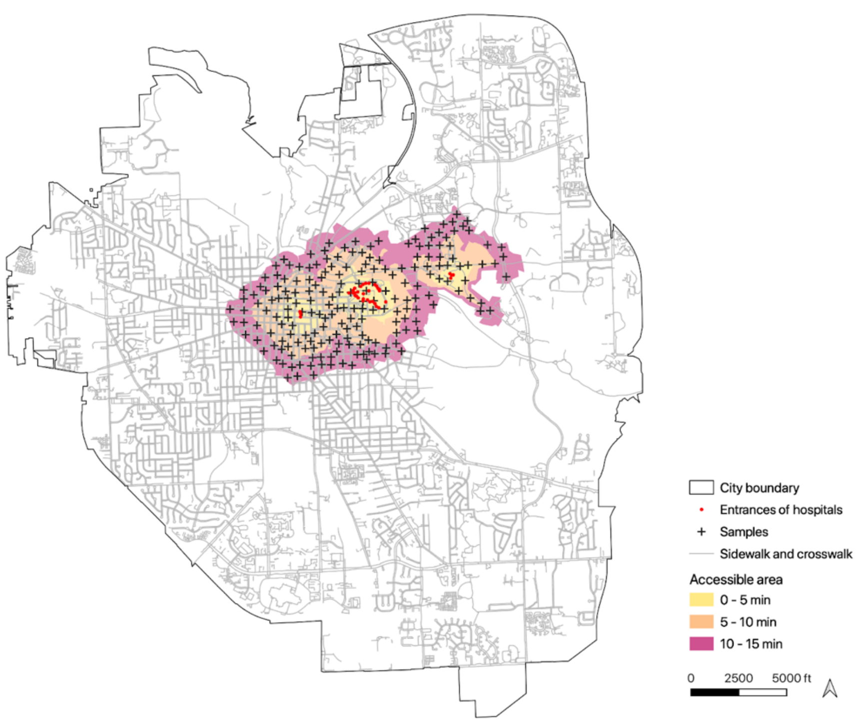


Figure 3.1: Study area and the locations of samples

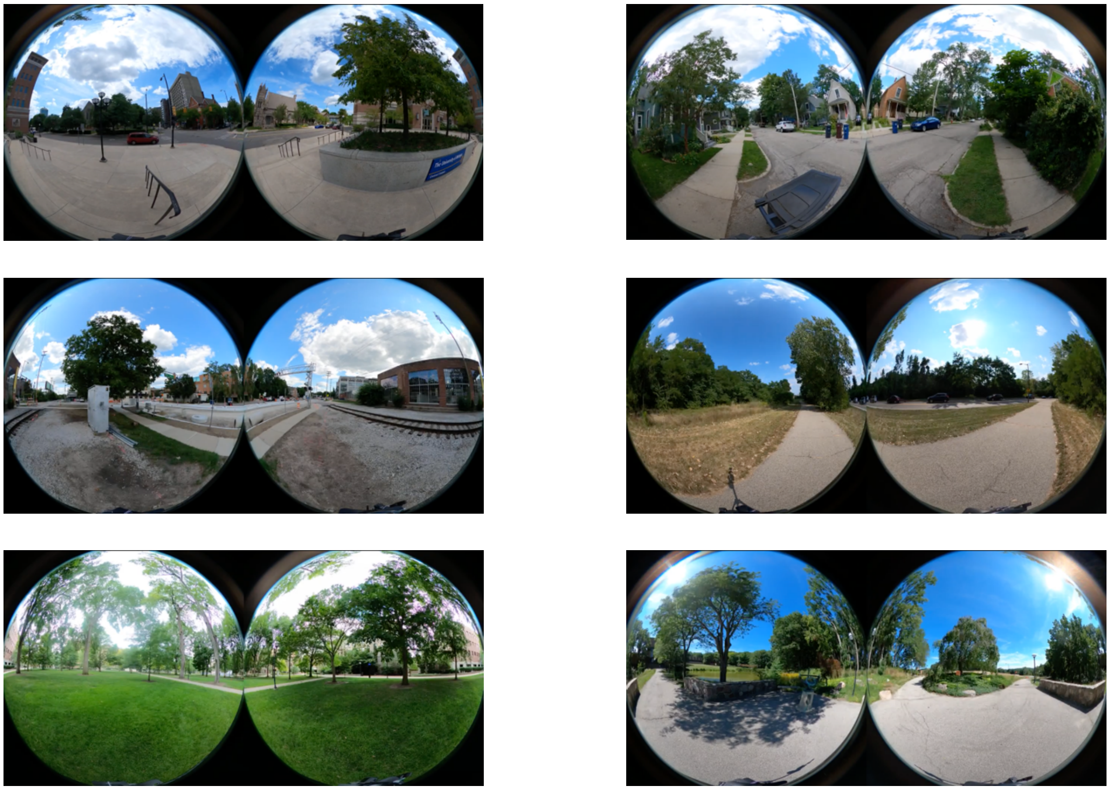


Figure 3.2: Example locations of the 200 samples. Top left: Downtown street; Top right: Residential; Middle left: Mian road; Middle right: Suburban; Button left: University campus in downtown; Bottom right: University campus in suburban.

3.2 Sound Recording

To test the affective qualities of soundscape, we randomly sampled the ambient sounds of 200 locations across a gradient of urban greenspace. Sounds were recorded for 5 minutes using a Zoom H6 recorder[2]. Sound samples were captured from 2:00 pm to 5:30 pm from July to August 2020. During the data collection, the sound recorder was connected to a tripod, which was adjusted to a height of 170 cm to best approximate the average listener's perspective of the soundscape. The operator remained silent during the recording to avoid any sound activity related to data capture. These audio clips were developed into spectrograms for valence and arousal prediction.

2 https://zoomcorp.com/en/us/handheld-recorders/handheld-recorders/h6-audio-recorder/

3.3 Dataset

The Emo-soundscape dataset created by Fan et al. was used to train the prediction model of emotion perceived from soundscape (Fan et al., 2017). This dataset was created for soundscape emotion recognition, which aims at the automatic recognition of emotions perceived from the spectrograms of soundscape recordings. The Emo-soundscape dataset consists of 1,213 environmental sound clips that were collected from Freesound[3], a social media platform with more than 500,000 sounds from 8 million registered users.

Each audio clip from Emo-soundscape includes a 6-second sound file and contains quantitative annotations of valence (pleasantness) and arousal (eventfulness), annotated by 1,182 trusted respondents from 74 different countries. Sounds are classified using a support vector regression model based on the emotional responses as a training dataset. Predicted accuracy assess arousal is (MSE = 0.048 and R2 = 0.855) much better than that for valence (MSE = 0.124 and R2 = 0.629). Numerous additional studies have contributed to the prediction accuracy of the Emo-soundscape dataset by optimizing deep learning methods (Ntalampiras et al. 2020). Meanwhile, some previous works have achieved higher classification accuracy of environmental sounds by using combined features rather than single features. For example, Peng et al. improved the recognition performance of environment sound classification based on horizontally combined images of LM (Log-Mel Spectrogram) and MFCC (Mel frequency cepstral coefficient), which can extend the breadth of audio information. (Peng et al., 2020). Similarly, Yu et al. vertically connected the images of audio features (LM, MFCC, Chroma, Spectral Contrast, and Tonnetz features) and achieved 94.6% classification accuracy on the UrbanSound8K dataset (Salamon et al., 2014), which is higher than previous models (Piczak, 2015, Zhang et al., 2017). In this study,

[3] https://freesound.org/

we used the same feature combination to train deep learning models. First, we used Librosa, a python library, to respectively extract LM, MFCC, Chroma, Spectral Contrast, and Tonnetz features as RGB-based images from each audio clip in the Emo-soundscape and then vertically combined the images (Fig. 3.3 a). The resolution of the final spectrogram combinations used for the prediction was 1110 x 344 pixels.

3.4 Auditory Data Processing

To predict soundscape emotion metrics, the sound recordings of each of the 200 locations were converted to spectrogram combinations using the same Emo-soundscape method (see Fig. 2.3 above). Each 5-minute recording was split into 50 6-second clips to match length of each clip in Emo-soundscape dataset, with the mean pleasantness and eventfulness computed by prediction models for each location based on the 50 spectrograms. Therefore, for prediction purpose, the clipped sound recordings of each sampled sites were transformed into the combined spectrogram using the same method applied to processing the data in Emo-soundscape dataset.

3.5 Constructing and Training Models

PyTorch-based package Fastai (https://www.fast.ai/), a layered API for deep learning was used to create and train a state-of-the-art vision model based on the Emo-soundscape dataset (Fan et al., 2017) for our goal of predicting two emotion metrics of perceived soundscapes separately (Fig. 3.4). The dataset was split with Fastai, into a training set and a testing set using a pre-trained CNN model. Each spectrogram of sound recoding has a corresponding numeric value of soundscape emotion (pleasantness or eventfulness) (Fig. 3.3 b). A ResNet-34 model pre-trained on the ImageNet dataset from PyTorch for image feature extraction and the one-cycle policy (learning rate changes during the training) was used to determine optimum learning rates in the training dataset and computational efficiency (He et al., 2016, Smith et al., 2018). The r-square score was used to evaluate model performance. Using this deep learning approach, we built two predictive models of pleasantness and eventfulness.

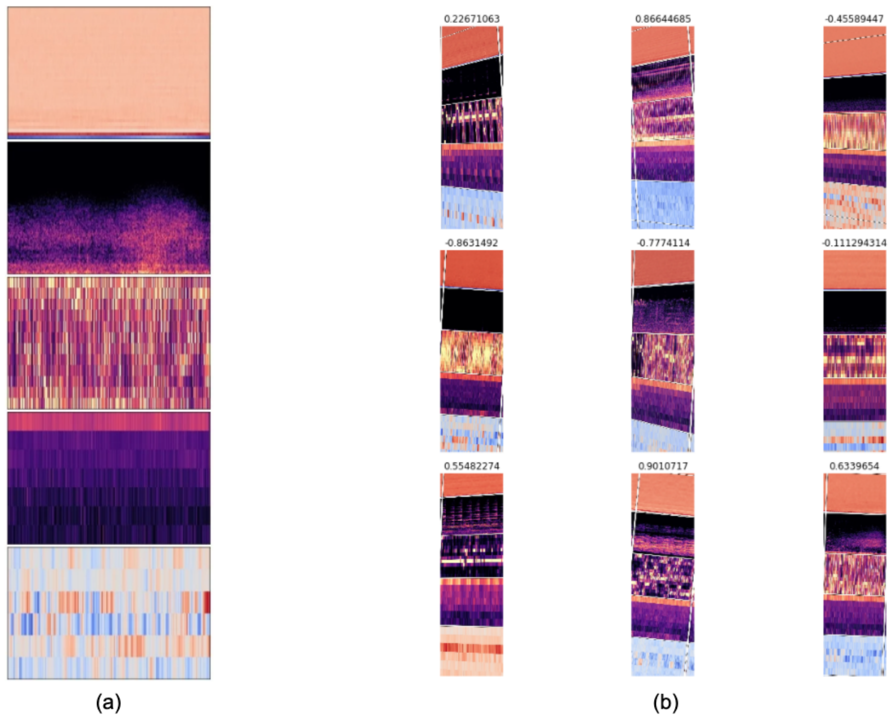


Figure 3.3: Example spectrograms (combinations) extracted from sound clips. (a): the spectrograms from up to bottom are MFCC, Log-Mel Spectrogram, Chroma, Spectral Contrast, and Tonnetz features. (b): the training dataset includes combined audio features corresponding to labeled emotion metrics (valence or arousal).

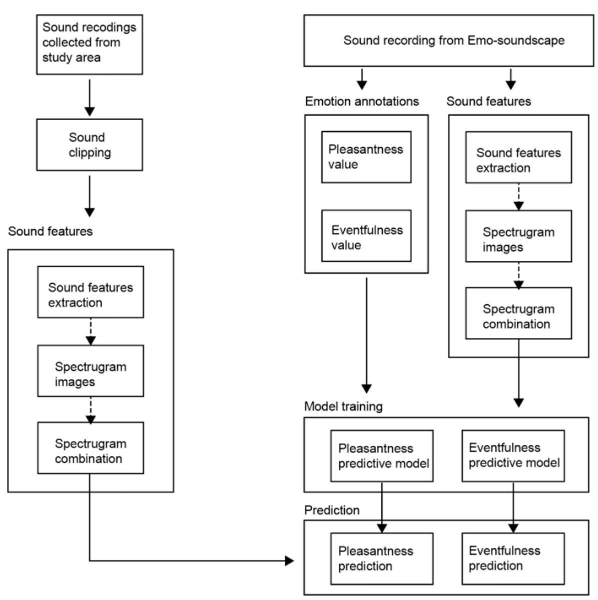


Figure 3.4: Data processing and training for prediction

3.6 Urban Landscape and Soundscape Emotion Modeling

To analyze the relationship between pleasantness and eventfulness and urban green space, we created a high-resolution map of the tree canopy (see Appendix A) using LiDAR (light detection and ranging) data. The classification was based on the highest returns of the point cloud data and computed using the lidR package in R (Roussel, 2021). The package enables the analysis of LiDAR while also supporting segmentation including ground classification and canopy segmentation (Roussel et al, 2020). Pixel mixing and misclassification were post-processed and corrected using building footprints (see Appendix B) from the City of Ann Arbor Data Catalog (https://www.a2gov.org/services/data/Pages/default.aspx) to remove some areas in building footprints that were incorrectly classified as the canopy. For statistical purposes, we generated a square polygon grid over the study area, where each square is 30m by 30m. Then, we used zonal statistical analysis to calculate the percentage of canopy over the square polygon grid based on the high-resolution map of the tree canopy (Fig. 3.5 a1).

In order to explore the spatial relationships between the soundscape emotion metrics and the derived canopy, we interpolated point distributions of valence and arousal of 200 samples to create heat maps (see Appendix C) using the kernel density estimation (KDE) in Qgis, an open-resource GIS platform. As one of the most popular point pattern analysis methods (Bailey et al., 1995, Silverman, 2018), KDE is used to produce a smooth density surface based on the distribution of point events over space by computing event intensity as density estimation (Xie et al., 2008). Based on the heat maps, we used zonal statistical analysis to average the means of both pleasantness and eventfulness through the generated square polygon grid in the similar way we used to get the percentage of the canopy cover (Fig. 3.5 a2), and this was overlain on the canopy layer (Fig. 4.6). Based on the zonal statistic analysis of soundscape emotions and canopy cover, bivariate maps were developed to visualize the spatial distribution of soundscape emotion and the spatial relationship between soundscape emotion metrics and canopy cover (Fig. 3.5 b1). For mapping purposes, the zonal statistic values of canopy, pleasantness, and eventfulness were broken down to five graduated classes based on the quantile method (see Appendix D). A Spearman's rank correlation was calculated over the squares in the 30m-by-30m polygon grid to measure spatial associations between the soundscape emotion metrics and the canopy area across the whole study area and different accessible areas of hospitals (Fig. 3.5 b2)

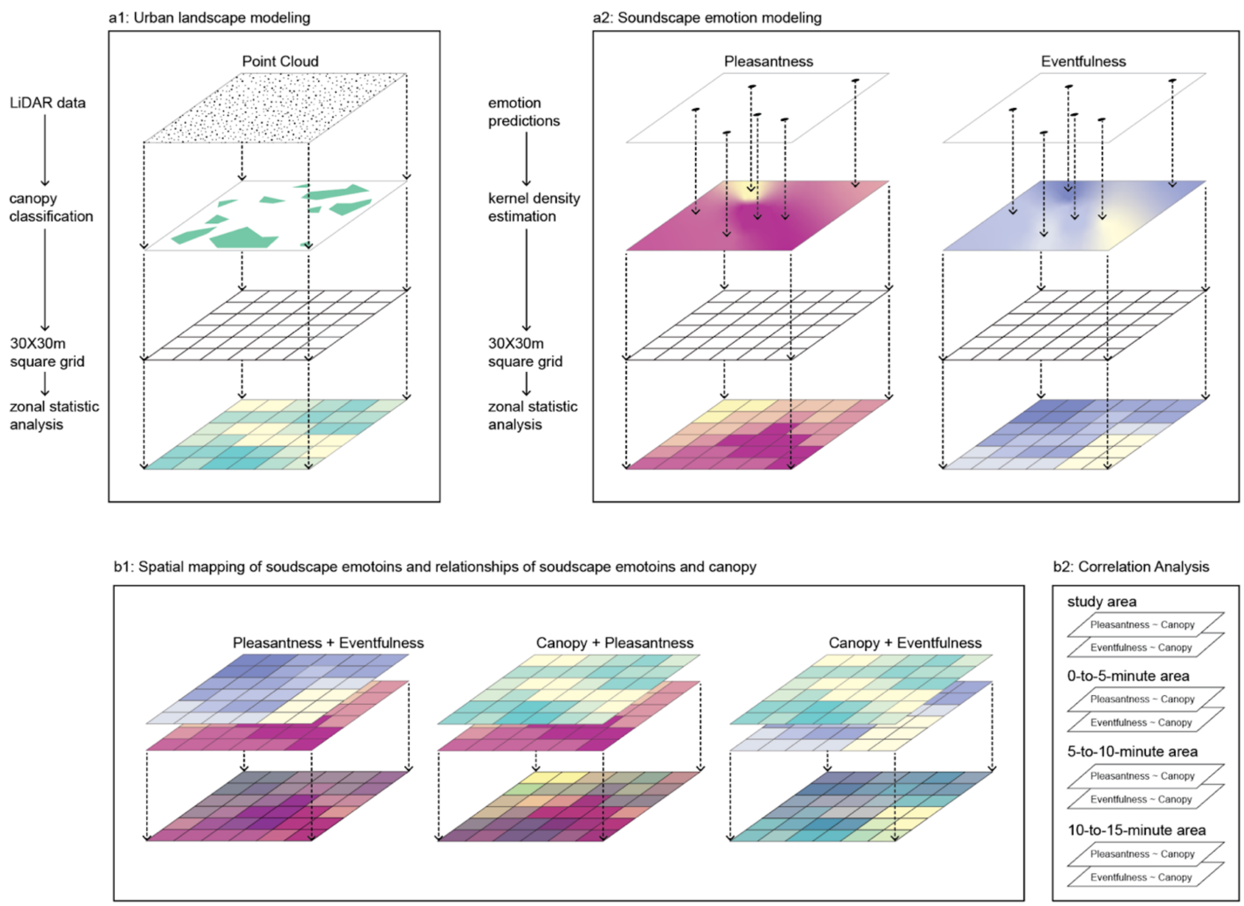


Figure 3.5: Mapping and analysis of canopy cover and soundscape emotions (a1: the percentage of canopy cover area in each 30X30m square was calculated based on LiDAR-based canopy distribution, a2: the means of pleasantness and eventfulness were extracted from the interpolations based on emotion predictions of 200 sites, b1: bivariate maps were generated based on zonal statistical analysis using multiply overlaying, b2: correlation analysis was conducted over the whole study area and three accessible areas of hospitals)

Chapter 4 Result

Result analysis includes emotion predictive models, soundscape emotion mappings, and correlations between green canopy and perceived emotions. Both Eventfulness and pleasantness are significantly associated with canopy distributions.

4.1 Prediction Models

Comparing our test data with the model of the training dataset resulted in prediction accuracy (r^2 score) for eventfulness of 0.857 and 0.666 for pleasantness. This is comparable to other studies predicting the emotional response to environmental sounds (Fig. 4.1) (Lundén et al., 2010, Fan et al., 2015, Fan et al., 2017).

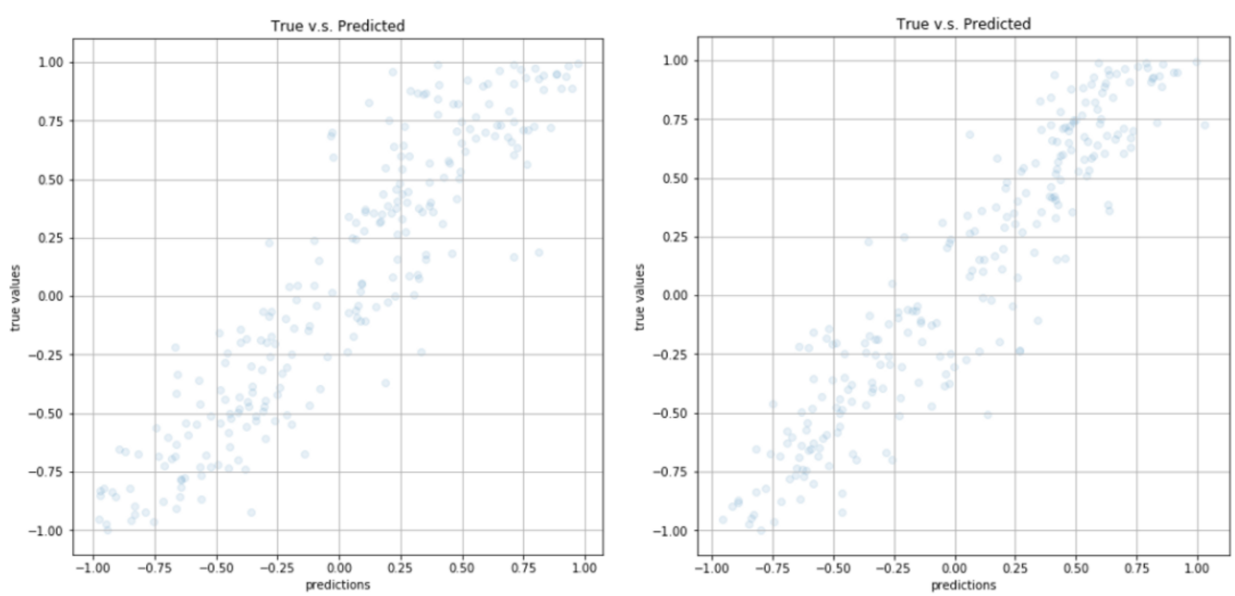


Figure 4.1: Training results of prediction models (left: pleasantness, right: eventfulness)

4.2 Soundscape Emotion Prediction

The mean values of pleasantness and eventfulness vary across all accessible areas of hospitals. According to the soundscape dimensions, the majority of the sites are highly eventful and rate low for pleasantness (i.e., located in the ‘chaotic’ quadrant) whereas the highly pleasant areas are less eventful (i.e., in the ‘calm’ quadrant) (Fig. 1.1, Fig. 4.2). Most sites assessed as eventful and unpleasant (chaotic) fall in the 5-minute accessible area of hospitals, where the most eventful site was found. In 5-minute and 10-minute accessible areas, the emotional responses of the sites vary on the entire eventful scale and negative side of pleasant scale. However, the most pleasant and uneventful areas (Fig. 4.2), those in the calm quadrant, are within walking distance of 10 to 15 minutes of the hospital. Specifically, the most unpleasant site is located within the 10-to-15-minute accessible area. There is no site falling in the ‘exciting’ quadrant, which means that there is not a site with a combination of positive pleasantness and positive eventfulness.

According to the distribution of pleasantness across three accessible areas (Fig. 4.3), the 5-minute accessible area respectively has the lowest first quartile, lowest median, and lowest third quartile. The 10-to-15-minute accessible area has the highest first quartile, highest median, and highest third quartile. The distribution of pleasantness in the 5-to10-minute accessible area is between the 5-minute area and 10-to-15-minute area. For the distribution of eventfulness, the 5-to-10-minute accessible area has the lowest median, and lowest third quartile whereas the first quartile is slightly higher than that of the 5-to10-minute accessible area. The 10-to15-minute accessible area has the highest first quartile, highest median, and highest third quartile.

In the map of the distribution of soundscape emotion of samples (Fig. 4.4), there is one location with relatively pleasant soundscapes that are also somewhat eventful in the city center near the university campus and hospitals among relatively more eventful and less pleasant (relatively chaotic) locations within the 0-to5-minute accessible area. Within the 10-to15-minute accessible area, most sites with relatively high pleasantness and low eventfulness are concentrated in the south. The most chaotic (eventful and unpleasant) soundscapes are also found along some roads. This can be explained by a previous study that found that traffic noise is a dominant sound source influencing the soundscape quality (Hong and Jeon, 2015).

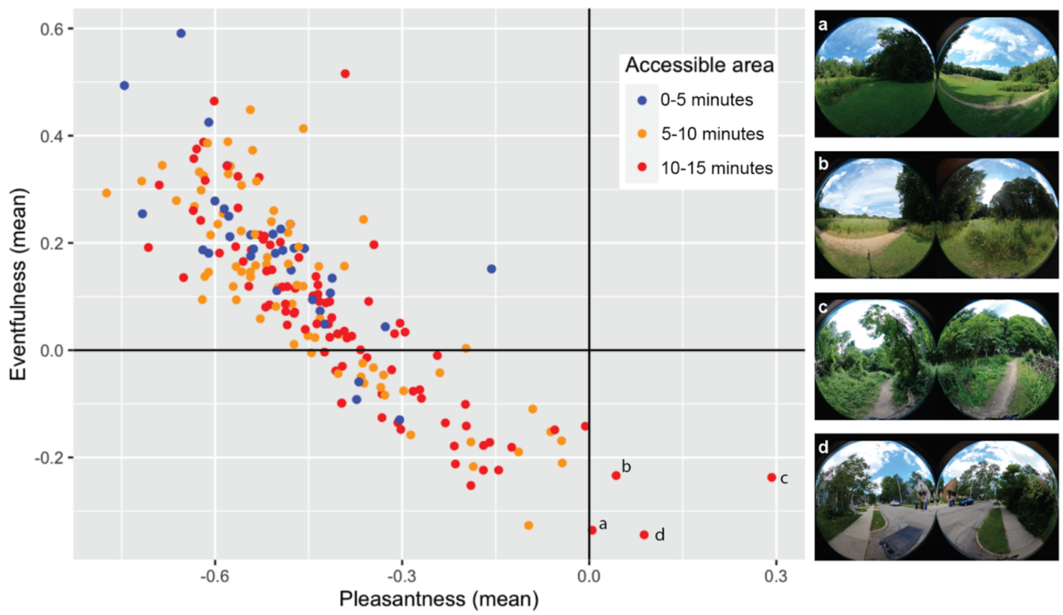


Figure 4.2: Prediction results of pleasantness (x-axis) and eventfulness (y-axis) based on accessible areas. Results indicated that the locations were classified as unpleasant and varying in eventfulness, aside from four locations (a, b, c, and d) that were pleasant and uneventful.

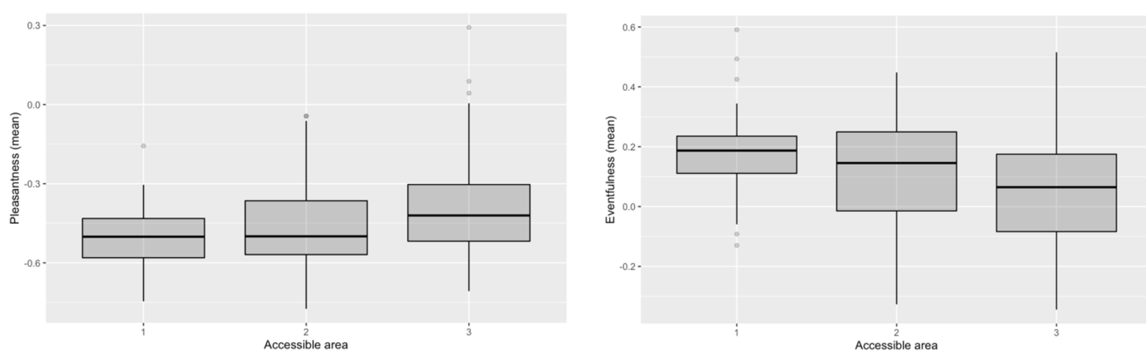


Figure 4.3: Sample distributions of pleasantness (left) and eventfulness (right) based on accessible areas (1: 0-5 minutes, 2: 5-10 minutes, 3: 10-15 minutes

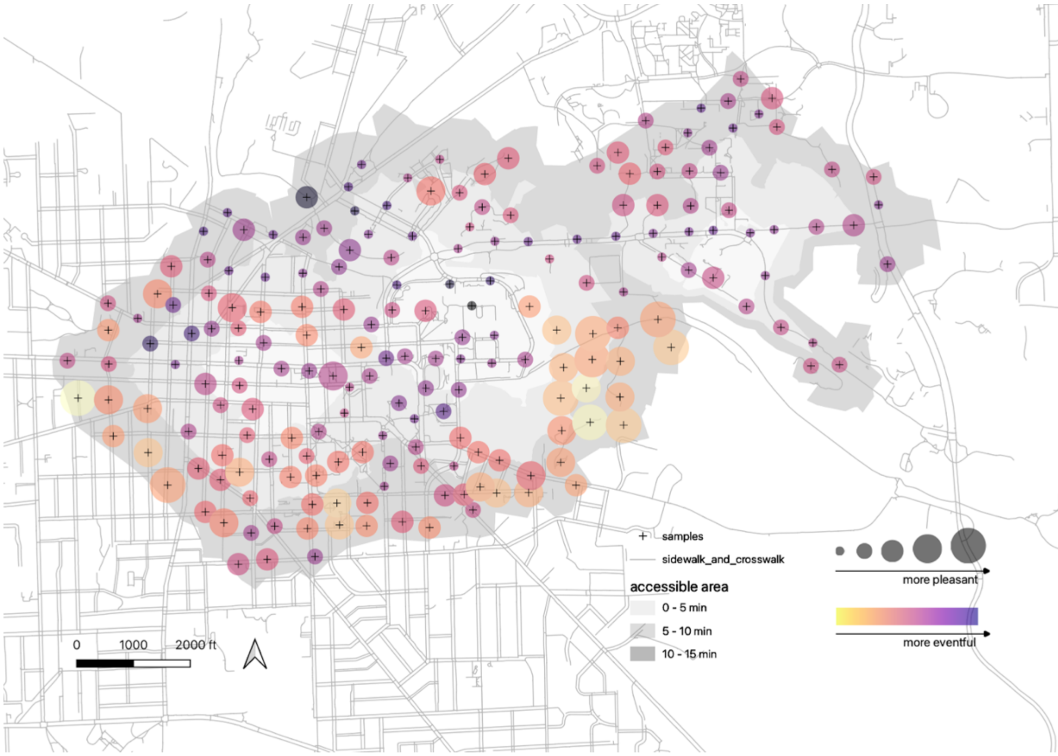


Figure 4.4: Distribution of soundscape emotion of samples. The larger circle indicates a more pleasant emotional response and the darker color means a more eventful emotional response.

Many areas with relatively lower predicted pleasantness scores have relatively higher eventfulness scores (shown in purplish-red) over residential and open space within 10-15 walking distance of hospitals and the reverse (shown in blue-grey) is also true (Fig. 11). The intersection of the least pleasant and most eventful areas, the chaotic quadrant, is located near the university campus, commercial areas, and residential areas within 0-to10 minute walking distance. Compared to these areas, the soundscape over the commercial streets across the downtown area is predicted to be relatively eventful, being more pleasant. The occurrence of relatively chaotic soundscapes over campus, residential areas, and commercial areas suggest that sound sources from different types of activities or mechanical sounds could affect the soundscape perception. Additionally, the least pleasant and eventful areas (shown in yellow) over residential areas are lying between the walking distances of 5-10 minutes and 10-15 minutes in the south of the study area (Fig. 4.5).

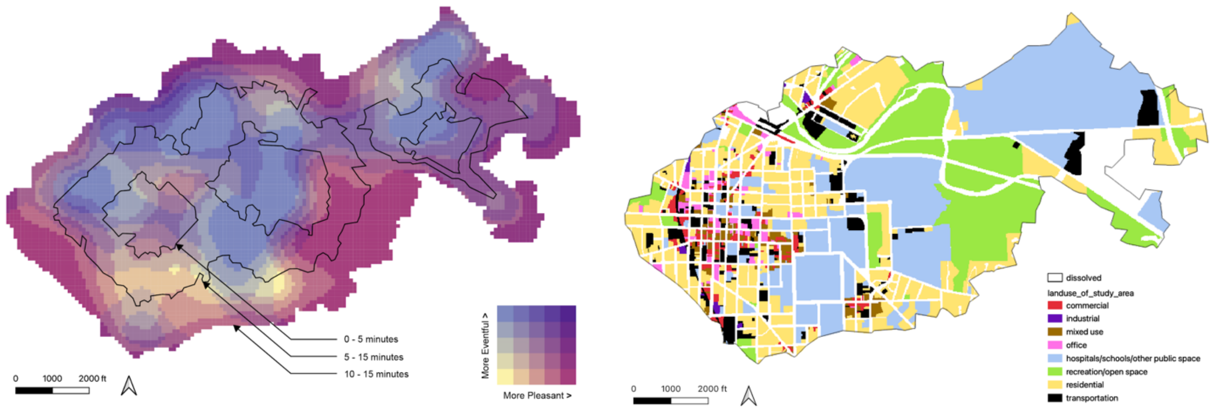


Figure 4.5: Bivariate map of pleasantness and eventfulness (left) and land use map (right)

4.3 Urban Green Space

Based on the zonal statistic of percentage of the canopy (Fig. 4.6), the distribution of canopy cover in three accessible areas of hospitals (Fig. 4.7) shows that the 0-to-5-minute accessible area has the lowest first quartile, lowest median, and lowest third quartile whereas the 10-to-15-minute accessible area has the highest first quartile, highest median, and highest third quartile. The distribution of canopy in the 5-to10-minute accessible area is between the 5-minute area and 10-to-15-minute area. Overall, the further away from the hospital, the higher the canopy cover.

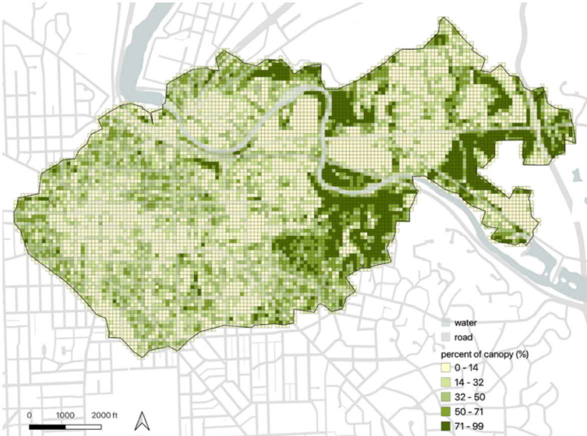


Figure 4.6: Zonal statistic map of percentage of canopy in 30X30m square grid

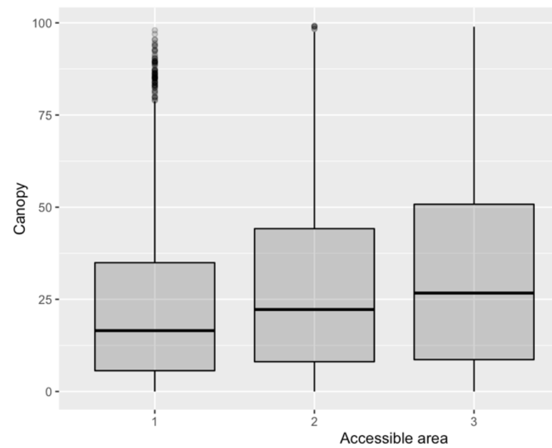


Figure 4.7: Distribution of percentage of canopy based on accessible areas (1: 0-5 min, 2: 5-10min, 3: 10-15min)

4.4 Spatial Relationships of Soundscape Emotion and Urban Green Space

Overall, many areas, which tend to be more eventful and less pleasant, have less canopy cover than other areas. The areas where the canopy is more concentrated, connected, and dense, tend to be more pleasant and less eventful. Many of the locations with eventful and pleasant soundscapes are in the most accessible area of hospitals (10-minute walking distance) even with its canopy cover (Fig. 4.8). This probably implies that the soundscapes near the hospitals are highly influenced by urban contexts, including traffic noise.

For the Spearman's rank correlation between pleasantness and canopy cover the p-value is 2.210-16, which is less than the significance level alpha = 0.05 with a correlation coefficient of 0.3186 (Fig. 4.9). Therefore, pleasantness is positively associated with canopy cover percent. For the Spearman's rank correlation between eventfulness and canopy cover the p-value of the test is 2.210-16, which is less than the significance level alpha = 0.05 with a correlation coefficient of -0.3078 (Fig. 4.10). Therefore, eventfulness is negatively associated with canopy cover. Interestingly, within each walking distance of hospitals, the spatial correlation between both pleasantness and eventfullness and canopy cover is compatible with the spatial correlation over the whole study area. Specifically, the correlations between soundscape emotions and canopy (R=0.34 for pleasantness-canopy and R=-0.31 for eventfullness-canopy) in the 0-5 minutes walking distance is stronger than those (R=0.26 for pleasantness-canopy and R=-0.22 for eventfullness-canopy) in the 10-15 minutes walking distance. The correlations between soundscape emotions and canopy (R=0.36 for pleasantness-canopy and R=-0.36 for eventfullness-canopy) in the 5-10 minutes walking distance is slightly stronger than those in the 0-5 minutes walking distance (Fig. 4.11).

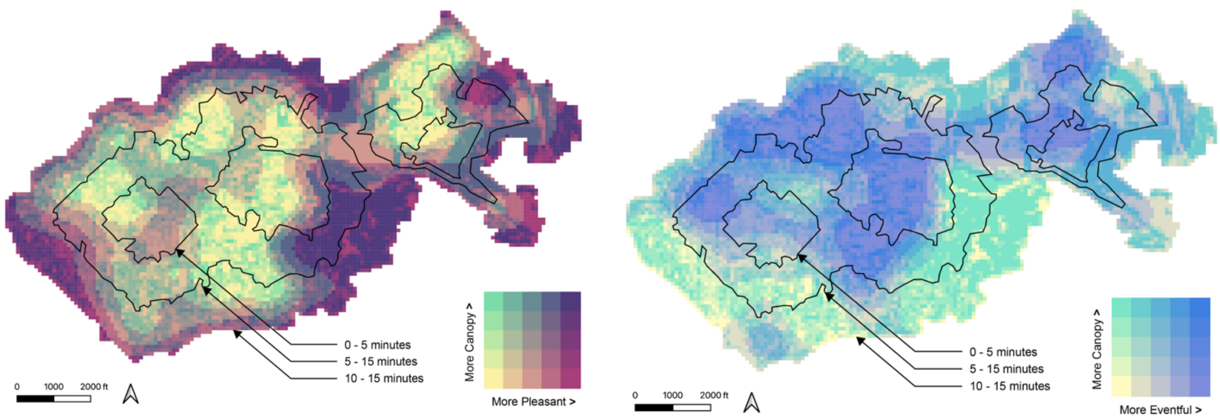


Figure 4.8: Bivariate maps of pleasantness (left) / eventfulness (right) and canopy cove

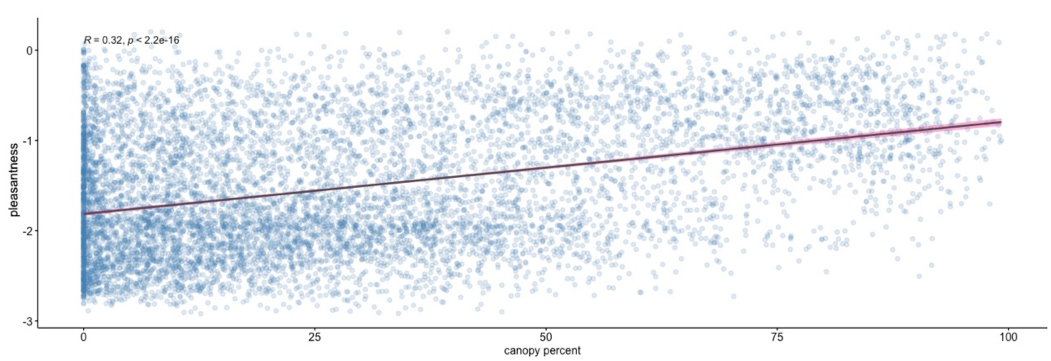


Figure 4.9: Spearman's rank correlation of pleasantness and canopy

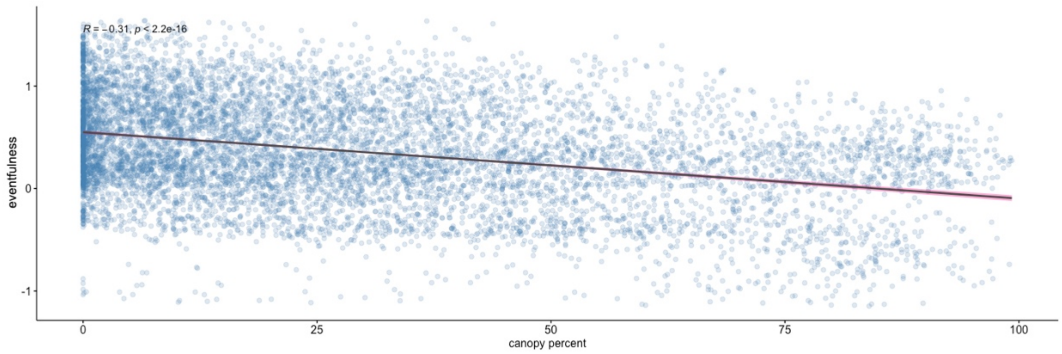


Figure 4.10: Spearman's rank correlation of eventfulness and canopy

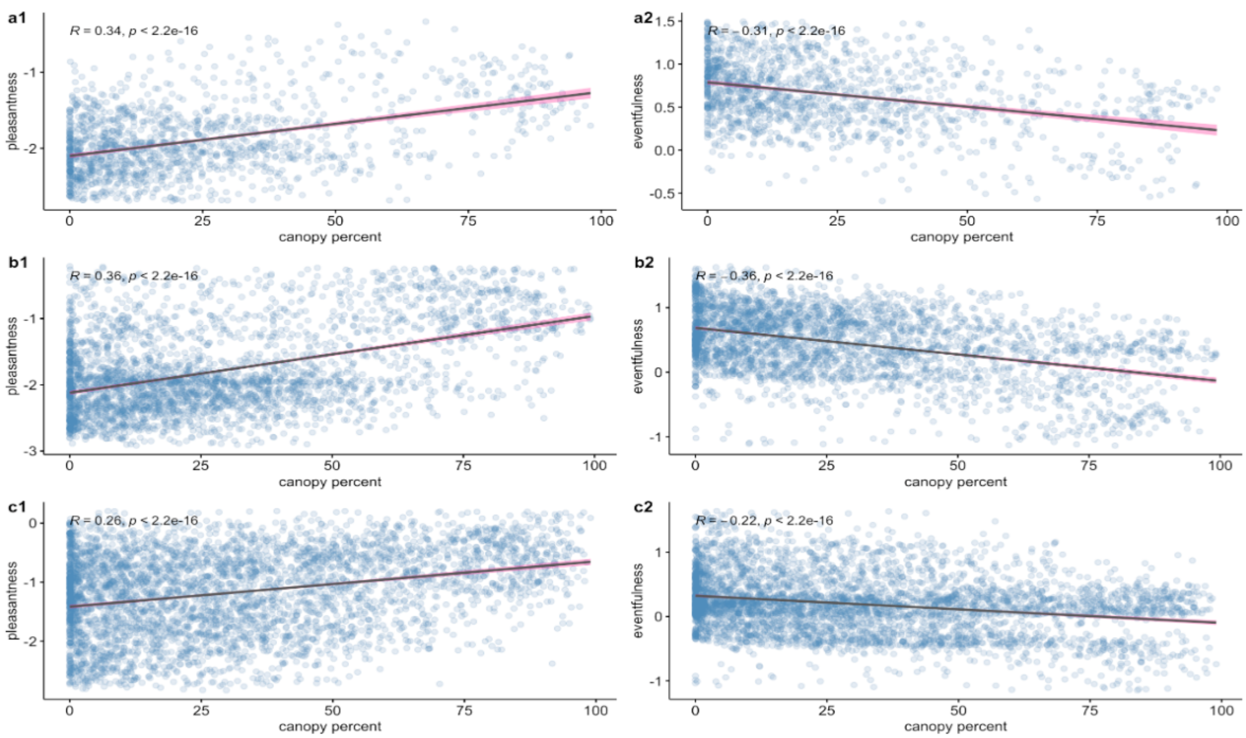


Figure 4.11: Spearman's rank correlation of pleasantness (left) / eventfulness (right) and canopy cover in three accessible areas of hospitals (a1-2: 0-5 minutes, b1-2: 5-10 minutes, c1-2,: 10-15 minutes)

Chapter 5 Discussion

5.1 Methodological Considerations

There are several research gaps that need to be filled: (i) landscape's restorative benefits have not been spatially assessed in the urban-scale acoustic environment, (ii) the spatial association between perceived emotion from soundscapes and landscapes is still unclear, (iii) and a machine learning method for quantitatively mapping sound emotions needs to be established. To fill these gaps, this study proposed a quantitative mapping solution based on an image-based deep learning approach to explore spatial relationships between soundscape emotions and green space.

The urban-scale soundscape mapping and assessment have been a key focus of well-being-related landscape and urban planning in recent decades. However, there has been a lack of effective methods to model perceived response to soundscape due to the spatiotemporal limitation of the self-report methods such as situ questionnaires, soundwalks, and laboratory assessments. Based on recent studies of soundscape prediction leveraging machine learning and the ground-truth dataset, we explored the possibility of training an annotated soundscape dataset to identify areas with high and low emotional responses on a large scale. Built on this exploration, we proposed a framework that combines perception-based assessment of perceived emotion and GIS-based landscape modeling of high-resolution spatial data to study the therapeutic potential of landscapes. This method was applied to an urban area covering gradients of urban impervious and natural landscapes. Previous studies have not predicted soundscape emotion based on such large samples across a large area (Kang et al., 2018).

5.2 Soundscape Emotion and Urban Green Space

What is the correlation between urban green spaces and soundscape emotions? The correlation measures the strength and direction of association between the degrees of pleasantness and eventfulness and the percentage of the green canopy. The emotion predictions of both pleasantness and eventfulness are consistent with the conclusions of traditional field-based research that used surveys and participatory methods (Alvarsson et al., 2010, Erfanian et al., 2021). In our study, weak correlations were found between soundscape emotions and green spaces. Pleasantness is positively associated with canopy cover, which means that trees can potentially contribute to improving soundscape quality. On the contrary, tree canopy cover is negatively related to eventfulness. Therefore, we find a chaotic emotional response in low tree cover areas, given their location in the top-left quadrant defined by the dimensions of unpleasantness and eventfulness. The correlations can be explained by the effect of the green belt on the mitigation of traffic-generated noise in parks and suburbans (Pathak et al., 2011, Herrington, 1974). Additionally, natural soundscapes are rated as more restorative than urban acoustic environments (Krzywicka & Byrka, 2017; Payne, 2008, 2013) and green spaces serve as habitats of species that enrich natural soundscapes. Previous studies have demonstrated that bird species richness that is important to calm is supported by the ecological quality and vegetation structure of urban green spaces (Uebel et al, 2021, Irvine et al., 2009). Although people are not exclusively concerned with acoustic perception directly, exposure to poor soundscapes could lead to adverse impacts on mental health such as reduced sleep, stress, and impaired cognitive abilities (Schwela, 2000, Hume et al., 2012, Jennings and Cain, 2013). Our results here add to the growing evidence base on the value of increasing canopy cover to positively affect the restorative potential of natural soundscapes (Van Renterghem, 2019). Additionally, our methodology can help operationalize this research in urban-scale planning applications.

What is the ideal urban configuration for the improvement of psychological benefit in urban soundscapes? To achieve optimal urban configurations for positive health outcomes, it is important to understand the distribution of health-related soundscapes and the structure of spatial configuration (Yamada, 2006, Djenaihi et al., 2021). This study further explored the distribution of landscapes and perceived soundscape emotions, which can inform urban planning and policymaking to improve soundscape quality and accessibility. Our study found that the soundscape tends to be perceived as chaotic within 10-minute walking distance of hospital facilities in spite of the existence or proximity of the green canopy. However, our soundscape emotion mapping also shows that both pleasantness and eventfulness are linked to urban land cover and land use including transportation, commercial areas, university campuses, residential, etc. In a small range of areas within the same distance, soundscapes on the commercial streets and university campus could be perceived as relatively pleasant and eventful, resulting in the prediction of an excited emotional state. This might be because the traffic volumes are slower in the city core due to traffic congestion, and the university campus has relatively more vegetation and fewer people density than the surrounding areas. Automobile traffic patterns around the center of the downtown area may contribute to a chaotic soundscape that is less pleasant, while suburban areas and urban parks are likely to have relatively less traffic noise. Besides, the dominant sound source and soundscape quality could be different depending on the main functions of places. For example, eventful sounds from human activities in commercial areas could be positive but adverse in residential areas (Hong and Jeon, 2015). Therefore, strategies of soundscape planning should consider the various contexts in the urban center. Our study also indicates that both pleasant and uneventful soundscapes with woods are greater than a 10-minute walking distance, implying that the current accessibility of restorative soundscapes for hospitals in our study area may be insufficient. Previous studies showed overall agreement on the well-being benefits of nature-based health interventions but still rarely considered these interventions for enhancing health outcomes of soundscape for people (Shanahan et al., 2019, Buxton et al., 2021). Our results support the evidence that urban green space soundscapes with less urban sonic intervention have greater potential for restorative benefits to psychological health. For existing greenspace beside hospitals in urban centers, introducing birdsong by planting trees may contribute to noise mitigation and traffic reduction. Meanwhile, urban planners and policymakers may want to increase speed limits or introduce new traffic plans in the proximity of existing large parks to protect restorative soundscapes.

5.3 Limitation and Future Study

We note that our study has limitations due to the sound sample procedure, the possibility of omitted variables that are important to explaining soundscape emotion, and errors in prediction. On the one hand, the sound recording of each position only lasts 5 minutes, which is only a short period in a day and is not long enough to represent the holistic acoustic environment of each location. On the other hand, there may be other factors like land use, land cover, traffic volume, species, visual experiences, and demographics, which could explain emotional response to soundscape (Hong and Jeon, 2015, Hong, 2016, Uebel et al., 2021). Also, the acoustic environment can be very dynamic and different from day to night (Hong et al., 2014), which means the temporal patterns of emotional response also need to be explored in further study. Moreover, the absolute value of sound emotion predicted by the deep learning model might not wholly reflect the real-world perceived soundscape emotion due to the limited accuracy of prediction models and the limited samples of sound recordings. But it could be worthwhile to use this method to study spatial relative relationships between a specific variable of soundscape emotion and urban contexts like green spaces, urban forms, buildings, etc. (Okba et al., 2021). Additionally, the collection of the sound recordings that we used to model soundscape emotion costs a lot of time and labor, which will still limit the capacity of large-scale soundscape prediction and mapping. However, the growing crowdsourcing data of environmental sounds with geotags such as Freesound, etc. will make our framework promising for future therapeutic landscape research in a larger urban context because sound data could be directly downloaded by API instead of on-site recording.

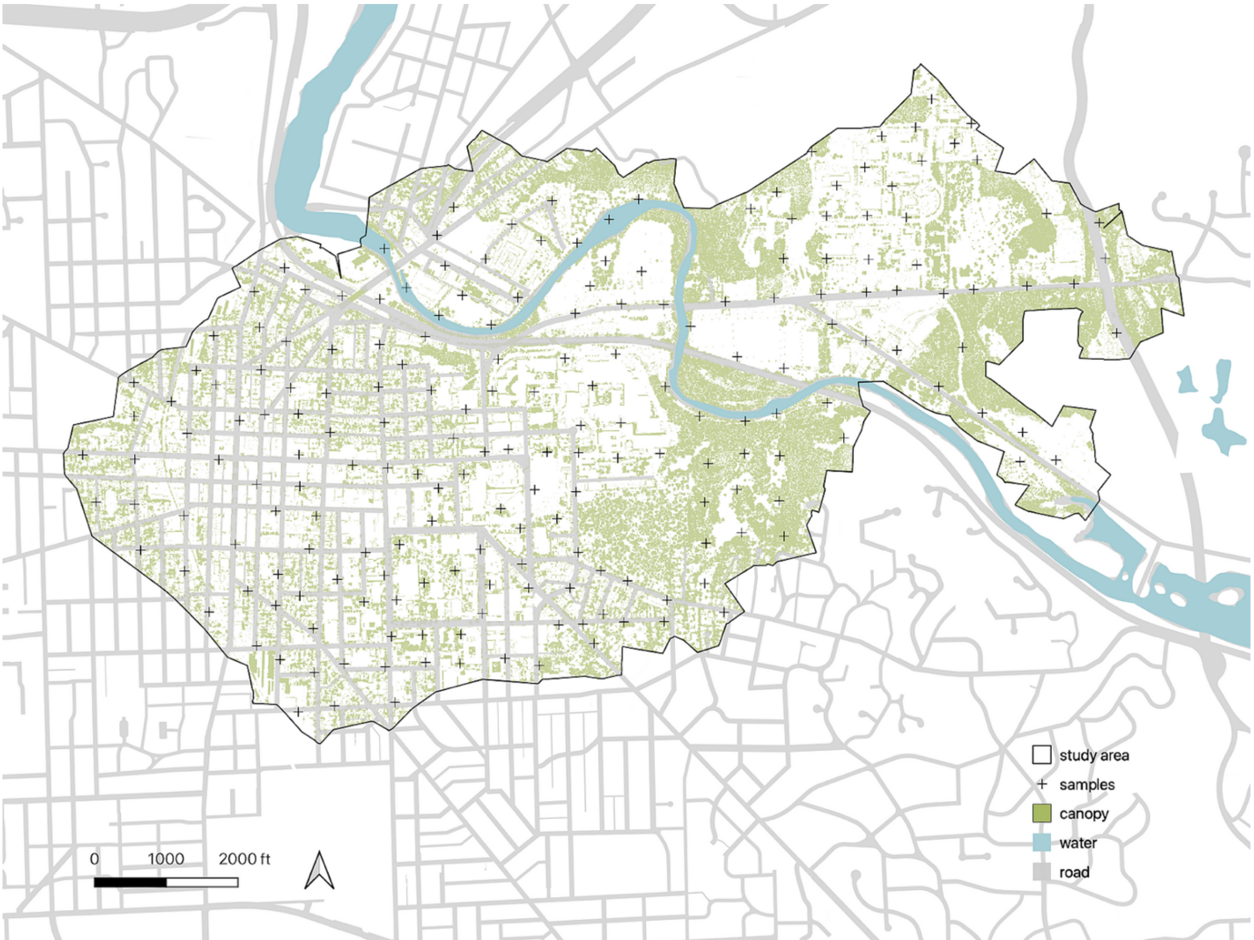
The approach proposed in this paper has the potential to include more physical environment factors to discover more pathways in which environmental configuration affects the emotional response to soundscapes. In order for the approach to be useful in design decision-making for urban environments, it is necessary to expand the prediction scope and reduce the bias that can be caused by spatial and acoustic variances and social differences. In terms of spatial and acoustic variances across cities, the audio data can be sampled from the larger domain to test correlation. Meanwhile, more correlations of spatial variables and perceived soundscape emotion should be examined. The variables like street view, urban form, and viewshed could be utilized as independent variables to build the prediction model based on learned relationships between soundscape emotions and environment configurations. Tabrizian et al. have proposed an approach for predicting the restorative potential of the landscape using visible parameters of viewshed as the input independent variables (Tabrizian et al., 2020). For social differences, future studies should consider using data such as social media data with georeferencing and demographics of the neighborhood.

Chapter 6 Conclusion

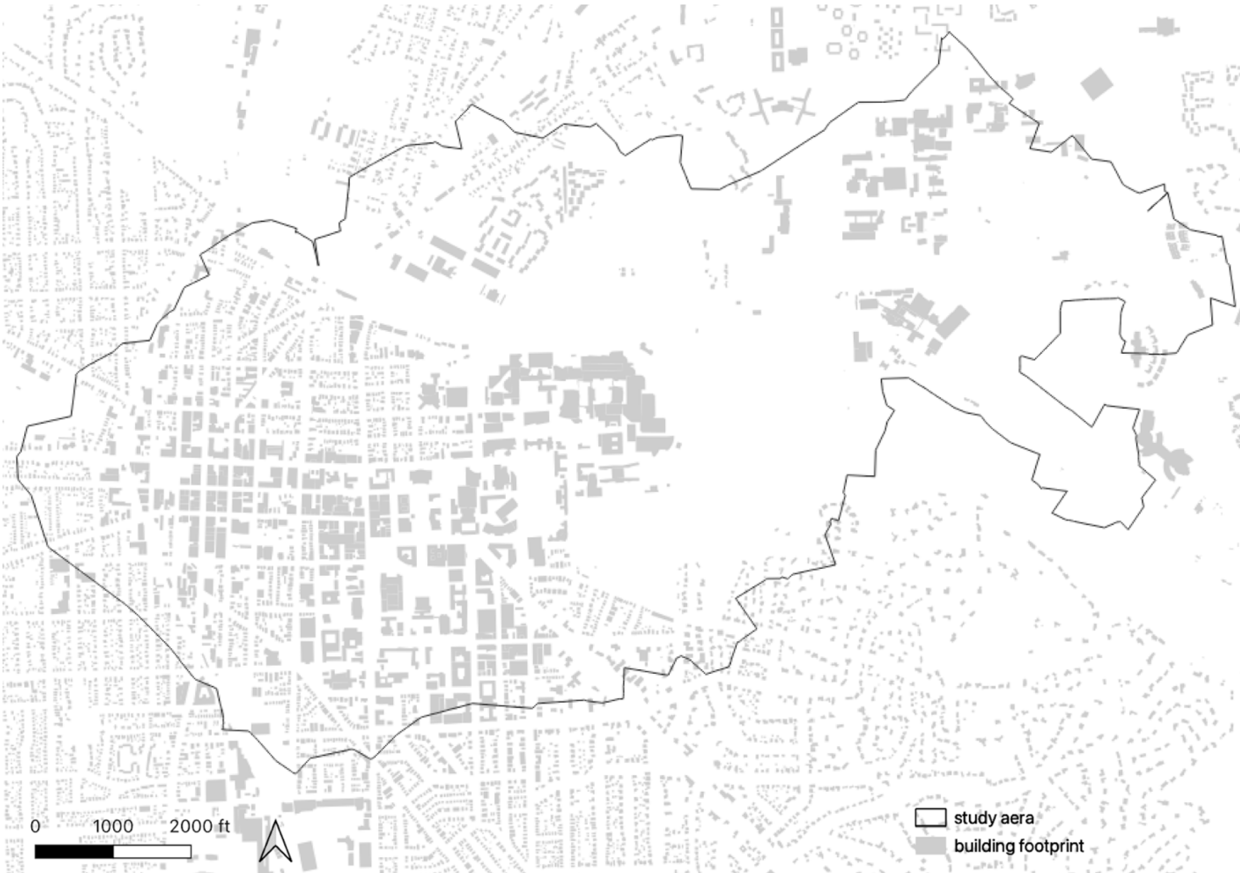
In summary, our method relies on five main parts, namely: (1) sound source collection, (2) prediction of perceived emotion, (3) implementation of soundscape emotion maps, (4) model of spatial landscape, (5) spatial correlation of soundscape emotion and landscape. For the second part, we demonstrated techniques for developing a predictive model of emotion using the dataset with ground-truth labels to automate the prediction of the large-scale soundscape emotions. Specifically, the dataset consists of environmental sound clips that were evaluated by experts with measures of pleasantness and eventfulness. Besides, this step novelly applied the conversion of audio files into images to soundscape assessment tasks, which is still a very rare approach (Mushtaq et al., 2021). For the third and fourth parts, we precisely modeled and mapped pleasantness and eventfulness and high-resolution green canopy using the zonal statistic method. The last part was built on the former two parts, we uncovered the spatial relationships and statistical associations between emotion and greenspace across the urban area.

As an experiment, the application of deep learning in the prediction of emotional response to soundscape has proven to be valuable not only in soundscape quality assessment but can be adapted to large-scale correlation analysis of soundscape and urban configurations and social factors in future landscape studies. Based on Spearman's rank correlation, both pleasantness and eventfulness have a significant correlation with greenspace. Pleasantness is positively correlated with greenspace, whereas eventfulness and greenspace are negatively correlated. We found that the dominant soundscape near hospitals and the transportation line tends to be chaotic whereas most calm soundscapes were found in urban parks. Our findings suggest that soundscape emotions are spatially associated with urban greenspace and the improvement of the urban landscape can enlarge the mental health benefit of the soundscape.

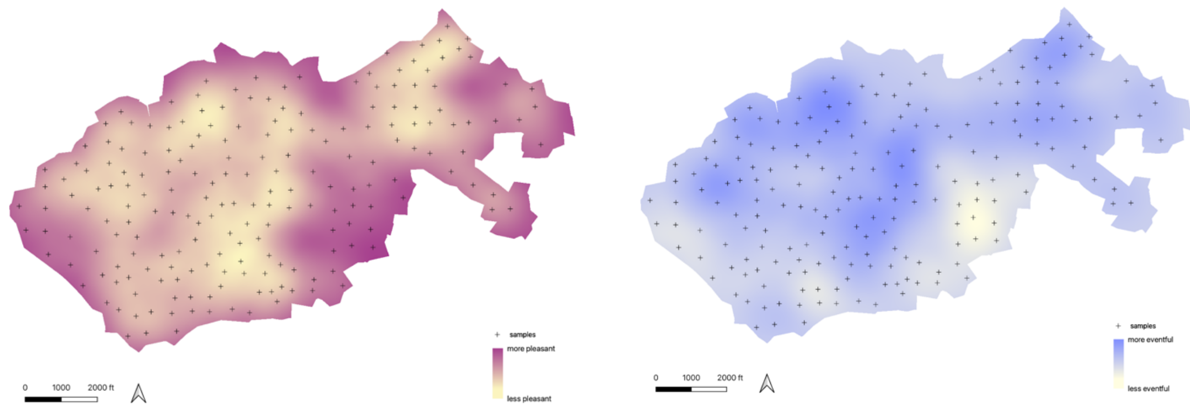
Appendices



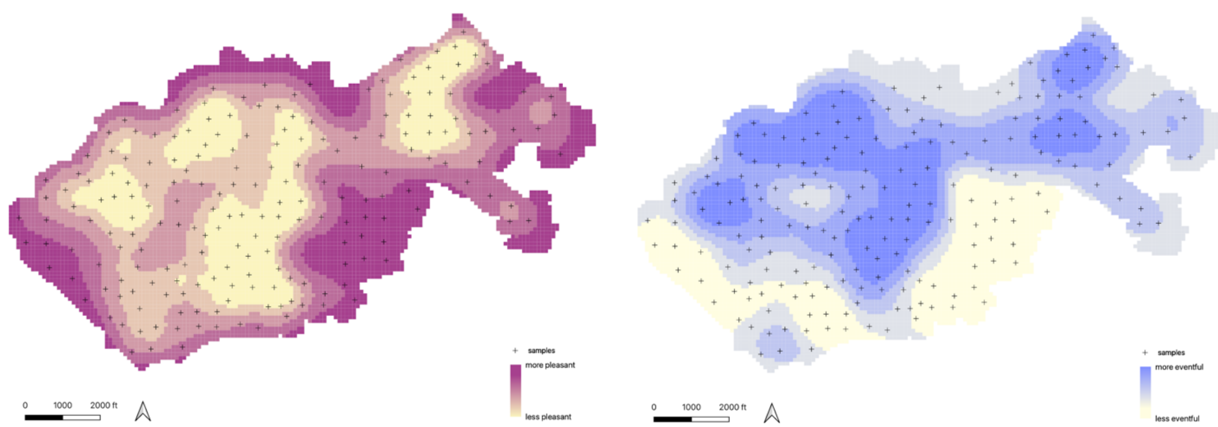
Appendix A: Distribution of canopy cover



Appendix B: Building footprint



Appendix C: Heat maps of pleasantness (left) and eventfulness (right)



Appendix D: Zonal statistical maps of pleasantness (left) and eventfulness (right)

Bibliography

Anderson, L. M., Mulligan, B. E., Goodman, L. S., & Regen, H. Z. (1983). Effects of sounds on preferences for outdoor settings. Environment and behavior, 15(5), 539-566.

Axelsson, Ö., Nilsson, M. E., & Berglund, B. (2010). A principal components model of soundscape perception. The Journal of the Acoustical Society of America, 128(5), 2836-2846.

Alvarsson, J. J., Wiens, S., & Nilsson, M. E. (2010). Stress recovery during exposure to nature sound and environmental noise. International journal of environmental research and public health, 7(3), 1036-1046.

Annerstedt, M., Jönsson, P., Wallergård, M., Johansson, G., Karlson, B., Grahn, P., ... & Währborg, P. (2013). Inducing physiological stress recovery with sounds of nature in a virtual reality forest—Results from a pilot study. Physiology & behavior, 118, 240-250.

Aletta, F., Oberman, T., & Kang, J. (2018). Associations between positive health-related effects and soundscapes perceptual constructs: A systematic review. International journal of environmental research and public health, 15(11), 2392.

Aletta, F., Guattari, C., Evangelisti, L., Asdrubali, F., Oberman, T., & Kang, J. (2019). Exploring the compatibility of "Method A" and "Method B" data collection protocols reported in the ISO/TS 12913-2: 2018 for urban soundscape via a soundwalk. Applied Acoustics, 155, 190-203.

Abri, F., Gutiérrez, L. F., Namin, A. S., Sears, D. R., & Jones, K. S. (2020, December). Predicting Emotions Perceived from Sounds. In 2020 IEEE International Conference on Big Data (Big Data) (pp. 2057-2064). IEEE.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry, 25(1), 49-59.

Bailey, T. C., & Gatrell, A. C. (1995). Interactive spatial data analysis (Vol. 413, No. 8). Essex: Longman Scientific & Technical.

Brocolini, L., Lavandier, C., Marquis-Favre, C., Quoy, M., & Lavandier, M. (2012, April). Prediction and explanation of sound quality indicators by multiple linear regressions and artificial neural networks. In Acoustics 2012 (pp. 2121-2126).

Brown, A. L., Gjestland, T., & Dubois, D. (2016). Acoustic environments and soundscapes. Soundscape and the built environment, 1-16.

Buxton, R. T., Pearson, A. L., Allou, C., Fristrup, K., & Wittemyer, G. (2021). A synthesis of health benefits of natural sounds and their distribution in national parks. Proceedings of the National Academy of Sciences, 118(14).

Cassidy, T. (2013). Environmental psychology: Behaviour and experience in context. Psychology Press.

Ciuk, D., Troy, A., & Jones, M. (2015). Measuring emotion: Self-reports vs. physiological indicators. Physiological Indicators (April 16, 2015).

Collins, R. M., Spake, R., Brown, K. A., Ogutu, B. O., Smith, D., & Eigenbrod, F. (2020). A systematic map of research exploring the effect of greenspace on mental health. Landscape and Urban Planning, 201, 103823.

de Kluijver, H., & Stoter, J. (2003). Noise mapping and GIS: optimising quality and efficiency of noise effect studies. Computers, environment and urban systems, 27(1), 85-102.

Duzgoren-Aydin, N. S. (2007). Sources and characteristics of lead pollution in the urban environment of Guangzhou. Science of the Total Environment, 385(1-3), 182-195.

Davies, W. J., Adams, M. D., Bruce, N. S., Cain, R., Carlyle, A., Cusack, P., ... & Poxon, J. (2013). Perception of soundscapes: An interdisciplinary approach. Applied acoustics, 74(2), 224-231.

Deng, L., Luo, H., Ma, J., Huang, Z., Sun, L. X., Jiang, M. Y., ... & Li, X. (2020). Effects of integration between visual stimuli and auditory stimuli on restorative potential and aesthetic preference in urban green spaces. Urban Forestry & Urban Greening, 53, 126702.

Djenaihi, W. M., Zemmouri, N., Djenane, M., & Van Nes, A. (2021). Noise and Spatial Configuration in Biskra, Algeria—A Space Syntax Approach to Understand the Built Environment for Visually Impaired People. Sustainability, 13(19), 11009.

Del Campo, M., Carlson, A., & Manninger, S. (2021). Towards Hallucinating Machines-Designing with Computational Vision. International Journal of Architectural Computing, 19(1), 88-103.

Eigenfeldt, A., Bizzocchi, J., Thorogood, M., & Bizzocchi, J. (2015). Applying valence and arousal values to a unified video, music, and sound generative multimedia work. In Generative Art Conference (GA).

Engel, M. S., Hummell, B., Pfaffenbach, C., & Fels, J. (2019, September). Perceptual studies on Landscape and Soundscape in Aachen: Healthy Urban Planning Approach. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Vol. 259, No. 5, pp. 4595-4604). Institute of Noise Control Engineering.

Erfanian, M., Mitchell, A., Aletta, F., & Kang, J. (2021). Psychological well-being and demographic factors can mediate soundscape pleasantness and eventfulness: A large sample study. Journal of Environmental Psychology, 77, 101660.

Fan, J., Thorogood, M., Riecke, B. E., & Pasquier, P. (2015). Automatic recognition of eventfulness and pleasantness of soundscape. In Proceedings of the Audio Mostly 2015 on Interaction With Sound (pp. 1-6).

Fan, J., Thorogood, M., & Pasquier, P. (2016). Automatic soundscape affect recognition using a dimensional approach. Journal of the Audio Engineering Society, 64(9), 646-653.

Fan, J., Thorogood, M., & Pasquier, P. (2017, October). Emo-soundscapes: A dataset for soundscape emotion recognition. In 2017 Seventh international conference on affective computing and intelligent interaction (ACII) (pp. 196-201). IEEE.

Fiebig, A., Jordan, P., & Moshona, C. C. (2020). Assessments of acoustic environments by emotions–The application of emotion theory in soundscape. Frontiers in Psychology, 3261.

Facts & Figures | Michigan Medicine. (2021). University of Michigan Health. https://www.uofmhealth.org/about+umhs/facts-figures

Frescura, A., & Lee, P. J. (2022). Emotions and physiological responses elicited by neighbours sounds in wooden residential buildings. Building and Environment, 210, 108729.

Granö, J. G., & Granv, J. G. (1997). Pure geography. Johns Hopkins University Press.

Goel, N., & Etwaroo, G. R. (2006). Bright light, negative air ions and auditory stimuli produce rapid mood changes in a student population: a placebo-controlled study. Psychological medicine, 36(9), 1253-1263.

Gould van Praag, C. D., Garfinkel, S. N., Sparasci, O., Mees, A., Philippides, A. O., Ware, M., ... & Critchley, H. D. (2017). Mind-wandering and alterations to default mode network connectivity when listening to naturalistic versus artificial sounds. Scientific reports, 7(1), 1-12.

Gemmeke, J. F., Ellis, D. P., Freedman, D., Jansen, A., Lawrence, W., Moore, R. C., ... & Ritter, M. (2017, March). Audio set: An ontology and human-labeled dataset for audio events. In 2017 IEEE international conference on acoustics, speech and signal processing (ICASSP) (pp. 776-780). IEEE.

Giannakopoulos, T., Orfanidi, M., & Perantonis, S. (2019, January). Athens urban soundscape (athus): a dataset for urban soundscape quality recognition. In International Conference on Multimedia Modeling (pp. 338-348). Springer, Cham.

Greenberg, N. (2020). Mental health of health-care workers in the COVID-19 era. Nature Reviews Nephrology, 16(8), 425-426.

Herrington, L. P. (1974). Trees and acoustics in urban areas. Journal of Forestry, 72(8), 462-465.

Hetherington, J., Daniel, T. C., & Brown, T. C. (1993). Is motion more important than it sounds?: The medium of presentation in environment perception research. Journal of environmental psychology, 13(4), 283-291.

Hume, K. I., Barrett, H., Ip, A., McDonagh, T., Davies, W. J., Adams, M. D., ... & Plack, C. (2008). Physiological responses and subjective estimates of sounds: initial results of pilot study. Proc. I. o. A., 30(2).

Hume, K. I., Brink, M., & Basner, M. (2012). Effects of environmental noise on sleep. Noise and health, 14(61), 297.

Hume, K., & Ahtamad, M. (2013). Physiological responses to and subjective estimates of soundscape elements. Applied Acoustics, 74(2), 275-281.

Hall, D. A., Irwin, A., Edmondson-Jones, M., Phillips, S., & Poxon, J. E. (2013). An exploratory evaluation of perceptual, psychoacoustic and acoustical properties of urban soundscapes. Applied Acoustics, 74(2), 248-254.

Hong, J. Y., & Jeon, J. Y. (2014). Soundscape mapping in urban contexts using GIS techniques.

Hong, J. Y., & Jeon, J. Y. (2015). Influence of urban contexts on soundscape perceptions: A structural equation modeling approach. Landscape and Urban Planning, 141, 78-87.

Hanna-Attisha, M., LaChance, J., Sadler, R. C., & Champney Schnepp, A. (2016). Elevated blood lead levels in children associated with the Flint drinking water crisis: a spatial analysis of risk and public health response. American journal of public health, 106(2), 283-290.

Hong, J. Y. (2016, August). Relationship between soundscapes and landscape factors in urban commercial spaces. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Vol. 253, No. 3, pp. 5425-5428). Institute of Noise Control Engineering.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

Hedblom, M., Knez, I., Ode Sang, Å., & Gunnarsson, B. (2017). Evaluation of natural sounds in urban greenery: potential impact for urban nature preservation. Royal Society open science, 4(2), 170037.

Hedblom, M., Gunnarsson, B., Schaefer, M., Knez, I., Thorsson, P., & Lundström, J. N. (2019). Sounds of nature in the city: no evidence of bird song improving stress recovery. International journal of environmental research and public health, 16(8), 1390.

Hedblom, M., Gunnarsson, B., Iravani, B., Knez, I., Schaefer, M., Thorsson, P., & Lundström, J. N. (2019). Reduction of physiological stress by urban green space in a multisensory virtual experiment. Scientific reports, 9(1), 1-11.

Hong, X., Jiang, Y., Wu, S., Zhang, L., & Lan, S. (2019, July). Study on evaluation model of soundscape in urban park based on Radial Basis Function Neural Network: A case study of Shiba Park and Kamogawa Park, Japan. In IOP Conference Series: Earth and Environmental Science (Vol. 300, No. 3, p. 032036). IOP Publishing.

Hunter, M. R., Gillespie, B. W., & Chen, S. Y. P. (2019). Urban nature experiences reduce stress in the context of daily life based on salivary biomarkers. Frontiers in psychology, 10, 722.

Howard, J., & Gugger, S. (2020). Fastai: a layered API for deep learning. Information, 11(2), 108.

Irvine, K. N., Devine-Wright, P., Payne, S. R., Fuller, R. A., Painter, B., & Gaston, K. J. (2009). Green space, soundscape and urban sustainability: an interdisciplinary, empirical study. Local environment, 14(2), 155-172.

Irwin, A., Hall, D. A., Peters, A., & Plack, C. J. (2011). Listening to urban soundscapes: Physiological validity of perceptual dimensions. Psychophysiology, 48(2), 258-268.

Jahncke, H., Hygge, S., Halin, N., Green, A. M., & Dimberg, K. (2011). Open-plan office noise: Cognitive performance and restoration. Journal of Environmental Psychology, 31(4), 373-382.

Jennings, P., & Cain, R. (2013). A framework for improving urban soundscapes. Applied Acoustics, 74(2), 293-299.

Job, J. R., Myers, K., Naghshineh, K., & Gill, S. A. (2016). Uncovering spatial variation in acoustic environments using sound mapping. PloS one, 11(7), e0159883.

Jiang, L., & Kang, J. (2017). Perceived integrated impact of visual intrusion and noise of motorways: influential factors and impact indicators. Transportation Research Part D: Transport and Environment, 57, 217-223.

Keyel, A. C., Reed, S. E., McKenna, M. F., & Wittemyer, G. (2017). Modeling anthropogenic noise propagation using the Sound Mapping Tools ArcGIS toolbox. Environmental modelling & software, 97, 56-60.

Koprowska, K., Łaszkiewicz, E., Kronenberg, J., & Marcińczak, S. (2018). Subjective perception of noise exposure in relation to urban green space availability. Urban Forestry & Urban Greening, 31, 93-102.

Kaplan, R., & Kaplan, S. (1989). The experience of nature: A psychological perspective. Cambridge university press.

Kang, J., Aletta, F., Margaritis, E., & Yang, M. (2018). A model for implementing soundscape maps in smart cities. Noise Mapp 5 (1): 46–59.

Liu, J., Kang, J., Behm, H., & Luo, T. (2014). Effects of landscape on soundscape perception: Soundwalks in city parks. Landscape and urban planning, 123, 30-40.

Lundén, P., & Hurtig, M. (2016, August). On urban soundscape mapping: A computer can predict the outcome of soundscape assessments. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Vol. 253, No. 6, pp. 2017-2024). Institute of Noise Control Engineering.

Lopes, P., Liapis, A., & Yannakakis, G. N. (2017). Modelling affect for horror soundscapes. IEEE Transactions on Affective Computing, 10(2), 209-222.

Larkin, A., & Hystad, P. (2019). Evaluating street view exposure measures of visible green space for health research. Journal of exposure science & environmental epidemiology, 29(4), 447-456.

Liu, Y., Hu, M., & Zhao, B. (2019). Audio-visual interactive evaluation of the forest landscape based on eye-tracking experiments. Urban Forestry & Urban Greening, 46, 126476.

Li, H., & Lau, S. K. (2020). A review of audio-visual interaction on soundscape assessment in urban built environments. Applied acoustics, 166, 107372.

Lindquist, M., Maxim, B., Proctor, J., & Dolins, F. (2020). The effect of audio fidelity and virtual reality on the perception of virtual greenspace. Landscape and Urban Planning, 202, 103884.

Lei, Y., Davies, G. M., Jin, H., Tian, G., & Kim, G. (2021). Scale-dependent effects of urban greenspace on particulate matter air pollution. Urban Forestry & Urban Greening, 61, 127089.

Mennitt, D., Fristrup, K. M., Sherrill, K., & Nelson, L. (2013, September). Mapping sound pressure levels on continental scales using a geospatial sound model. In Proceedings of Inter-Noise (Vol. 4, pp. 1-11).

Marselle, M. R., Irvine, K. N., & Warber, S. L. (2013). Walking for well-being: are group walks in certain types of natural environments better for well-being than group walks in urban environments?. International journal of environmental research and public health, 10(11), 5603-5628.

Murphy, E., & King, E. A. (2014). Environmental noise pollution: Noise mapping, public health, and policy. Newnes.

Medvedev, O., Shepherd, D., & Hautus, M. J. (2015). The restorative potential of soundscapes: A physiological investigation. Applied Acoustics, 96, 20-26.

Marselle, M. R., Irvine, K. N., Lorenzo-Arribas, A., & Warber, S. L. (2015). Moving beyond green: exploring the relationship of environment type and indicators of perceived environmental quality on emotional well-being following group walks. International journal of environmental research and public health, 12(1), 106-130.

Maehr, A. M., Watts, G. R., Hanratty, J., & Talmi, D. (2015). Emotional response to images of wind turbines: A psychophysiological study of their visual impact on the landscape. Landscape and urban Planning, 142, 71-79.

Mushtaq, Z., Su, S. F., & Tran, Q. V. (2021). Spectral images based environmental sound classification using CNN with meaningful data augmentation. Applied Acoustics, 172, 107581.

Naomi, A. S. (2020). Access to nature has always been important; with COVID-19, it is essential.

Ntalampiras, S. (2020). Emotional quantification of soundscapes by learning between samples. Multimedia Tools and Applications, 79(41), 30387-30395.

Noviyanti, A., Sudarsono, A. S., & Kusumaningrum, D. (2019, December). Urban soundscape prediction based on acoustic ecology and MFCC parameters. In AIP Conference Proceedings (Vol. 2187, No. 1, p. 050005). AIP Publishing LLC.

Okba, B., Cutini, V., Leccese, F., Salvadori, G., & Zemmouri, N. (2021, August). Relation between soundscape and spatial configuration in different urban contexts. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Vol. 263, No. 5, pp. 1405-1414). Institute of Noise Control Engineering.

Pathak, V., Tripathi, B. D., & Mishra, V. K. (2011). Evaluation of anticipated performance index of some tree species for green belt development to mitigate traffic generated noise. Urban forestry & urban greening, 10(1), 61-66.

Piczak, K. J. (2015, September). Environmental sound classification with convolutional neural networks. In 2015 IEEE 25th international workshop on machine learning for signal processing (MLSP) (pp. 1-6). IEEE.

Pheasant, R. J., & Watts, G. R. (2015). Towards predicting wildness in the United Kingdom. Landscape and Urban Planning, 133, 87-97.

Peng, N., Chen, A., Zhou, G., Chen, W., Zhang, W., Liu, J., & Ding, F. (2020). Environment Sound Classification Based on Visual Multi-Feature Fusion and GRU-AWS. IEEE Access, 8, 191100-191114.

Russell, J. A. (1980). A circumplex model of affect. Journal of personality and social psychology, 39(6), 1161.

Radicchi, A. (2013). Emotional geography & soundscape studies: beyond the cognitive approach in (sound) mapping urban spaces. In Conference: the EAEA11 Envisioning Architecture: Design, Evaluation, Communication.

Richardson, E. A., Pearce, J., Mitchell, R., & Kingham, S. (2013). Role of physical activity in the relationship between urban green space and health. Public health, 127(4), 318-324.

Romero, V. P., Maffei, L., Brambilla, G., & Ciaburro, G. (2016). Modelling the soundscape quality of urban waterfronts by artificial neural networks. Applied Acoustics, 111, 121-128.

Roussel, J. R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R., Meador, A. S., ... & Achim, A. (2020). lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. Remote Sensing of Environment, 251, 112061.

Radicchi, A., & Grant, M. (2021). From noise to soundscape in the service of urban health. Cities & Health, 5(1-2), 15-19.

Ratcliffe, E. (2021). Sound and soundscape in restorative natural environments: A narrative literature review. Frontiers in Psychology, 12, 963.

Roussel, J. R. (2021). Airborne LiDAR Data Manipulation and Visualization for Forestry Applications [R package lidR version 3.0. 4].

Schwela, D. H. (2000). The World Health Organization guidelines for environmental noise. Noise News International, 8(1), 9-22.

Stansfeld, S., Hygge, S., Clark, C., & Alfred, T. (2010). Night time aircraft noise exposure and children's cognitive performance. Noise and Health, 12(49), 255.

Salamon, J., Jacoby, C., & Bello, J. P. (2014, November). A dataset and taxonomy for urban sound research. In Proceedings of the 22nd ACM international conference on Multimedia (pp. 1041-1044).

Stevens, F., Murphy, D. T., & Smith, S. L. (2017). Soundscape categorisation and the self-assessment manikin. In Proceedings of the 20th International Conference on Digital Audio Effects.

Sudarsono, A. S., Lam, Y. W., & Davies, W. J. (2017). The validation of acoustic environment simulator to determine the relationship between sound objects and soundscape. Acta Acustica united with Acustica, 103(4), 657-667.

Sanchez, G. M. E., Van Renterghem, T., Sun, K., De Coensel, B., & Botteldooren, D. (2017). Using Virtual Reality for assessing the role of noise in the audio-visual design of an urban public space. Landscape and Urban Planning, 167, 98-107.

Salem, T., Zhai, M., Workman, S., & Jacobs, N. (2018). A multimodal approach to mapping soundscapes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 2524-2527).

Silverman, B. W. (2018). Density estimation for statistics and data analysis. Routledge.

Smith, L. N. (2018). A disciplined approach to neural network hyper-parameters: Part 1--learning rate, batch size, momentum, and weight decay. arXiv preprint arXiv:1803.09820.

Su, Y., Zhang, K., Wang, J., & Madani, K. (2019). Environment sound classification using a two-stream CNN based on decision-level fusion. Sensors, 19(7), 1733.

Serebrennikova, A., Na, L., Qi, M., & Kang, J. (2019, September). A cross-cultural comparison in sound perception between Chinese and Russian students in urban open spaces in China.

In INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Vol. 259, No. 3, pp. 6591-6602). Institute of Noise Control Engineering.

Shanahan, D. F., Astell–Burt, T., Barber, E. A., Brymer, E., Cox, D. T., Dean, J., ... & Gaston, K. J. (2019). Nature–based interventions for improving health and wellbeing: The purpose, the people and the outcomes. Sports, 7(6), 141.

Shu, S., & Ma, H. (2020). Restorative effects of urban park soundscapes on children's psychophysiological stress. Applied Acoustics, 164, 107293.

Schulte-Fortkamp, B. (2021). Soundscape and its contribution to health in the city. Cities & Health, 5(1-2), 71-73.

Thorogood, M., & Pasquier, P. (2013, May). Impress: A Machine Learning Approach to Soundscape Affect Classification for a Music Performance Environment. In NIME (pp. 256-260).

Ulrich, R. S. (1986). Human responses to vegetation and landscapes. Landscape and urban planning, 13, 29-44.

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.

Uebel, K., Marselle, M., Dean, A. J., Rhodes, J. R., & Bonn, A. (2021). Urban green space soundscapes and their perceived restorativeness. People and Nature.

Van Renterghem, T., & Botteldooren, D. (2016). View on outdoor vegetation reduces noise annoyance for dwellers near busy roads. Landscape and urban planning, 148, 203-215.

Verma, D., Jana, A., & Ramamritham, K. (2019). Classification and mapping of sound sources in local urban streets through AudioSet data and Bayesian optimized Neural Networks. Noise Mapping, 6(1), 52-71.

Van Renterghem, T. (2019, September). Using vegetation to tackle environmental noise problems-combining exposure level reduction and noise perception improvement. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Vol. 259, No. 4, pp. 5095-5099). Institute of Noise Control Engineering.

Verma, D., Jana, A., & Ramamritham, K. (2020). Predicting human perception of the urban environment in a spatiotemporal urban setting using locally acquired street view images and audio clips. Building and Environment, 186, 107340.

Versümer, S., Steffens, J., Blättermann, P., & Becker-Schweitzer, J. (2020). Modeling evaluations of low-level sounds in everyday situations using linear machine learning for variable selection. Frontiers in psychology, 2593.

World Health Organization. (2010). Why urban health matters (No. WHO/WKC/WHD/2010.1). World Health Organization.

Wang, B., & Kang, J. (2011). Effects of urban morphology on the traffic noise distribution through noise mapping: A comparative study between UK and China. Applied Acoustics, 72(8), 556-568.

Weber, M. (2014). Assessing and improving the soundscape of urban parks. ICBEN2014, Nara, Japan.

Wolch, J. R., Byrne, J., & Newell, J. P. (2014). Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. Landscape and urban planning, 125, 234-244.

World Health Organization. (2018). Environmental noise guidelines for the European region.

White, M. P., Alcock, I., Grellier, J., Wheeler, B. W., Hartig, T., Warber, S. L., ... & Fleming, L. E. (2019). Spending at least 120 minutes a week in nature is associated with good health and wellbeing. Scientific reports, 9(1), 1-11.

Wang, R., & Zhao, J. (2020). Effects of evergreen trees on landscape preference and perceived restorativeness across seasons. Landscape Research, 45(5), 649-661.

Xie, Z., & Yan, J. (2008). Kernel density estimation of traffic accidents in a network space. Computers, environment and urban systems, 32(5), 396-406.

Xing, Z., Baik, E., Jiao, Y., Kulkarni, N., Li, C., Muralidhar, G., ... & Pouliot, C. (2017). Modeling of the latent embedding of music using deep neural network. arXiv preprint arXiv:1705.05229.

Yang, W., & Kang, J. (2005). Acoustic comfort evaluation in urban open public spaces. Applied acoustics, 66(2), 211-229.

Yamada, Y. (2006). Soundscape-based forest planning for recreational and therapeutic activities. Urban Forestry & Urban Greening, 5(3), 131-139.

Yu, L., & Kang, J. (2009). Modeling subjective evaluation of soundscape quality in urban open spaces: An artificial neural network approach. The Journal of the Acoustical Society of America, 126(3), 1163-1174.

Zhang, X., Zou, Y., & Shi, W. (2017, August). Dilated convolution neural network with LeakyReLU for environmental sound classification. In 2017 22nd International Conference on Digital Signal Processing (DSP) (pp. 1-5). IEEE.

Zhao, J., Xu, W., & Ye, L. (2018). Effects of auditory-visual combinations on perceived restorative potential of urban green space. Applied Acoustics, 141, 169-177.

Zhang, Y., & Kang, J. (2020, March). The development of emotional indicators for the soundscape in urban open public spaces. In IOP Conference Series: Materials Science and Engineering (Vol. 780, No. 5, p. 052006). IOP Publishing.

Zhu, H., Yang, F., Bao, Z., & Nan, X. (2021). A study on the impact of Visible Green Index and vegetation structures on brain wave change in residential landscape. Urban Forestry & Urban Greening, 64, 127299.