12-13-2016

Investigating Environmental Inequities in Terms of Street Greenery using Google Street View

Xiaojiang Li
University of Connecticut - Storrs, xiaojiang.li@uconn.edu

Investigating Environmental Inequities in Terms of Street Greenery using Google Street View

Xiaojiang Li, Ph. D.

University of Connecticut, [2016]

As an important component of the urban ecosystem, the urban greenery provides a series of benefits to urban residents and plays an important role in maintaining the urban sustainability. Unequal access to urban greenery represents environmental disparities when some urban residents are deprived of the benefits provided by urban greenery. As an important component of the urban greenery, the street greenery provides a series of benefits to urban residents, such as energy saving, provision of shade and aesthetic values. In addition, the street greenery is a kind of publicly financed amenity and the spatial distribution of the street greenery is influenced heavily by different policies. In this study, I studied the distribution of street greenery in dozens of major American cities and investigated whether racial/ethnic minorities and economically disadvantaged groups are living in neighborhoods with less street greenery. The modified green view index (MGVI), which literally represents the visibility of street greenery or how much street greenery people can see and feel on the ground, was used to represent the distribution of street greenery. The MGVI was calculated based on the publicly accessible Google Street View (GSV) images captured at different horizontal and vertical view angles. Tens of millions GSV images were downloaded for all the selected cities based on the static Google Street View images API to calculate MGVI in the study areas. The environmental inequity in terms of street greenery was further investigated by examining the relationships between the spatial distributions of residential street greenery and socioeconomic variables in different cities at census tract level. Results showed that people with various social conditions have different amounts of street greenery in their living environments in different cities. Generally, people with

higher incomes tend to live in places with more street greenery. The percentage of home ownership also plays a positive role in the spatial distribution of street greenery. In summary, this study contributes to literature by providing insights into the living environments of urban residents in terms of street greenery, and it generates a valuable reference data for future urban greening programs.

Keywords: Urban greenery, Google Street View, Environmental inequity, Street greenery.

Investigating Environmental Inequities in Terms of Street Greenery using Google Street View

Xiaojiang Li

B.S., Henan University, China, **[2010]**

M.S., University of Chinese Academy of Sciences, China, **[2013]**

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

[2016]

APPROVAL PAGE

Doctor of Philosophy Dissertation

Investigating Environmental Inequities in Terms of Street Greenery using Google Street View

Presented by

Xiaojiang Li, B.S., M.S.

Major Advisor _____
Chuanrong Zhang

Associate Advisor _____
Daniel Weiner

Associate Advisor _____
Robert Cromley

Associate Advisor _____
Daniel Civco

Associate Advisor _____
Debarchana Ghosh

University of Connecticut
2016

Acknowledgements

It is definitely not easy to get a Ph. D degree. I am lucky to finish my Ph. D program in Geography department at University of Connecticut in three and half years with the help of my family and friends. In the last three years, I worked hard and enjoyed the life in Connecticut.

On my way to Ph. D degree, I have many people to thank. I want to first thank my wife and my parents. My wife Guoqing is currently a Ph. D student in University of Massachusetts, Boston. She is the most important person for me. We applied to graduate school and came to United States together. I am very lucky to have the support from my wife all the time. In the last 3 years, we shared happiness and frustration. I cannot imagine what my life is going to be without my wife.

I also want to express my gratitude to my advisor, Prof. Chuanrong Zhang. She is really a very good advisor. She did all she can to help her students. She gave me enough freedom to do research independently. In the last three years, she revised my manuscripts and replied to my Emails 24/7. I will put this kind of passion on my students in future. I would also like to thank my committee members, Prof. Daniel Weiner, Prof. Robert Cromley, Prof. Daniel Civco, and Dr. Debarchana Ghosh.

I also want to thank other graduate students in UConn geography department. They are Weixing Zhang, Kate Johnson, Shuowei Zhang, Karen Johnson, Jose Torres, Qinglin Hu, and Stephanie Walker. In my first year in UCONN, Weixing drove me to campus every day. I really appreciate his help. There are many challenges in Connecticut, such as no public transportation,

wild cold winter, far away from my family and previous friends. With the help of my friends in UConn, all of these are not big problems anymore.

I would like give my special thanks to Michael Howser and other colleagues in MAGIC, Connecticut state data center. Michael is not just my boss but also my friend. He is very kind; he provides me a very good working environment and many opportunities. I think that probably the reason why all students who have ever worked in MAGIC are very successful.

At last, I would also want to thank my landlady Winnie and landlord Ed in Brooklyn, Connecticut. Winnie is a professor in Rhodes Island School of Design. I learned a lot in the conversation with her.

Xiaojiang Li,
Department of Geography,
University of Connecticut, Storrs

Table of Contents

Contents

List of Tables

List of Figures

Publications

Xiaojiang Li, Zhang, C, Li, W, Kuzovkina, Y, 2016, Environmental inequities in terms of different types of urban greenery in Hartford, Connecticut, USA. *Urban Forestry and Urban Greening*, 18, 163-172.

Zhang W., W. Li, C. Zhang, and **Xiaojiang Li**. Incorporating Spectral Similarity into Markov Chain Geostatistical Cosimulation for Reducing Smoothing Effect in Land Cover Post-Classification". *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, doi: doi: 10.1109/JSTARS.2016.2596040.

Xiaojiang Li, Weidong Li, Qingyan Meng, Chuanrong Zhang, Tamas Jancso & Kangli Wu (2016): Modeling building proximity to greenery in a three-dimensional perspective using multi-source remotely sensed data, *Journal of Spatial Science*, 1-15.

Xiaojiang Li, Zhang, C, Li, W, Kuzovkina, Y, Weiner, D. Who lives in greener neighborhoods? The distribution of street greenery and its association with residents' social conditions in Hartford, Connecticut, USA, *Urban Forestry and Urban Greening*, 2015, 14(4).

Xiaojiang Li, Zhang, C., & Li, W. (2015). Does the Visibility of Greenery Increase Perceived Safety in Urban Areas? Evidence from the Place Pulse 1.0 Dataset. *ISPRS International Journal of Geo-Information*, 4(3), 1166-1183.

Xiaojiang Li, Zhang, C., Li, W., Ricard, R., Meng, Q., Zhang, W. Assessing street-level urban greenery using Google Street View and a modified green view index, *Urban Forestry and Urban Greening*, 2015, 14(3).

Xiaojiang Li, Q. Meng, W. Li, C. Zhang, T. Jansco, and S. Mavromatis. 2014. "An explorative study on the proximity of buildings to green spaces in urban areas using remotely sensed imagery." *Annals of GIS* 20, 3 (2014): 193-203.

Xiaojiang Li, Qingyan Meng, Xingfa Gu, Tamas Jasco. A hybrid method combining pixel based and object based methods and its applications in Hungary using Chinese HJ-1 Satellite image, *International Journal of Remote Sensing*, 2013, 34(13), 4655-4667.

Xiaojiang Li, Qingyan Meng, Chunmei Wang, Miao Liu. A hybrid model of object- and pixel based classification of Remotely sensed data, *Geo-spatial Information Science*, 2013, 15(5) (In Chinese).

Chapter 1 Introduction

1.1. Urban greenery

Urban areas are the places of mass interactions between human and the natural world. Urban areas are also homes to a large proportion of the world population. In 1900, only 10% of the global population lived in urban areas, the percentage exceeds 50% in 2008 (Grimm et al., 2008). The number of urban population is still increasing rapidly. The increasing population and the spatial prominence of urban areas make them an important focus of many studies (Pickett et al., 2011). Human beings are increasingly living in urban areas, while continuing to depend on the natural world for survival (Bolund and Hunhammar, 1999). Urban ecosystems serve as a foundation for human's survival in cities.

Based on definition of Konijnendijk et al (2006), urban greenery mainly includes urban parks, woodlands, street and square trees, lawns, and other kinds of vegetation. Urban greenery is an important part of urban ecosystem and has long been recognized for their importance in urban environment (Li et al, 2015a). Urban greenery provides a lot of environmental, economic, social,

and health benefits (Chen et al, 2006; Jim and Chen, 2008; Onishi et al, 2010; Miller, 1997; Gidlow et al, 2012; van Dillen et al, 2012; Wendel et al, 2011) meeting diverse and overlapping goals (Bain et al., 2012). The existing of urban greenery is regarded as an important environmental amenity (Dwivedi et al., 2009; Nichol and Wong, 2005; Seymour et al., 2010).

However, there are only a few of studies have examined the distribution of street greenery in residential areas. As one component of urban greenery, residential street greenery makes an important contribution to the attractiveness and walkability of streets (Schroeder and Cannon, 1983; Wolf, 2005; Bain et al., 2012). Street trees growing on Rights-of-Way provide a range of health benefits by promoting outdoor exercises (Wolch et al., 2005; Takano et al., 2002). Street greenery also provides sensory functions addressing the visual effects of greenery. It can mitigate the visual intrusions of vehicular traffics, and contribute to the beauty of cityscapes. This dissertation mainly focuses on the spatial distribution of residential street greenery in different cities of United States.

1.2. Environmental injustice and urban greenery

Uneven distribution of environmental amenities and disamenities in cities lead to unequal distribution of social benefits and burdens across people and places (Landry and Chakraborty, 2009). Since 1980s, with growth of U.S environmental justice movements, urban environmental injustice has received considerable attentions in urban studies. Initially, the environmental justice studies focused majorly on disproportionate exposure to environmental burdens (include locally unwanted land uses, air pollution, and hazardous waste risks) of racial/ethnic minorities and

economically disadvantaged groups (Been, 1994; Liu, 1997; Mohai and Bryant, 1992; Brainard et al, 2002). Recently, the accident of toxic tap water in Flint, Michigan reminds us the environmental injustice is still a serious problem in United States. Proliferating results show that the minorities are living in neighborhoods with more hazardous waste risks and air pollution. A study conducted by Clark et al (2014) at national scale shows that people of color are disproportionately hurt by air pollution in United States. Based on the National Land Cover Dataset, Jesdale et al. (2013) examined the distribution of heat risk–related land cover across racial/ethnic groups at the national scale and found that racial/ethnic minorities tend to live in neighborhoods with higher heat risk-related land cover.

The environmental justice studies were later broadened to include the environmental goods or amenities, like urban parks (Boone et al, 2009; Dai, 2011; Wolch et al, 2005), recreational facilities (Hewko et al, 2002; Wells et al, 2008). As an important kind of environmental amenities and an important element of urban socio-ecosystem (Nichol and Wong, 2005; Dwivedi et al., 2009; Seymour et al., 2010), the uneven distribution of urban greenery or unequal accessibility of urban greenery represents environmental disparities when some urban residents are deprived of the benefits that the urban greenery provides. Greenery on private lands usually results from natural colonization and private investments, however, trees on public Right-of-Way areas are majorly planted and maintained by public agencies (Landry and Chakraborty, 2009). Therefore, the spatial distribution of the publicly financed street greenery is more affected by the public investment, and the uneven distribution of the street greenery may reflect the potential environmental injustice or environmental racism.

However, there are still a few studies about the uneven distribution of street greenery in United States (Landry and Chakraborty, 2009; Li et al., 2015b). In this study, I studied the distributions of residential street greenery in dozens of major American cities. This study also investigated whether the racial/ethnic minorities and the economically disadvantaged groups are living in neighborhoods with the same amount of street greenery as other groups.

1.3. The problem of green metrics based on overhead view data

Remote sensing seems to be one of the most commonly used objective methods for measuring the distribution of urban greenery (Gupta et al. 2012). Metrics derived from remotely sensed data have been widely used as indicators of the spatial distribution of urban greenery and the proxies for environmental conditions (Pearsall and Christman, 2012). This is probably due to a number of virtues (e.g., repeatability, synoptic view, and larger area coverage) of remotely sensed data. Percentage of vegetation cover, green space/built area ratio, green space density and other measures, have been calculated for analyzing, assessing, and visualizing urban greenery (Ruagrit and sokhi 2004; Faryadi and Taheri 2009; Li et al., 2014; Li et al., 2016a; Zhu et al. 2003).

The aesthetic attractiveness of a neighborhood is greatly influenced by the amount of greenery that can be visually and aesthetically enjoyed. Studies have shown that urban greenery with more visible vegetation can obtain stronger public support than that with less visible vegetation, even though they may have the same coverage (Yang et al., 2009). The visibility of greenery helps to increase the satisfaction of citizens to their residential environments and plays

an important role in comforting citizens. While green indices derived from remotely sensed data may be good for quantifying urban greenery, they are not suitable for assessing profile views of street greenery. The profile view of street greenery that people can see on the ground is different from the overhead view captured by remotely sensed data (Li et al., 2015a). By using Stand Visualization System developed by the USDA Forest Service, Yang et al. (2009) showed that two urban forests with the same canopy cover look completely different in their profile views. **Fig. 1.1** (a) shows the profile view of a green wall on ground. When viewing a green wall from above using remotely sensed data, the wall is missed. In addition, an overhead view from remotely sensed imagery may miss the shrubs and lawns under tree canopies in case of a multi-layer green space (**Fig. 1.1** (b)). Therefore, while aerial/space remotely sensed imagery might provide useful information for measuring urban greenery, it usually fails to acquire what people actually see at street-level on ground.

(a) Profile view of a green wall (b) Profile view of a multi-layer green space

Fig.1.1. Profile views of different types of green spaces: (a) profile view of a green wall, (b) profile view of a multi-layer green space (Li *et al*, 2015a).

Previous studies show that there is little agreement between human perceived greenness and the objectively derived greenness from remotely sensed images (Leslie et al., 2010). However, it is the perceived greenness has a direct connection with the benefits provided by street greenery**.** Leslie et al. (2010) criticized the mixed findings on the relationship between neighborhood greenness and physical activities in terms of the discrepancy between perceived and objectively measured greenness. The greenness indicators based on remotely sensed images or aerial photographs may not fully represent the neighborhood greenness, especially the greenness perceived by residents. The green metrics derived from remotely sensed data may not fully represent neighborhood greenness (Li et al., 2015b).

In addition, for studying fine scale street greenery, high spatial resolution remotely sensed data are usually required (Li et al., 2014; Pham et al., 2012, 2013; Landry and Chakraborty, 2009; Zhou and Troy, 2008). However, in real applications, high-resolution remotely sensed datasets are not always available and are expensive to collect. New and cheap data sources are in need for the street greenery study.

1.4. Street level images for urban greenery study

To date it has been difficult to efficiently and accurately represent and quantify street greenery. Using color photographs or slides as surrogates for the natural environment has been chosen as a cost-effective method for evaluating urban greenery (Yao et al., 2012; Meitner, 2004; Stamps, 1990). This method had been validated by various independent studies (Daniel and Boster, 1976; Shuttleworth, 1980; Stamps, 1990). Using street view images to map the amount of

street greenery represents a new and promising approach (Yang 2009; Li et al. 2015a, b). Recently, Yang et al. (2009) used color pictures to evaluate the visibility of surrounding urban forests as representative of pedestrians' view of greenery through developing a Green View Index (GVI). Four pictures were taken in four directions (north, south, east, and west) for each street intersection in their study area. Those pictures were processed to extract green areas, which were further used for calculating their proposed GVI. A drawback, however, is that the data collection and processing processes in their study are tedious and time consuming because the whole workflow (from collection of the pictures to extraction of the green areas) was conducted manually. This limits the application of the GVI to only a very small area. In addition, people's perception to surrounding environment is influenced by hemispherical scene (Asgarzadeh et al., 2012; Bishop, 1996). The method proposed by Yang et al. (2009) has limitations in measuring the amount of visible greenery because only pictures in four directions were used to calculate the GVI, which cannot cover the spherical view field of an observer.

To overcome those limitations of overhead view datasets, this study proposed to use Google Street View (GSV) images (which have view angles similar to those of pedestrians and open access) for assessing the street greenery. GSV, first introduced in 2007, is a library of video footage captured by cars driven down streets (Rundle et al., 2010). By stitching the pictures together, GSV images can create a continuous 360-degree image of a streetscape. GSV creates what feels like a seamless tour of city streets and it is quite similar to what you or we can see when exploring a city by cars, bikes, or foot.

GSV images have been proposed as an effective potential data source for urban studies

(Rundle et al., 2010). These applications include: identification of commercial entities (Zamir et al., 2011), 3D city modeling (Torii et al., 2009; Mičušík and Košecká, 2009), public open space audit (Edwards et al., 2013; Taylor et al., 2011), and neighborhood environmental audit (Charreire et al., 2014; Rundle et al., 2011; Odgers et al., 2012; Griew et al., 2013).

Different from previous studies using the canopy cover or vegetation indices (Grove et al., 2006; Mennis, 2006; Landry and Chakraborty, 2009; Leslie et al., 2010; Jenerette et al., 2013), this study used the GSV-based modified green view index (MGVI) to represent the distribution of greenness in residential areas. The MGVI is the averaged value of the area proportions of green vegetation in street-level images captured in eighteen different directions. It quantitatively represents how much greenery a pedestrian can see from ground level (Li et al., 2015a).

GSV covers streetscapes of most of American cities, and it provides a new tool for national scale street greenery study. In this study, the GSV based MGVI was applied to map the spatial distribution of street greenery for dozens of major cities in United States. In order to investigate the environmental inequities in terms of street greenery in residential areas, the MGVI maps of different cities were further compared with social variables, which were derived from American Census Survey (ACS) data.

1.5. The highlights and dissertation outline

The highlights of the study include:

1. This study first investigated the national scale street greenery in major American cities using Google Street View (GSV) data. The GSV data is publicly accessible and has global

coverage. This study presents an example to utilize the GSV for large-scale urban studies.

2. Different from traditional green metrics derived from remotely sensed imagery with overhead view or zoning data, this study used the GSV based MGVI for measuring the street greenery. The GSV based MGVI covers the profile view of street greenery, which would be more suitable to represent the visibility of street greenery.

3. This study then investigated the environmental inequities in terms of street greenery in dozens of major cities in different regions of United States. The results of this study could provide a reference for future urban greening projects to narrow the gap of environmental qualities of different neighborhoods with different social groups.

The outline of the dissertation is listed as following:

Chapter 2 reviews the environmental, economic, social, and esthetic benefits provided by urban greenery. Since this study focuses on the street greenery, the benefits of street greenery were specified. Previous studies about environmental inequities in terms of urban greenery were also reviewed in this section.

Chapter 3 describes the chosen 26 major cities and the data sources used in this study. Ten major cities, which located in different regions of U. S., were finally selected for the further analysis.

Chapter 4 introduces the methods for geographical sampling, Google Street View static images collection, static images processing, and calculation of the MGVI. This chapter also analyzes the spectral signatures of green vegetation in different seasons for different cities located in different climate zones.

Chapter 5 compares the MGVI with traditional derived green metrics derived from remotely sensed data. Hartford and Boston were selected as the case study area to compare the MGVI and traditional green metrics. The relations and discrepancies between these metrics were analyzed.

Chapter 6 analyzes the environmental inequities in terms of street greenery in ten major cities of United States. Bivariate correlation analysis and regression models were deployed to analyze the relationships between street greenery index and social variables derived from census data.

Chapter 7 comes to the conclusions and potential directions for future study.

Chapter 2 Literature Review

2.1 Urban greenery and its benefits

As an integral part of urban ecosystem, urban greenery is important for the stability and sustainability of urban ecosystem. Urban greenery can absorb carbon dioxide and release oxygen (Hough, 1984), which is very important for urban dwellers. The existing of urban greenery helps to regulate urban microclimate and mitigate the urban heat island (Chen et al., 2006; Onishi et al., 2010). In hot summers, the tree canopies can block the sunshine from radiating the ground directly and provide shades for pedestrians. Urban trees help to filter airborne pollutants and particulates from the air (Lawrence, 1995; Jim et al., 2008). The filtering capability increases with availability of the more leaf areas (Givoni, 1991). The urban greenery helps to attenuate storm-water runoff and reduce surface water runoff, which will then prevent the floods after heavy rainfalls (Zhang et al., 2012; Liu et al., 2014). This is extremely important for densely urban areas, where are paved with impervious surfaces. The attenuated water runoff can further help to reduce the soil erosion. In addition, the root system of greenery helps to hold soil in the place and keeps sediments out of lakes, streams, both of which help to decrease the possibility of

urban floods and protect the water quality.

In terms of the economic benefits, the existing of urban greenery helps to increase the value of properties (Jim and Chen, 2006; Kong and Nakagoshi, 2007; Mansfield et al, 2005). Urban forest in central business areas positively affects people's judgment of visual quality and further significantly influences consumer responses and behaviors (Wolf, 2005). Compared with stores along streets with no street greenery, those stores along streets with more street greenery could attract more customers (Wolf, 2005). In addition, the existing of urban greenery could help to decrease the energy consumption in summer by providing shades and mitigating urban heat island effects (Akbari et al, 2001).

About the social benefits, urban greenery offers urban residents opportunities for recreations, physical exercises (e.g., walking and bicycling), and social activities (Zhou and Kim, 2013; Maas et al, 2006; Ellaway et al, 2005; Dai, 2011; Wolch et al, 2011). All of these further benefit human mental health (Leslie et al., 2010; Lee and Maheswaran, 2011; Bain et al., 2012; Coutts, 2008) and reduce aggression and crimes (Kuo and Sullivan, 2001; Troy et al., 2012; Wolfe and Mennis, 2012). However, the role of urban greenery in reducing aggression and crimes appears ambiguous. The urban greenery may provide hiding places for potential criminal activities (Fisher and Nasar, 1992; Nasar et al., 1993). In addition, shrubs could obstruct the "eyes on street", which could further facilitate the potential crime activities. Evidences from Donovan and Prestemon (2012) shows that low trees obstructing views from first floor windows on private lots are associated with increased crime occurrences.

The urban greenery also brings some health benefits to urban residents (Gidlow et al., 2012;

van Dillen et al., 2012; Wendel et al., 2011). People's accessibility to views of greenery seems to influence their recovery from surgery, increase restorative potential, and improve psychological wellbeing (Ulrich, 1984; Pazhouhanfara and Kamal, 2014; Kaplan, 2001). Urban greenery also provides a range of health benefits by promoting outdoor exercises (Wolch et al, 2005; Takano et al, 2002; Arbogast, 2009).

In addition, urban greenery also adds to the aesthetics of urban areas. There are evidences that living in a greener environment makes people feel closer to nature. Urban street greenery makes an important contribution to the attractiveness and walkability of residential streets (Schroeder and Cannon, 1983; Wolf, 2005; Bain et al., 2012). Street greenery also provides aesthetic benefits by mitigating visual intrusions of traffics in densely populated urban areas (Li et al, 2015). The existence of urban greenery usually increases people's aesthetic rating of urban scenes (Camacho-Cervantes et al., 2014; Balram and Dragićević, 2005).

Different types of urban greenery play different roles in providing benefits (Li et al., 2016b). As a public facility, urban parks are important for the quality of life in cities. Urban parks provide public places for recreations, physical exercises, and social activities, which can promote both personal health and social cohesion within communities (Zhou and Kim, 2013; Maas et al., 2006; Ellaway et al., 2005; Dai, 2011; Wolch et al., 2011). The private backyard vegetation is usually managed by private owners and not directly accessible to other people (Lachowycz and Jones, 2013; Li et al, 2014). View of backyard greenery through window is helpful to increasing restorative potentials and improving psychological wellbeing (Ulrich, 1984; Pazhouhanfar and Kamal, 2014; Kaplan, 2001). In addition, residential tree canopy cover reduces cooling energy

use in summer by providing shades (Akbari et al, 2001). Street greenery makes an important contribution to the attractiveness and walkability of residential streets (Schroeder and Cannon, 1983; Wolf, 2005; Bain et al., 2012; Lachowycz and Jones, 2013; Li et al., 2014). Planting street trees may provide more benefits to urban residents than planting trees in parks and private backyards (Kardan et al., 2015). Street greenery also provides a range of health benefits by promoting outdoor exercise (Wolch et al., 2005; Takano et al., 2002) and beautifies a neighborhood while mitigating the visual intrusion of traffics (Li et al., 2015a).

However, urban greenery is not always an environmental amenity. It increases the budget for cleaning the dead leaves and branches. The roots of the street trees could break the road conditions along the streets, especially the walkways. This study assumes the urban greenery as kind of amenity in urban areas.

2.2 Environmental inequities in terms of urban greenery

Previous studies have reported the environmental inequities in terms of urban greenery in North American cities (Heynen et al, 2006; Boone et al., 2009; Zhou and Kim, 2013; Dai, 2011; Pham et al., 2012; Landry and Chakraborty, 2009; Li et al., 2015b; Li et al., 2016b). Current environmental inequity studies mainly focus on the uneven distribution of vegetation coverage/indices (Pham et al., 2011, 2012, 2013; Jennings et al., 2012; Zhou and Kim, 2013; Jesdale et al., 2013; Leslie et al., 2010; Landry and Chakraborty., 2009; Jensen et al., 2004) and visiting distances to green spaces (Zhou and Kim, 2013; Boone et al., 2009; Leslie et al., 2010; Lotfi and Koohsari, 2011). Proliferating evidences show that racial/ethnic minorities,

low-income people, and underprivileged populations are living in neighborhoods with disproportionally less vegetation coverage than affluent groups across North American cities (Pham et al., 2012; Jesdale et al., 2013; Zhou and Kim, 2013). Heynen et al. (2006) found a negative relationship between urban-forest canopy cover and proportion of Hispanics in Milwaukee, Wisconsin, but a positive relationship between urban-forest canopy cover and proportion of non-Hispanic Whites. Jensen et al. (2004) found a positive relationship between leaf area index and median household income in Terre Haute, Indiana. Jesdale et al. (2013) investigated the distribution of heat risk–related land cover across racial/ethnic groups at the national scale, and found vegetation coverage disparities among different racial/ethnic groups in the United States. However, the findings on disparities of accessibility to green spaces are not consistent in different cities. Boone et al. (2009) examined the distribution proximity to parks in Baltimore, Maryland, and found that a higher proportion of African Americans have access to parks within walking distances than do other groups, while Whites have access to a larger acreage of parks. Dai (2011) evaluated the disparities in green space access in Atlanta, Georgia across different racial/ethnic and socio-economic groups, and found neighborhoods with higher proportions of African Americans have significantly less access to green spaces**.** Zhou and Kim (2013) studied the accessibility to urban parks in six Illinois cities based on the Google Map application programming interface. Results show that there is no significant accessibility difference among different racial/ethnic groups in those six cities.

Different types of urban greenery play different roles in providing benefits to urban residents (Li et al., 2016b). As an important component of urban greenery, street greenery makes streets

more beautiful and walkable. The street greenery is an important publicly financed amenity and would be affected by the different policies. Landry and Chakraborty (2009) investigated the environmental equity implication of public right-of-way trees in Tampa, Florida, USA. Results show that neighborhoods with higher proportions of African Americans, low-income residents have significantly lower proportions of public right-of-way tree cover. Different from previous studies using the canopy cover or vegetation indices (Grove et al., 2006; Mennis, 2006; Landry and Chakraborty, 2009; Leslie et al., 2010; Jenerette et al., 2013), Li et al (2015b) used a green view index to represent the distribution of street greenery in residential areas and checked whether or not the minorities and economically disadvantaged groups live in places with less street greenery in Hartford, Connecticut. The green view index is calculated based on street-level images, and it quantifies how much street greenery people can see and feel on the ground (Li et al, 2015a). Results show that people with various social conditions have different amounts of street greenery in their living environments in Hartford. Those people with higher incomes live in neighborhoods with more street greenery and there exists no significant relationship between the street greenery and racial/ethnic variables.

Chapter 3 Study areas and Data Sources

3.1 Study areas

Most of the chosen cities (**Table 3.1.1**) are large cities in United States. In addition, some other cities, which have environmental inequities reported previously, were also included in the list. Different climate zones could affect the cost of maintaining urban greenery and the cost of urban greening projects. In addition, the phenology of urban greenery in different climate zones varies, which may further affect the benefits it provides to urban residents. Therefore, the corresponding climate zones were also included in this study based on previous studies (Kottek et al., 2006; Yang et al., 2015) (**Fig. 3.1**). More details about the phenology of urban greenery are showed in Chapter 4.

Table 3.1.1

The climate zones of the chosen cities, the climate zone is adopted from Kottek et al (2006) and Yang et al (2015).

Climate	Cities
Humid continental (warm summer)	Boston, Philadelphia, Pittsburgh, Chicago, Detroit, Indianapolis, Cincinnati, Kansas City, St. Louis,

	Louisville
Humid continental (cool summer)	Minneapolis, Milwaukee
Humid subtropical	Washington D.C, Baltimore, Tampa, Dallas, Houston, San Antonio, Memphis
Mediterranean	Los Angeles, San Francisco
Tropical (wet/dry) season	Miami
Marine west coast	Seattle
Highland (alpine) climate	Denver
Midattitude desert	Phoenix

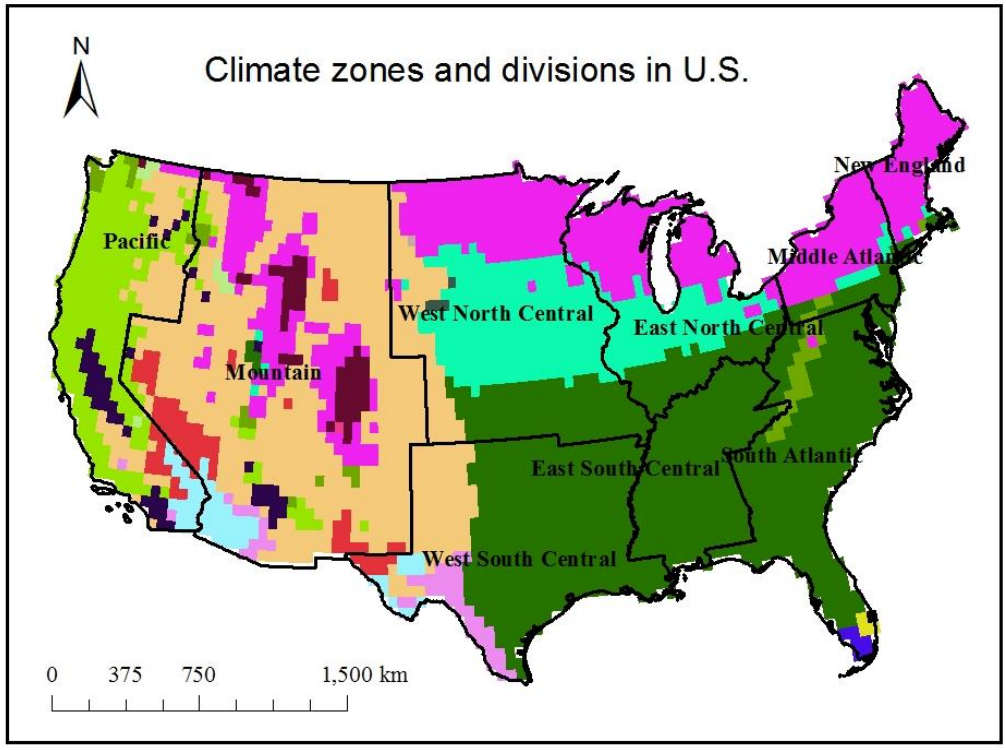


Fig.3.1. The climate zones and divisions in U.S. (modified from Kottek et al., 2006).

Table 3.1.2 lists the chosen cities and the corresponding divisions in United States, respectively. **Fig. 3.2** shows the locations of different cities in different regions and divisions of United States. The Northeastern region, which consists of ten states – Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont, New York, Pennsylvania, New Jersey and Delaware, is the nation's most economically developed and densely populated region. Three of the most populous cities (Philadelphia, Boston, and Pittsburgh) in Northeastern region were chosen in this study. As the most populous city in Northeast, New York City was not included in this study considering the different urban forms and the very different urban structures in New York City compared with other cities.

According to geographic region definition of US census bureau, the Midwest includes 12 states: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin. The major cities (Chicago, IL, Indianapolis, IN, Detroit, MI, Milwaukee, WI, Minneapolis, MN, Kansas City, MO, St. Louis, MO, and Cincinnati, OH) in this region were chosen in this study.

Several major cities from the South were also included in this study. These cities are Baltimore, MD, Washington D.C, Tampa, FL, Miami, FL, Memphis, TN, Atlanta, GA, Louisville, KY, Dallas, TX, Houston, TX, and San Antonio, TX.

The West is the largest and most geographically diverse region of United States. The West is split into two sub-regions, Pacific States (Washington, Oregon, California, Alaska, and Hawaii) and Mountain States (Montana, Wyoming, Colorado, New Mexico, Idaho, Utah, Arizona, and

Nevada). The major cities of this region chosen in this study include three cities from Pacific States (Los Angeles, CA, San Francisco, CA, and Seattle, WA) and two cities from the Mountain States (Phoenix, AZ and Denver, CO).

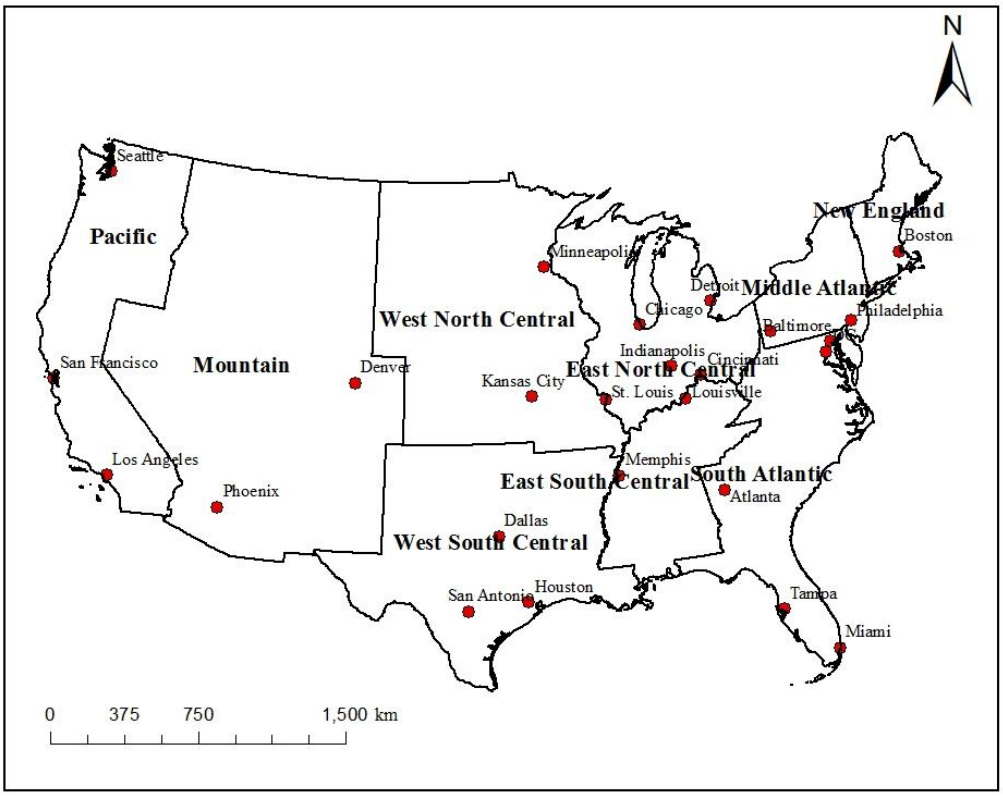


Fig.3.2 The selected cities from different U.S. census Bureau Regions and Divisions**.**

Table 3.1.2

The chosen major cities and their divisions in United States.

Cities	**States**	**Divisions**	**Regions**
Boston	Massachusetts	Division 1: New England	Northeast
Philadelphia	Pennsylvania	Division 2: Mid-Atlantic	
Pittsburgh	Pennsylvania		
Detroit	Michigan	Division 3: East North Central	Midwest
Chicago	Illinois		

Milwaukee	Wisconsin		
Indianapolis	Indiana		
Kansas City	Kansas		
Minneapolis	Minnesota		
St. Louis	Missouri		
Cincinnati	Ohio		
Tampa	Florida	Division 4: West North Central	South
Baltimore	Maryland	Division 5: South Atlantic	
D. C	Washington D.C.		
Miami	Florida		
Memphis	Tennessee	Division 6: East South Central	
Atlanta	Georgia		
Louisville	Kentucky		
Houston	Texas	Division 7: West South Central	
San Antonio	Texas		
Dallas	Texas		
Phoenix	Arizona	Division 8: Mountain	West
Denver	Colorado		
Los Angeles	California	Division 9: Pacific	
San Francisco	California		
Seattle	Washington		

3.2 Data sources

The major data sources in this study include administrative boundary maps, road maps, and land use or zoning maps. The administrative boundary and road maps for all cities were obtained from U.S. Census Bureau TIGER (Topologically Integrated Geographic Encoding and Referencing) products (https://www.census.gov/geo/maps-data/data/tiger.html). Because this study focuses on residential street greenery, only the residential streets were considered in this study. Other roads, like interstate high ways and state high ways, were removed from further

analysis.

Land use maps and zoning maps were used to delineate the residential areas for different cities. Land use maps and zoning maps for different cities were downloaded or requested from different geospatial data portal websites or municipal departments. **Table 3.1.3** shows the summary of the data sources of land use maps or zoning maps for all selected cities in this study.

Table 3.1.3

The data sources of land use map or zoning map of the chosen cities.

Cities	**Data types/year**	**Data sources**
Boston	Land use map, 2005	http://www.mass.gov/anf/research-and-tech/it-serv-and-support/application-serv/office-of-geographic-information-massgis/datalayers/lus2005.html
Philadelphia	Land use map, 2014	https://www.opendataphilly.org/dataset/land-use
Pittsburgh	Land use map, 2010	http://Pittsburghhpa.gov/dcp/gis/gis-data-new
Detroit	Zoning map, 2015	http://portal.datadrivendetroit.org/datasets/bbe90203edcd4c51af2f6c697ab5216c_0
Chicago	Zoning map, 2012	https://data.cityofchicago.org/
Milwaukee	Land use map, 2010	Requested from Milwaukee County Land Information Office
Indianapolis	Land use map, 2014	http://data1.indygis.opendata.arcgis.com/datasets/97398cf4efaa4556b0c57356d5818b76_3
Kansas City	Land use map, 2012	http://maps.kcmo.org/apps/parcelviewer/
Minneapolis	Primary zoning areas, 2015	http://opendata.minneapolismn.gov/datasets/eac15cee3f2d4ec4887e1f8995955ef1_0
St. Louis	Land use map, 2015	http://dynamic.stlouis-mo.gov/citydata/downloads/
Cincinnati		Not available
Tampa	Zoning map, 2015	http://city.tampa.opendata.arcgis.com/
Baltimore	Land use, 2008	https://data.baltimorecity.gov/
D. C	Land use map, 2010	http://opendata.dc.gov/
Miami		Not available
Memphis		Not available
Louisville	Land use map	https://portal.louisvilleky.gov/dataset/landuse-data
Houston	Land use map, 2014	http://data.ohouston.org/dataset/harris-county-land-use

San Antonio	Land use map, 2015	http://www.sanantonio.gov/GIS/GISData.aspx
Dallas	Land use map	http://www.dallascad.org/GISDataProducts.aspx, http://gis.dallascityhall.com/homepage/shapezip.htm
Phoenix	Zoning data, 2015	http://maps.phoenix.opendata.arcgis.com/datasets/d438c29d14ef407593279041e42fc015_0
Denver	Zoning data	http://www.denvergov.org/maps/map/zoning
Los Angeles	Zoning data, 2009	http://egis3.lacounty.gov/dataportal/2012/04/10/countywide-zoning/
San Francisco	Zoning data, 2012	https://data.sfgov.org/Geographic-Locations-and-Boundaries/Zoning-Districts/mici-sct2
Seattle		Not available

Chapter 4 Utilizing Google Street View

Google Street View (GSV), first introduced in 2007, is a free online service featured in Google Maps and Google Earth that provides panoramic views from positions along streets in the world (**Fig. 4.1**). It is a library of video footage captured by cars (**Fig. 4.1**) driven down streets (Rundle et al., 2010). GSV creates what feels like a seamless (if pixelated) tour of city streets and it can give one the feeling of "being there" (Li et al., 2015a). It is quite similar to what people on ground can see when exploring a city by cars, bikes, or foot. GSV panoramas are generated by stitching the pictures taken in different directions together (**Fig. 4.2**). GSV images can create a continuous 360-degree image of a streetscape.

GSV, which has a similar view angle with people on the ground, could be a very suitable dataset to study the urban environment. In this study, I developed a workflow to utilize static GSV images to measure the spatial distribution of street greenery. The workflow consists of four major steps: geographical sampling, GSV images collection, GSV images processing, and model developing for urban greenery assessment.

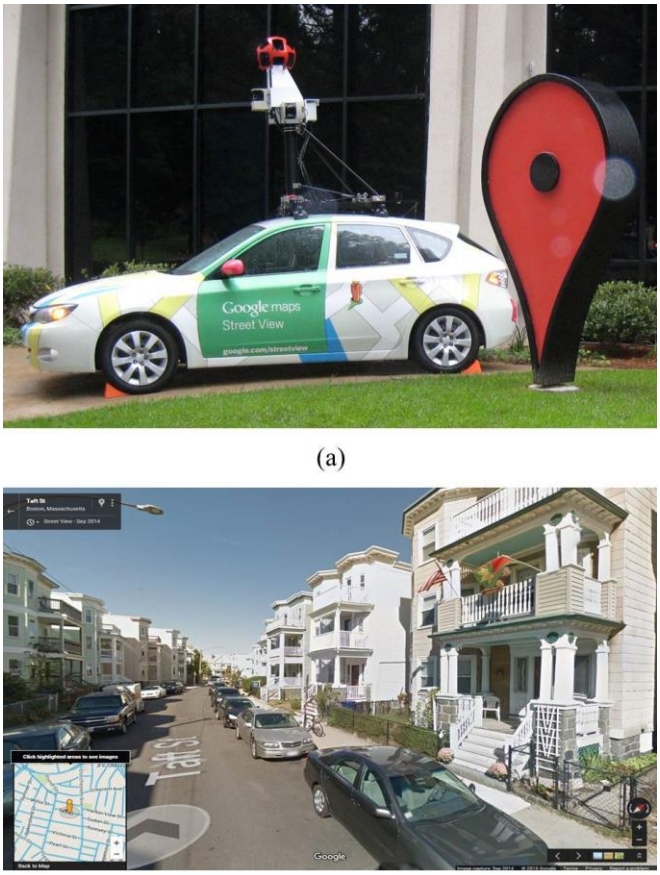


Fig.4.1. A GSV car and a snapshot of Google Maps.

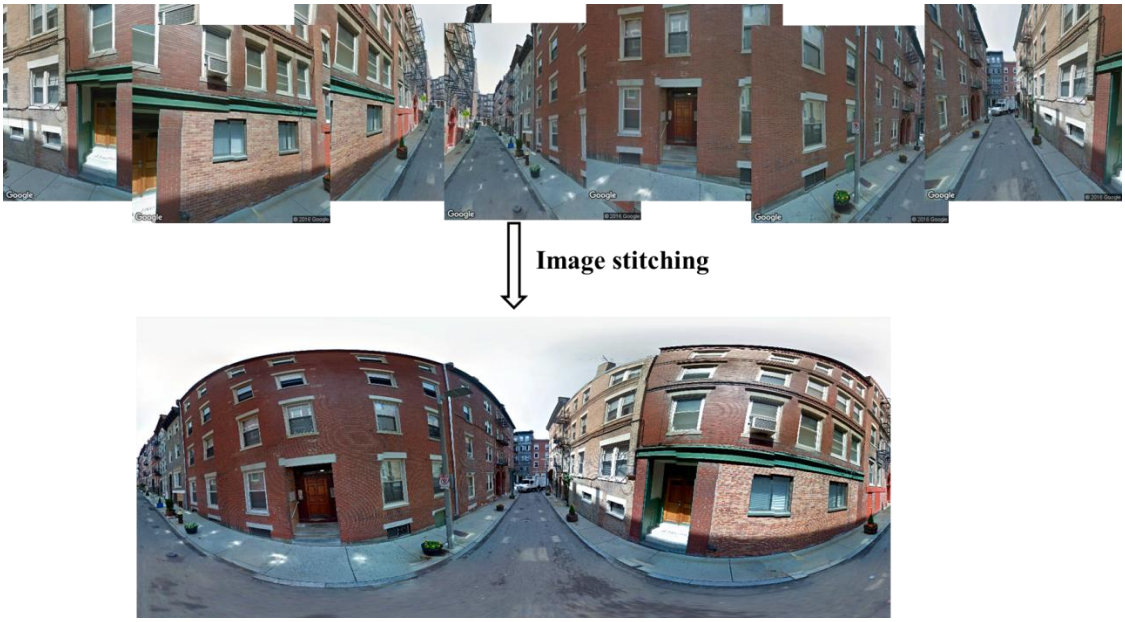
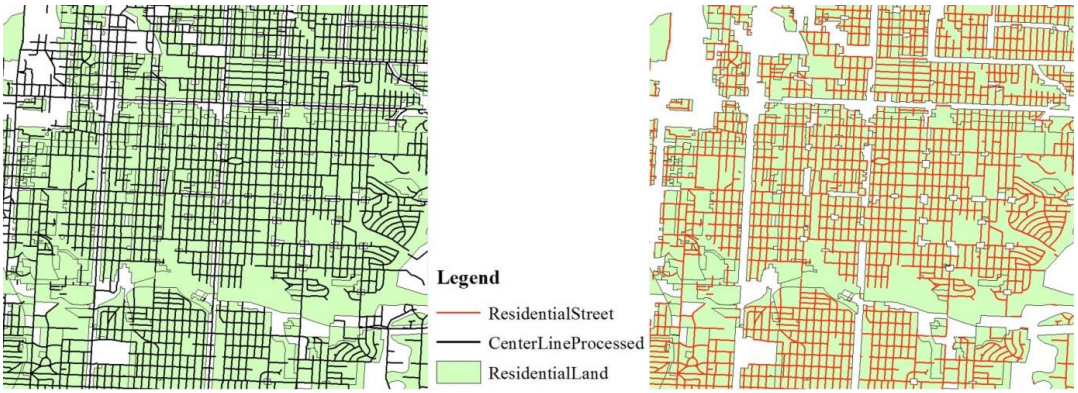


Fig.4.2. Image composition of a GSV panorama.

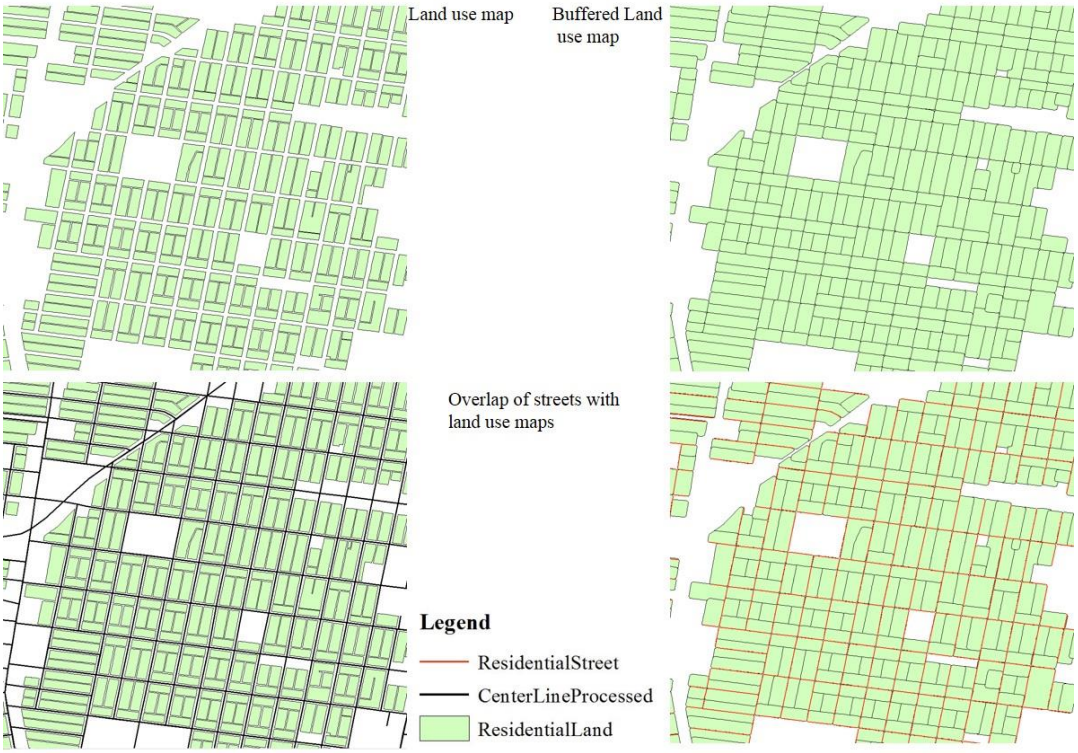
4.1 Geographical sampling

GSV panoramas are distributed discretely along roads. In order to collect available GSV panoramas in the study areas, geographical sampling is required. The geographical sampling is a process to create discrete sample sites along streets. In order to represent the overall greenness level of residential areas, only those streets in residential areas were selected in this study. The road maps were intersected with the corresponding residential land use maps to get the residential street maps for all chosen cities in this study.

However, the land use maps for different cities are in different forms. For regular residential land use maps, in which roads are not a separate land use type and the residential land use maps are continuous patches, the residential street maps were created by intersecting the residential land use maps and the road maps directly (**Fig. 4.3 (a)**). In some cities, the residential land use types are shown as residential parcels (**Fig. 4.3 (b)**). In this case, it is not suitable to use the feature intersection operations to extract the residential streets directly. Therefore, for those parcel-level land use maps, the buffer analysis was first conducted on the residential land use map with a buffer distance of 10m. Then, the residential street map was derived by intersecting the buffered land use map with the street map (**Fig. 4.3 (b)**).



(a) Intersection of street map with residential land use map



(b) Intersection of street map with buffered residential land use map

Fig.4.3. The overlap of residential land use map with street map.

Commercial software ESRI ArcGIS provides a tool – *CreateRandomPoints* for generation of random points in an extent window, inside polygon features, on point features, or along line features. Users can set the minimum distance between the created points. However, the created

points usually do not meet the minimum allowed distance restrictions and are usually too aggregated or too sparse. **Fig. 4.4 (a)** shows the created points along residential streets using the *CreateRandomPoints* tool of ArcGIS 10.2 in a small area of Milwaukee, WI. It can be seen clearly that some generated points are congested together and do not meet the condition of the 100 meters minimum allowed distance. Therefore, in this study, I used the *positionAlongLine* tool in console to create points evenly along streets. The minimum distance between any nearby points was set to 100m, in order to make sure there is one point every 100m along streets in residential areas in each city. **Fig 4.4 (b)** shows the created sample points along the residential streets using the *positionAlongLine* tool. Compared with the created sample points in **Fig 4.4 (a)**, the created sample points using the *positionAlongLine* are evenly distributed along the streets, and would be better to represent the greenness of residential neighborhoods.

The same workflow for geographical sampling was applied to all cities. **Fig 4.5** shows an example of the workflow for geographical sampling in Hartford, Connecticut.

Algorithm 1. Geographical sampling along residential streets using ArcGIS.

```
input: inputStreetMap': the input residential street map
output: outputSampleMap the create sample sites map
# comments: Chose appropriate projections for different regions
sr = arcpy.SpatialReference(102686)
points = []
for row in arcpy.da.SearchCursor(' inputStreetMap', ["SHAPE@"],spatial_reference=sr):
...     length = int(row[0].length)
...     for i in xrange(10m,length,10m):
...         point = row[0].positionAlongLine(i)
...         points.append(point)
... arcpy.CopyFeatures_management(points, outputSampleMap)
```

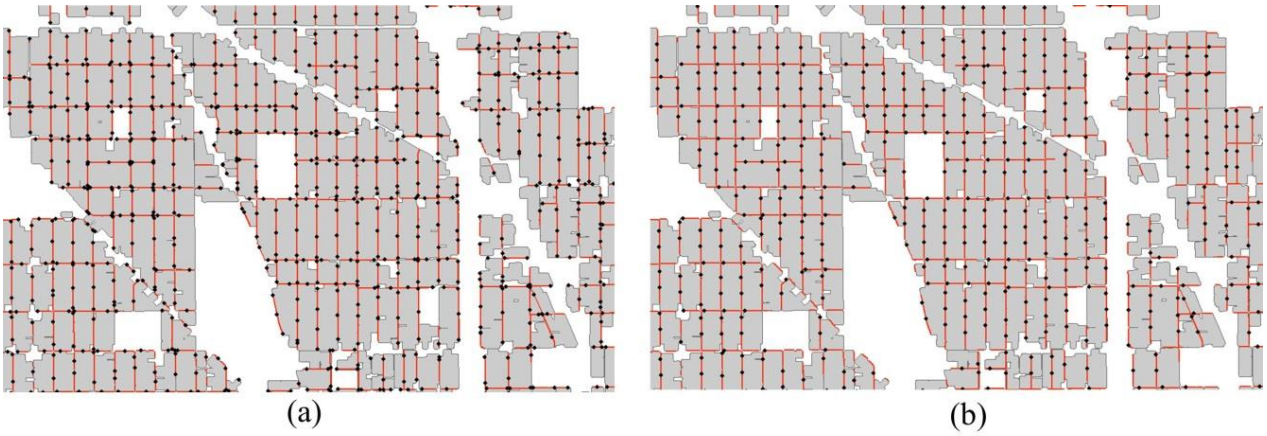


Fig.4.4. The difference of the created points using different methods in ArcGIS, (a) the created random sample sites using the *CreateRandomPoints* tool with the minimum distance of 100m; (b) the created sample sites using the *positionAlongLine* tool.

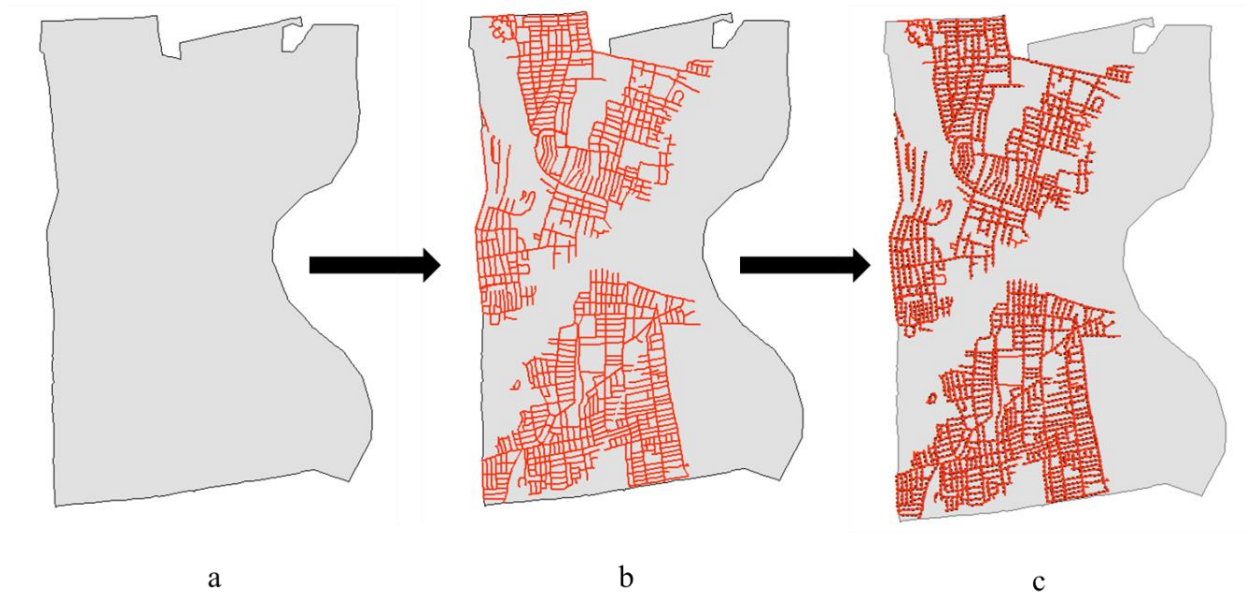


Fig.4.5. An example of geographical sampling in Hartford, Connecticut, (a) administrative boundary of Hartford, (b) residential street map, (c) created sample sites along the residential streets.

4.2 Google Street View images collection

GSV panorama is a 360° surrounding image generated by stitching together the eight original images captured by the eight horizontal cameras in sequences (Tsai and Chang, 2013). The static

GSV images can be requested using a HTTP URL form through the Google Street View Image API (Google, 2014). By specifying the coordinate, direction, and pitch angle in a HTTP URL requested form, users can get the corresponding static GSV image in any direction with any angle for any available site (Li et al., 2015a). An example of requesting a static GSV image is shown below.

GSV URL example:

http://maps.googleapis.com/maps/api/streetview?size=400x400&location=41.935,-87.80524300000002&fov=60&heading=180&pitch=0&sensor=false

Fig.4.6. A static Google Street View image.

Fig. 4.6 shows a static GSV image obtained using the above URL request. In this example, the parameter *size* specifies the output size of the requested GSV image, *location* provides the geo-location of the GSV image (the GSV Image API will snap to the panorama photographed

closest to this location), *heading* indicates the compass heading of the camera (the heading values range from 0 to 360), *pitch* specifies the up or down angle of the camera relative to the street view vehicle, and *fov* determines the horizontal field view of the image. Previous visual assessment studies chose the horizontal field view setting between 50° to 60° (Yang et al., 2009; Walker et al., 1990; Li et al, 2015a; Li et al, 2015b). Therefore, in this study, the *fov* was set to 60°, so that six images can cover the 360° horizontal surroundings.

4.3 Green vegetation extraction from GSV images

In this study, green vegetation extraction is a prerequisite step for utilizing GSV images for street greenery study. However, extracting green vegetation from GSV images is challenging due to many factors, such as the existence of shadows, seasonal variability, and the spectral confusion between vegetation and other manmade green features (e.g. green brands, green doors). What more important is that GSV images are stored in three dimensions using RGB color space, and have no near-infrared bands, which are commonly used for vegetation extraction. Thus, the limited spectral information makes extracting green vegetation from street view images more difficult.

Previous vegetation extract algorithms mainly focus on crop extraction (Guijarro et al., 2011; Woebbecke et al., 1995; Ribeiro et al., 2005). While the bi-classes (vegetation and soil) in crop images are simple and easy to be separated, urban features in street-level images are much more complex and difficult to differentiate. Many artificial features, such as green brands and green doors and windows, share similar spectral signatures with green vegetation in RGB bands. In

addition, green vegetation may show different spectral signatures in those street view images captured under different illumination conditions (e.g., sunny days with good illumination, cloudy days with poor illumination). Moreover, the shadow problem is much more serious in street-level images than in crop images due to the stronger heterogeneity of urban landscapes. All of these make the vegetation extraction from street-level images difficult using spectral information alone. In this study, a robust and time-efficient object based image analysis method was used to classify the green vegetation (Li et al, 2015b; Li et al, 2015c). The vegetation classification algorithm is based on a two-step methodology. The first step is to segment the original GSV images using the mean-shift segmentation algorithm (Comaniciu and Meer, 2002). A simple automatic threshold method – OTSU method (Otsu, 1979) was then used to extract the green vegetation based on the segmented images. Since GSV images were captured in different seasons and different years, therefore, prior to the image classification, the spectral analyses of GSV images at different seasons and years were also conducted.

4.3.1 Spectral Analyses of GSV images at different seasons

Green vegetation extraction from multispectral remotely sensed imagery has been studied for three decades (Almeer, 2012). The near infrared band and red band are the most frequently used bands for detecting vegetation because vegetation shows high reflectance at near infrared band but shows high absorption at red band (**Fig. 4.7**). However, GSV images only cover the red, green, blue bands, and the near infrared band is not available. By checking the Red-Green-Blue spectrum of green vegetation, it is can be seen clearly that green vegetation shows higher

reflectance at green band than other bands (**Fig. 4.7**), although the contrast is not as large as red and near infrared. This phenomenon is further proved by the spectral analysis of green vegetation using GSV images. In this study, 8,187 green vegetation pixels were chosen from the GSV images captured under different illumination conditions as samples for the spectral analysis. **Fig. 4.8 (b)** shows the mean values (μ_{rgb} = [0.301, 0.332, 0.230]) and the standard variances (δ_{rgb} = [0.186, 0.178, 0.221]) of the selected green vegetation pixels in the red, green, and blue bands.

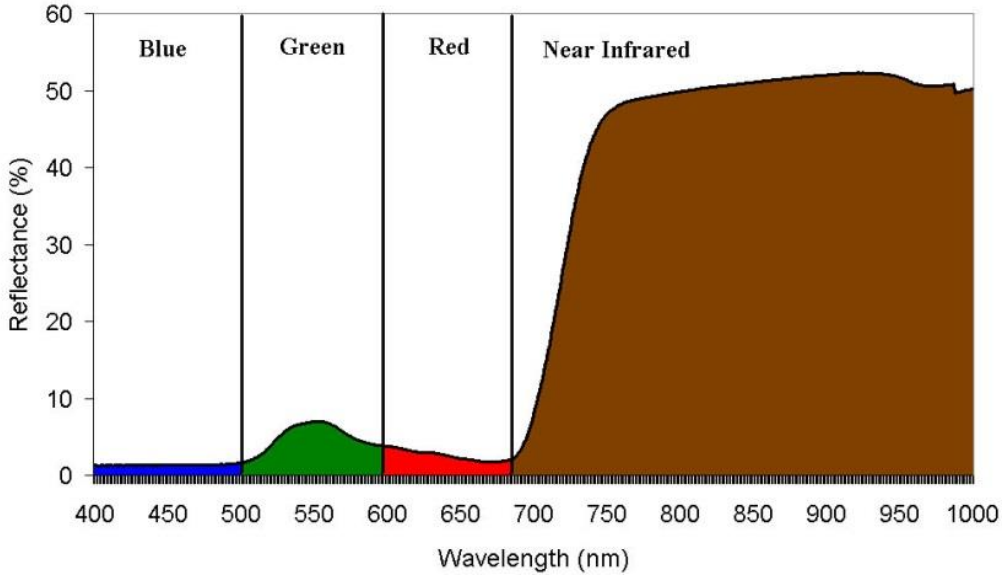
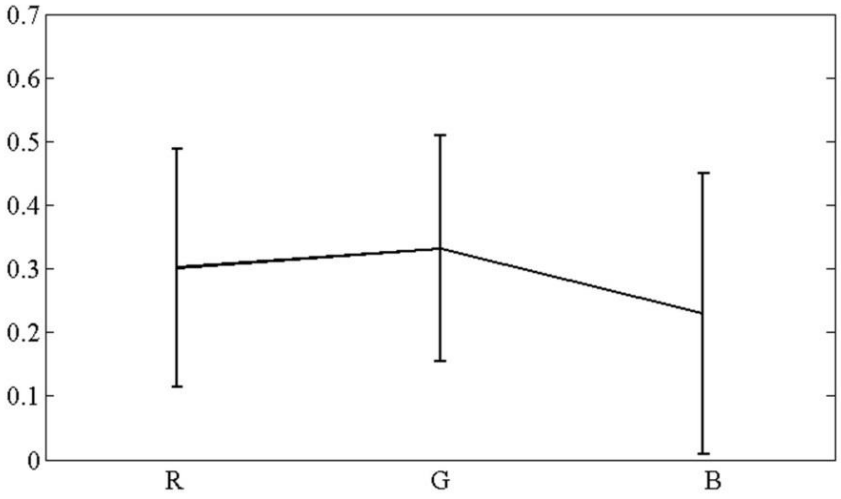


Fig.4.7. The spectral signature of vegetation, source from: http://www.seos-project.eu /modules/agriculture/agriculture-c01-s01.html.

(a)



(b)

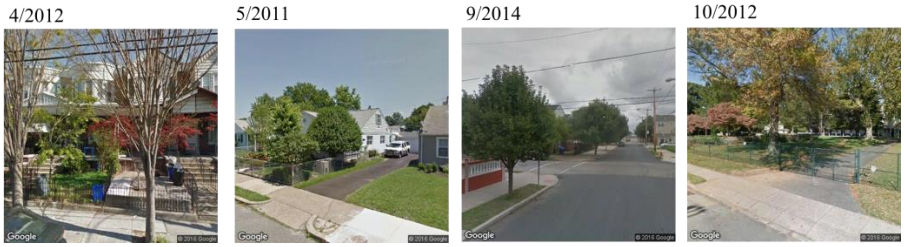
Fig.4.8. GSV images and spectral signature of green vegetation, (a) GSV images under different illumination conditions in green seasons, (b) averaged spectral signatures and corresponding standard variances of the selected vegetation samples in RGB bands.

However, the spectral signatures of vegetation could be affected by seasons. In different climate zones, urban greenery has different green seasons. In Northeast and Midwest, street trees

turn to yellow or red after the end of September (**Fig. 4.9**). However, in the Deep South or California and Arizona, street trees may keep the green in September or even October (**Fig. 4.9**). Table 4.3.1 lists the green seasons in different cities based on the visual inspection of the GSV images in those cities. In addition, different tree types also have very different spectral signatures. Pine trees usually are evergreen, and some street trees are not shown as green even in summer. For example, a cultivar of Norway maple, *Acer platanoides* 'Crimson King' has purple leaves in summer (**Fig. 4.10**).

Boston, Massachusetts

Philadelphia, Pennsylvania



Pittsburg, Pennsylvania

Detroit, Michigan

6/2014

9/2013

10/2014

Chicago, Illinois

5/2011

9/2015

10/2014

11/2014

Milwaukee, Wisconsin

5/2009

9/2014

10/2011

Indianapolis, Indiana

3/2012

4/2012

5/2012

9/2011

10/2011

Kansas City, Missouri

4/2015

5/2011

9/2011

10/2014

Minneapolis, Minnesota

St. Louis, Missouri

Baltimore, Maryland

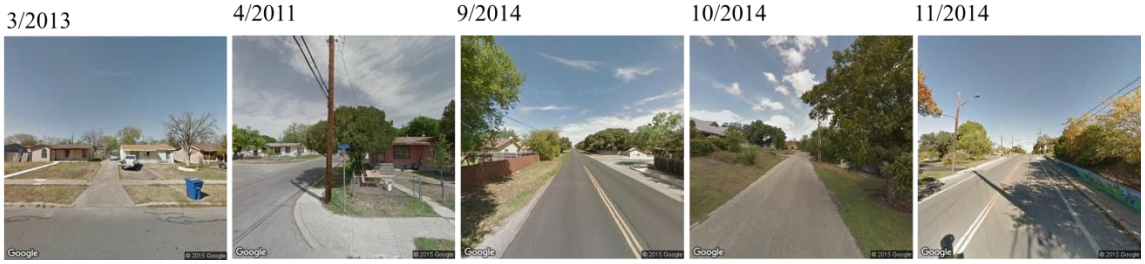
Washington D.C

Tampa, Florida

Louisville, Kentucky

San Antonio, Texas

Dallas, Texas



Denver, Colorado

Fig.4.9. GSV images of different seasons in different cities. Note: For some cities, no image taken in some months, therefore only GSV images in some available months are presented here.

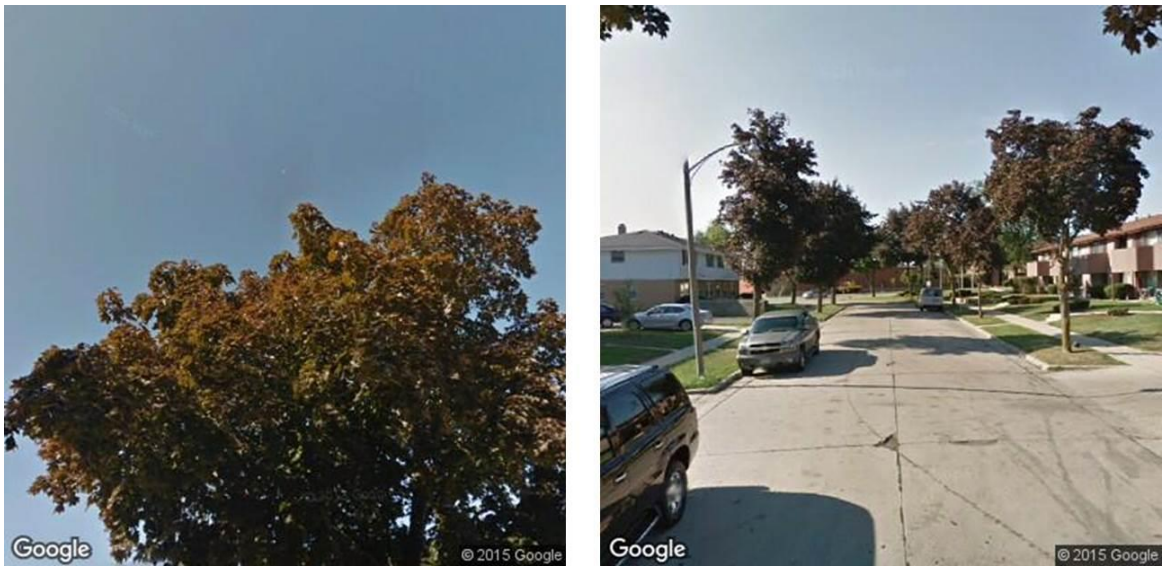


Fig.4.10. Non-green purple leaf Maple *Acer platanoides* 'Crimson King'.

Table 4.3.1
The different green seasons in different climate zones.

Cities	Green seasons	Regions
Boston	May, June, July, August, September	**Northeast**
Philadelphia	May, June, July, August, September	
Pittsburgh	May, June, July, August, September	
Detroit	May, June, July, August, September	**Midwest**
Chicago	May, June, July, August, September	
Milwaukee	June, July, August, September	
Indianapolis	May, June, July, August, September	
Kansas City	May, June, July, August, September, October	
Minneapolis	June, July, August, September	
St. Louis	May, June, July, August, September, October	
Baltimore	May, June, July, August, September	
Washington DC	May, June, July, August, September	
Tampa	January, February, March, April, May, June, July, August, September, October, November, December	**South**
Louisville	June, July, August, September, October	
Houston	--	
San Antonio	April, May, June, July, August, September, October,	

	November	
Dallas	April, May, June, July, August, September, October	
Phoenix	--	**West**
Denver	May, June, July, August, September	
Los Angeles	--	
San Francisco	--	
Seattle	--	

4.3.2 Timing GSV images

Unfortunately, the current version of the Google Street View static image API is not able to directly access to the time information of the GSV images for any specific location. However, Google has provided the time stamps for GSV images, and the time information of the GSV images is accessible through the Google Maps JavaScript API (Google, 2015). GSV panorama ID, longitude, latitude, and time information of the GSV panorama can be accessed through the Google Maps JavaScript API (Google, 2015). The following JavaScript code shows how to get the time information of the GSV images using the Google Map API (Google, 2015). Using the coordinates of the chosen sample sites as the input, the panorama IDs and time information can be accessed.

In the code, the *getPanoramaByLocation* is a function provided by the Google Maps API, the second parameter of this function is used to define the search area. Therefore, if the site (*longitude, latitude*) has no GSV panorama, then the function will snap a panorama from the surrounding 5 meters region. This method can help to guarantee that there are as many sites having available GSV panoramas as possible. The code can also save the panorama IDs and time

information of these panoramas together with the coordinates in JavaScript Arrays.

Pseudo code 1. Meta-data collection of GSV images

```
Input: Coordinates of sample sites
Output: arrays of panorama IDs and time information of panoramas
Comment: latlng is the coordinate of a sample site
Comment: panoIdArr is the array to store the panorama ID of a sample site
Comment: panoDateArr is the array to store the time information of a sample site
var sv = new google.maps.StreetViewService();
sv.getPanoramaByLocation(latlng, 5, storeGSV_Info);
function storeGSV_Info(data, status) {
      if (status == google.maps.StreetViewStatus.OK) {
            panoIdArr.push(data.location.pano);
            panoDateArr.push(data.imageDate);
      } else {
            console.log('street view is not available in this point');
      }
   }
}
```

Fig. 4.11 shows the tool I developed in this study to harvest the metadata of the GSV images. By imputing the coordinates of sample points in *geojson* format, the tool can return the metadata (pano ID, time information, and coordinates) of GSV panoramas for the input sample sites and save the metadata in a local text file.

Although Google has published the GSV images taken in different times, the static GSV image is only accessible for one time point. It is still impossible to access the GSV images at different time points for one site, which means that it is difficult to investigate the temporal changes of street greenery using the GSV. However, it could be possible to access static GSV images at different time points in future, since Google already collected historical GSV data and

published it online. Considering the fact that GSV images were all collected along streets, and not all cityscapes were covered, so, the GSV based method is more suitable for assessing street greenery but not suitable for other types of urban green spaces.

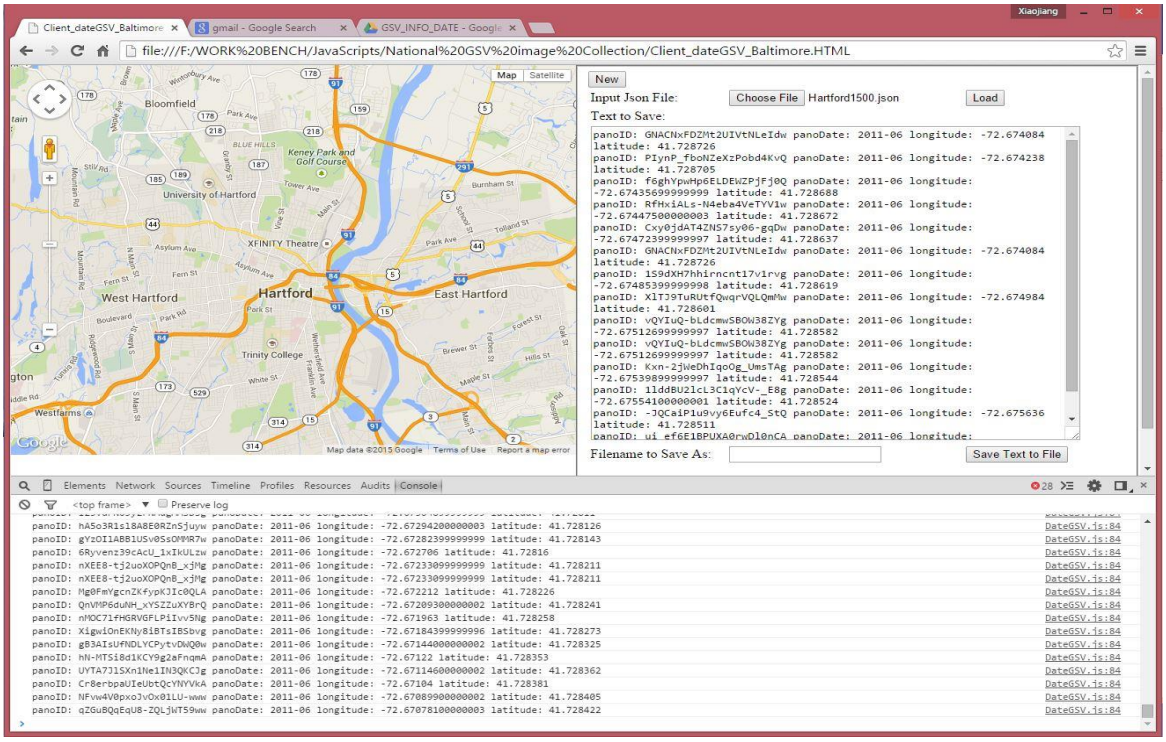


Fig.4.11. The developed tool for harvesting GSV metadata (panorama ID, panorama geo-coordinates, and time information). This tool can also save the panorama information locally for further analysis.

Coordinate: http://maps.googleapis.com/maps/api/streetview?size=400x400&location=41.935,-87.80524300000002&fov=60&heading=180&pitch=0&sensor=false

Pano ID: http://maps.googleapis.com/maps/api/streetview?size=400x400&pano=HfFYWKnZItamBeJxJPX5IA&fov=60&heading=180&pitch=0&sensor=false

Fig.4.12. A GSV image and its corresponding Uniform Resource Locators (URLs).

In the GSV static image API, instead of using the *location* parameter, GSV images can also be requested using *pano* parameter, which represents the panorama ID. Using the *pano* parameter is usually more stable than the *location* parameter (Google, 2015). **Fig. 4.12** shows the URLs of a static GSV image using both *location* and *pano* parameters.

A Python script was developed to download GSV images with meta-data based on the coordinates, panorama ID, and time information collected by *GSV_TIME_TOOL* (**Fig. 4.11**) for all selected sample sites. The script was used to collect GSV images and map the spatial distribution of time information of GSV images for all chosen cities in this study. **Fig. 4.13** (a-b) shows the spatial distribution and statistics of time information for all sample sites in Hartford, Connecticut. Of the 3,000 sample sites in Hartford, Connecticut, only 2,838 sites have GSV

panorama coverage. **Fig. 4.13(a)** shows the spatial patterns of image date information for all 2,838 GSV panorama sites. The GSV images for most of the sites (2,042) were taken in June 2011 (**Fig. 4.13(b)**). Other time points include July 2015 (390 sites), August 2012 (216 sites), July 2011 (79 sites), and October 2011 (84 sites). A few sites have GSV images captured in August 2007 (13 sites), August 2011 (10 sites), October 2012 (3 sites), and July 2008 (1 site).

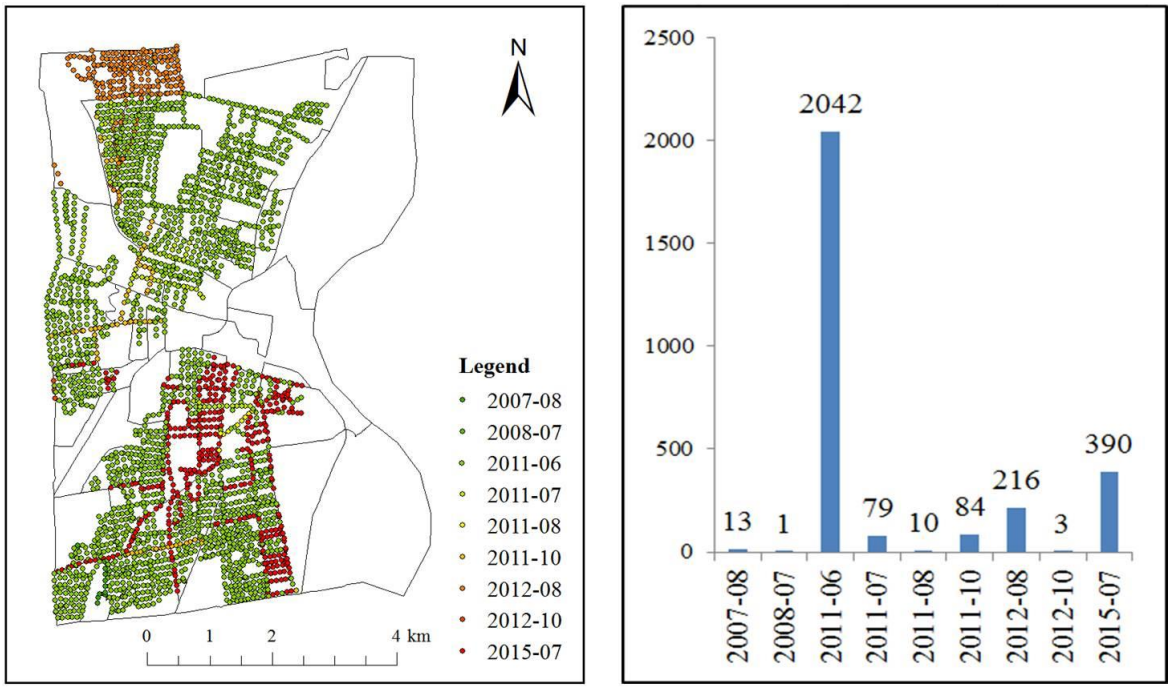


Fig.4.13. Selected sample sites and related image date information in a case study in Hartford, Connecticut: (a) the spatial distribution of date information for all chosen GSV images, (b) the statistics of the date information for all chosen samples.

4.3.3 Image segmentation

Image segmentation is a process of delineating an image into homogeneous polygons that are physically meaningful. It also can differentiate objects based on available geometric information

(Blaschke, 2010). The mean-shift image segmentation method is a simple iterative procedure for locating the maxima of density functions given discrete sample data. It was initially proposed by Fukunaga and Hostetler (1975). Later, this method was refined by Cheng (1995) through adding a kernel function, which further rekindled the interest in it. Since then, the mean-shift segmentation algorithm has been widely used in various applications due to its robustness and capability of generating qualified cluster results. In the mean-shift algorithm, the first step is to build a kernel function to estimate the possibility of the density function based on the original image. There are many methods to estimate the possibility of the density function. In this study, a normal kernel function

$$K(\mathbf{x};h) = \frac{(2\pi)^{-d/2}}{h^d} \exp(-\frac{1}{2} \| \frac{\mathbf{x}}{h} \|^2) \tag{4.1}$$

was used to estimate the probability distribution. In Eq. (4.1), d represents the dimensions of a space, $\mathbf{x}$ is the feature vector of the central pixel in a kernel, h is the window size of a kernel function and also acts as the scale parameter, and ||.|| is a norm.

For n data points $\mathbf{x}_i$ (i = 1, 2,..., n) in the d-dimensional space, the multivariate density function estimator is

$$f(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} K(\mathbf{x}_i - \mathbf{x}, h), \tag{4.2}$$

where $\mathbf{x}_i$ stands for the feature vector of each pixel. In this study, the $\mathbf{x}_i$ stands for pixel values at RGB bands, for example, $\mathbf{x}_i$ = [10, 255, 255].

Introducing the kernel function (4.1) into Eq. (4.2), we get the multivariate density function

estimator as

$$f(\mathbf{x}) = \frac{(2\pi)^{-d/2}}{nh^d} \sum_{i=1}^{n} \exp(-\frac{1}{2} \| \frac{\mathbf{x}-\mathbf{x}_i}{h} \|^2) \tag{4.3}$$

By writing $k(u) = \exp(-\frac{1}{2}u)$ and $C = \frac{(2\pi)^{-d/2}}{nh^d}$, we can simplify the notation of Eq. (4.3) to

$$f(\mathbf{x}) = C \sum_{i=1}^{n} k(\| \frac{\mathbf{x}-\mathbf{x}_i}{h} \|^2) \tag{4.4}$$

The density gradient for $f(\mathbf{x})$ or the derivative of $f(\mathbf{x})$ is,

$$\nabla f(\mathbf{x}) = \frac{2C}{h^2} \sum_{i=1}^{n} (\mathbf{x}-\mathbf{x}_i) k'(\| \frac{\mathbf{x}-\mathbf{x}_i}{h} \|^2) \tag{4.5}$$

Define the function $g(x) = -k'(x)$, we get

$$\begin{aligned} \nabla f(\mathbf{x}) &= \frac{2C}{h^2} \sum_{i=1}^{n} (\mathbf{x}_i - \mathbf{x}) g(\| \frac{\mathbf{x}-\mathbf{x}_i}{h} \|^2) \\ &= \frac{2C}{h^2} \left[\frac{\sum_{i=1}^{n} x_i g(\| \frac{\mathbf{x}-\mathbf{x}_i}{h} \|^2)}{\sum_{i=1}^{n} g(\| \frac{\mathbf{x}-\mathbf{x}_i}{h} \|^2)} - \mathbf{x} \right] \left[\sum_{i=1}^{n} g(\| \frac{\mathbf{x}-\mathbf{x}_i}{h} \|^2) \right] \end{aligned} \tag{4.6}$$

Since the $\sum_{i=1}^{n} g(\| \frac{\mathbf{x}-\mathbf{x}_i}{h} \|^2)$ is assumed to be a positive number (Comaniciu and Meer, 2002), the maximum of the density function $f(\mathbf{x})$ occurs when

$$\mathbf{x} = \frac{\sum_{i=1}^{n} \mathbf{x}_i g(\| \frac{\mathbf{x}-\mathbf{x}_i}{h} \|^2)}{\sum_{i=1}^{n} g(\| \frac{\mathbf{x}-\mathbf{x}_i}{h} \|^2)} \tag{4.7}$$

The mean-shift algorithm is an iterative procedure of updating x_i as Eq. (4.8) until it converges (Comaniciu and Meer, 2002).

$$\mathbf{x}_{i+1} = \frac{\sum_{i=1}^{n} \mathbf{x}_i g(\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \|^2)}{\sum_{i=1}^{n} g(\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \|^2)} \tag{4.8}$$

When Eq. (4.8) converges, the original image is segmented into different objects and those neighboring pixels sharing similar spectral signatures are clustered into one object. The size of window and the kernel function affect the mean-shift segmentation results. In this study, after trials and errors, the size of window was set to 5 for its best segmentation quality.

The meanshift Python module – *pymeanshift* was used in this study for image segmentation. The spectral *R*, *G* and *B* components in original 8-bit color RGB GSV images were first normalized to the range of [0,1] for segmentation. **Fig. 4.14** (b) shows the segmented results. Neighborhoring pixels in the original GSV images (**Fig. 4.14(a)**) were grouped together. After segmentation, the new thematic images (**Fig. 4.14(c)**) were generated by setting the attribute of each object to the average value of pixels in that object at each of the three RGB bands. Compared with the original GSV images, the thematic images are smoothed spectrally, and the contrast between green features and non-green features is enhenced. Both of these make the thematic images more suitable for vegetation classification, therefore, the thematic images were used to extract geen vegetation in the next step.

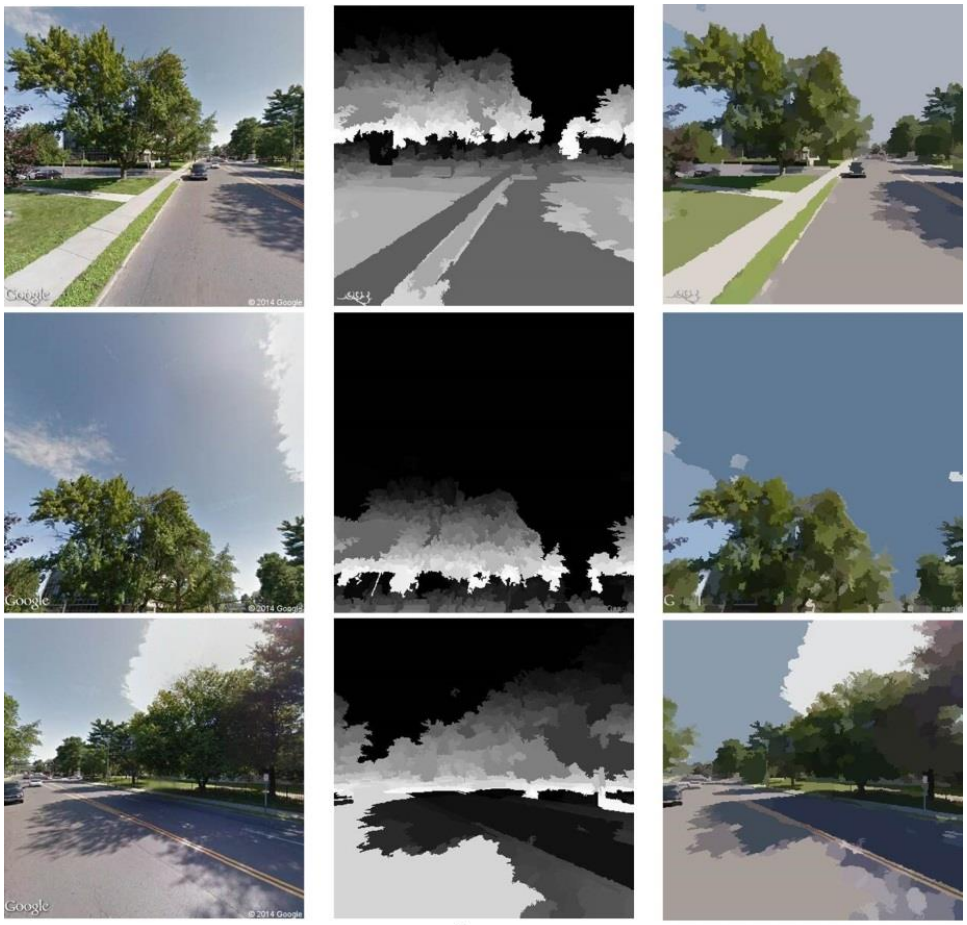


Fig.4.14. The image segmentation results for GSV images, (a) the original GSV images, (b) the segmented objects, (c) the thematic results (the value of each object is the mean value of all pixels in it).

4.3.4 Vegetation classification and validation

There are many developed image-processing algorithms for separating green vegetation from non-vegetation using RGB colorful images (Guijarro et al., 2011; Meyer et al. 2004; Zheng et al. 2009, 2010; Li et al., 2015a, 2015b). Current algorithms for green vegetation extraction fall into three categories (Guijarro et al., 2011). The first one comprises the visible spectral index based methods, including the excess green index (*ExG = 2G-R-B*) (Woebbecke et al., 1995; Ribeiro et al., 2005), the normalized difference index (*NDI = (G-R)/(G+R)*) (Woebbecke, 1992), the excess

red index (*ExR=1.4R-G*) (Meyer et al., 1998), and the excess green minus excess red index (*ExGR = ExG − ExR*) (Neto, 2004). The second category belongs to dynamic thresholding approaches, including the Otsu-based methods (Ling and Ruzhitsky, 1996; Shrestha et al., 2004) and the histogram entropy based methods (Tellaeche, 2008). The third category contains machine learning based methods, which use supervised or unsupervised algorithms to separate green vegetation from non-vegetated features (Meyer et al. 2004; Zheng et al. 2009, 2010).

In this study, based on the spectral signature of vegetation in RGB bands, a rule based on the spectral characteristics of green vegetation was set for the green extraction. The rule combines *Diff* based method (Li et al., 2015a) and *ExG* based method (Li et al., 2015b).

There are several steps in the *Diff* based method (Li et al., 2015a). Firstly, two difference images *Diff1* and *Diff2* were generated through the operations of *Diff1 = green band - red band* and *Diff2 = green band - blue band*, respectively. Then the two difference images were multiplied to generate the *Diff* image. Considering the fact that green vegetation pixels normally have higher values in the green band than in the other two bands, they will generally show positive values in the *Diff* image. Those pixels that have lower values in the green band than in the blue or red bands will show negative values in the *Diff* image. However, if a pixel has a lower value in the green band than in both the red and the blue bands, its corresponding value in the *Diff* image will still be positive. To avoid this confusion, an additional rule that the values of green vegetation pixels in the green band must be higher than in the red band is added to extract the green pixels.

Excess green index *ExG* was calculated through the operation of *ExG = 2green band – red band – blue band*. Considering the fact that green vegetation shows higher value in green band

than red band and blue band, the *ExG* would enhance the contrast between green vegetation and non-green urban features in the segmented GSV images. The Otsu algorithm was then used for choosing the optimum threshold to differentiate green vegetation and non-green features.

In the thematic GSV images, only those pixels meet both of these two rules (*Diff* rule and *ExG* rule) would be classified as the green vegetation. The pseudo code for green vegetation extraction algorithm is listed below as pseudo code 2.

Pseudo code 2. Algorithm for extracting greenery from segmented GSV images

Spectral rules for vegetation classification based on segmented GSV images

Comment: *green*, *red*, and *blue* are three bands in segmented images
Comment: *Vegetation* is the vegetation extraction result
ExG = 2green – red – blue
Diff1 = green – red
Diff2 = green – blue
Diff = Diff1× Diff2
Threshold = OTSU(ExG)
for *each pixel [i, j]:*
    if *ExG* [*i, j*] *> Threshold and Diff [i,j]>0 and Diff2>0:*
        Classify Vegetation [*i, j*] *as green vegetation*
Mask out pixels with values in green, red, blue bands higher than 0.7 in Vegetation image

Fig. 4.15 shows the image segmentation results and green vegetation classification results of three GSV images. The image segmentation algorithm first clustered nearby pixels, which have similar spectral characteristics into different objects (**Fig. 4.15** (b-c)). A comparison of segmentation results and the original GSV images shows that the image segmentation algorithm can help to keep the urban feature boundaries and smooth the noises in the GSV image. From the

vegetation extraction results (**Fig. 4.15** (d)), it can be seen that the proposed vegetation classification method extracts the green vegetation correctively from GSV images.

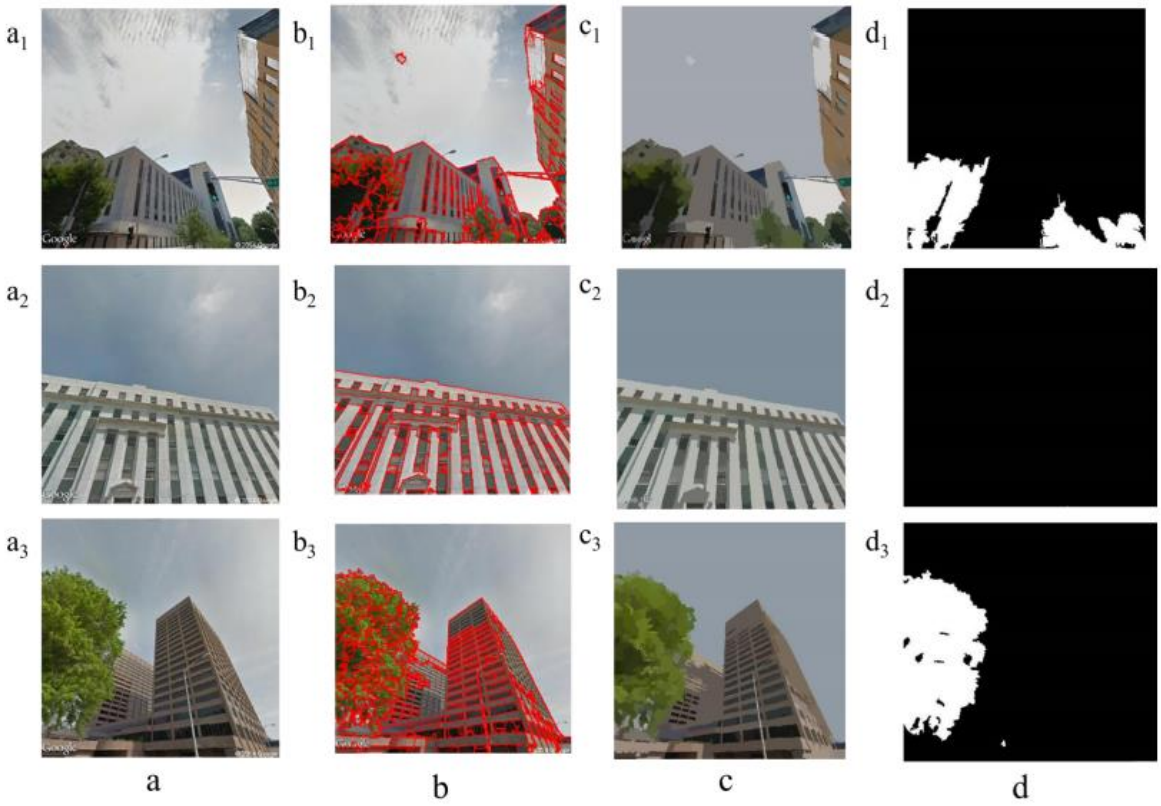


Fig.4.15. Vegetation extraction results, (a) the original GSV images, (b) segmentation results (red lines are the boundaries of objects), (c) thematic results, (d) the vegetation classification results.

This proposed vegetation classification algorithm only classified the vegetation shown as green in the GSV images. The proposed vegetation classification method is not suitable to classify vegetation from those GSV images captured in non-green seasons. Based on the spectral analysis of the green vegetation in different seasons, those sites have GSV images captured in non-green seasons were excluded from the analysis. The classification results of green seasons and non-green seasons for different studied cities can be found in **Table 4.3.1** and **Fig.4.9** in

section 4.3.1.

One hundred GSV images were randomly chosen to validate the classification results. The validation result shows that the overall accuracy of the vegetation extraction is higher than 85%, thus it is qualified for further analysis.

4.4 Green View indices

The sensory benefits provided by the street greenery are majorly through the visibility of the street greenery. Yang et al (2009) proposed a green view index (GVI) to evaluate the visibility of urban forests. Their GVI was defined as the percentage of the total green pixels from four pictures taken at a street intersection to the total pixel numbers of the four pictures, calculated using the Eq. (4.9),

$$GreenView = \frac{\sum_{i=1}^{4} Area_{g\_i}}{\sum_{i=1}^{4} Area_{t\_i}} \times 100\% \quad (4.9)$$

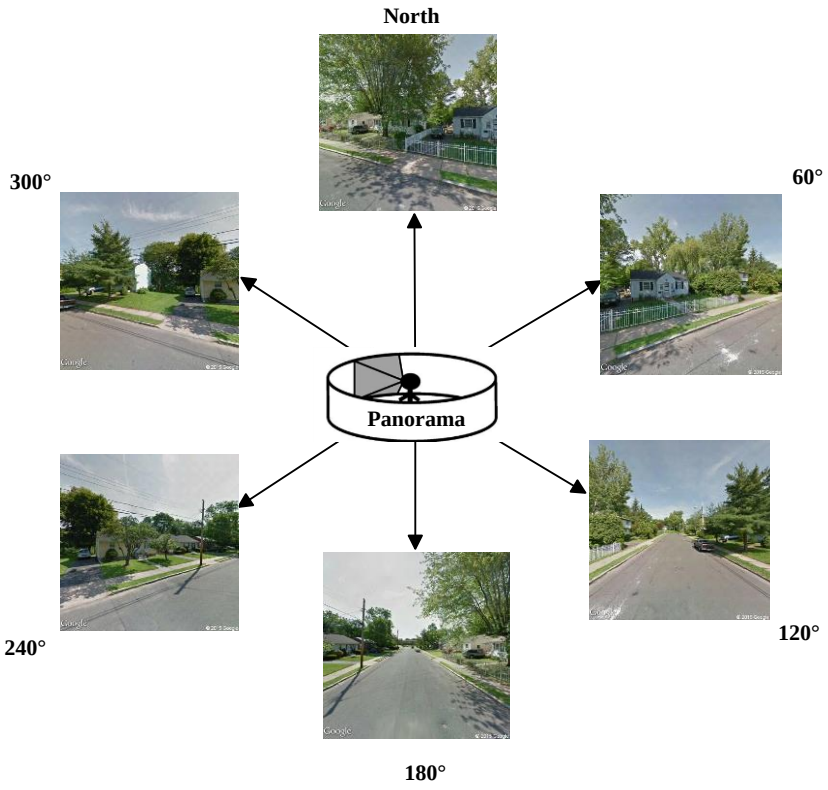
where $Area_{g\_i}$ is the number of green pixels in the picture taken in the *i*th direction among the four directions (true north, true east, true south and true west) for one intersection, and $Area_{t\_i}$ is the total pixel number of the picture taken in the *i*th direction.

The green view index was proposed to represent how much greenery people can see on the ground based on the street-level images. However, using the images captured in the four directions to calculate the green view index inevitably misses some vegetation around, because only four pictures at the field of view of 55° cannot cover the whole scene pedestrians can see

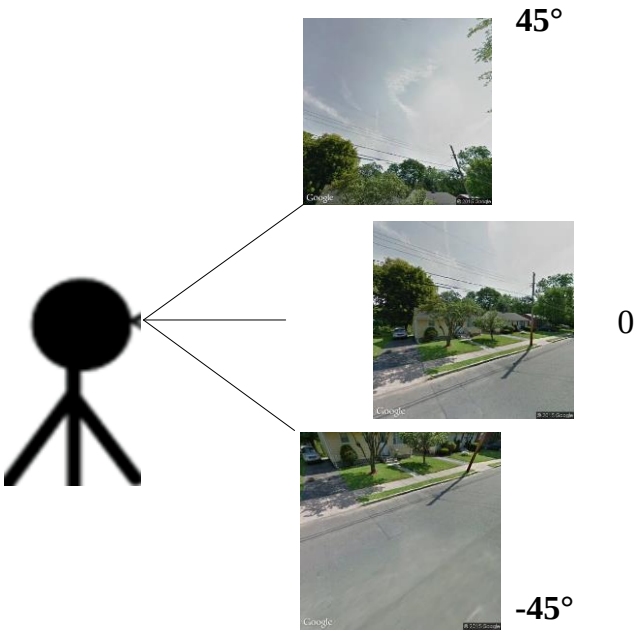
(Li et al, 2015a). In this study, the modified green view index (MGVI) was used. The MGVI uses six images covering the 360° horizontal surroundings to calculate the index for each sample site along streets. Furthermore, to effectively represent the surrounding greenness that pedestrians can see, three different vertical view angles (**Fig. 4.16**(b)) were also considered in each direction for calculating the MGVI (Li et al, 2015a). Consequently, the final MGVI used in this study was calculated using 18 GSV images for each site. Therefore, using the MGVI could better represent the distribution of street greenery, considering the fact that the Google Street View has a similar view angle with pedestrians on the ground. The MGVI calculation formula is written as,

$$GVI = \frac{\sum_{i=1}^{6}\sum_{j=1}^{3} Area_{g\_ij}}{\sum_{i=1}^{6}\sum_{j=1}^{3} Area_{t\_ij}} \times 100\% \tag{4.10}$$

where $Area_{g\_ij}$ is the number of green pixels in one of these images captured in 6 directions with three vertical view angles (-45°, 0°, 45°) for each sample site, and $Area_{t\_ij}$ is the total pixel number in one of the 18 GSV images.



(a)



(b)

Fig.4.16. GSV images captured in six horizontal directions at a sample site and three vertical view angles.

The GSV has a national coverage, and most cities in U.S. have the GSV coverage. Using the Google Street View static image API, street-level images can be collected for any site with GSV panorama available (Li et al, 2015a). This makes the GSV based MGVI can be calculated for any site with GSV available. In this study, the sample sites are located evenly along residential streets, which make the GSV based MGVI better to represent the street greenery in cities. In addition, there is no need to collect *in situ* street-level images to calculate MGVI using GSV.

Chapter 5 MGVI and vegetation characteristics

There are many developed green metrics to evaluate the distribution of urban greenery. However, most of the previously proposed methods study the urban greenery from the overhead view. What human being feel or see on the ground is very different from the overhead view captured or represented by remote sensing data and zoning data. The Google Street View (GSV) based method gives us a new perspective to study the urban greenery. It is important to investigate the relationship and discrepancy between the GSV based MGVI with the previously proposed green metrics and the vegetation characteristics.

Previous studies investigated the relationship between the metrics derived from street-level images and remotely sensed data (Yang et al., 2009; Chen et al, 2015; Li et al, 2015a). Yang et al (2009) first proposed to use green view index (GVI), which is calculated base on *in situ* street level pictures, to assess the visibility of urban forest. Correlation analysis shows that the GVI values have a strong correlation with the tree/shrub covers (Yang et al, 2009). The ANOVA analysis shows that the GVI is influenced by the size of trees, the distance between the trees and viewers, and other kinds of greenery. Chen et al (2015) built a regression model between the GVI

and vegetation parameters derived from LiDAR data. The distances between the trees and the viewer, the perimeter of tree crown, and canopy height parameters were found to be the best explaining variables of GVI. However, there exist a few studies about the association between the GSV based MGVI and the overhead view green metrics, except Li et al (2015a). In Li's paper, only the canopy cover and the grass cover were selected to compare with the MGVI. In addition, the comparison study was conducted in a very small area. In this chapter, the comparison study of the GSV based MGVI and overhead view data based green metrics was conducted on a large scale. In addition, this chapter also compared the difference between the distribution of MGVI and different types of urban greenery.

5.1 Case study areas and data sources

Boston, MA and Hartford, CT were chosen as the case study areas to compare the MGVI and the traditional green metrics. Boston and Hartford are the capital cities of Massachusetts and Connecticut, respectively (**Fig.5.1**). Boston is the largest cities in both Massachusetts and New England. According to recent census data (American Census Survey 2009-2014 data), Boston has a total population about 660,000. African American, Hispanics, and Asians account for 24.1%, 18.8, and 9.0% of the total population, respectively. Hartford is the capital city and fourth-largest city in Connecticut, USA, with population of approximately 125,000. Based on the 5-year aggregated census data from American Community Survey (US Census Bureau, 2012), African Americans and Hispanics are the two largest racial/ethnic groups in Hartford, which account for 37.65% and 43.05% of the total population, respectively. Recently, satellite imagery

based analysis shows that more than 2,870 acres of Hartford are covered by tree canopy, representing 26% of landmass in the city.

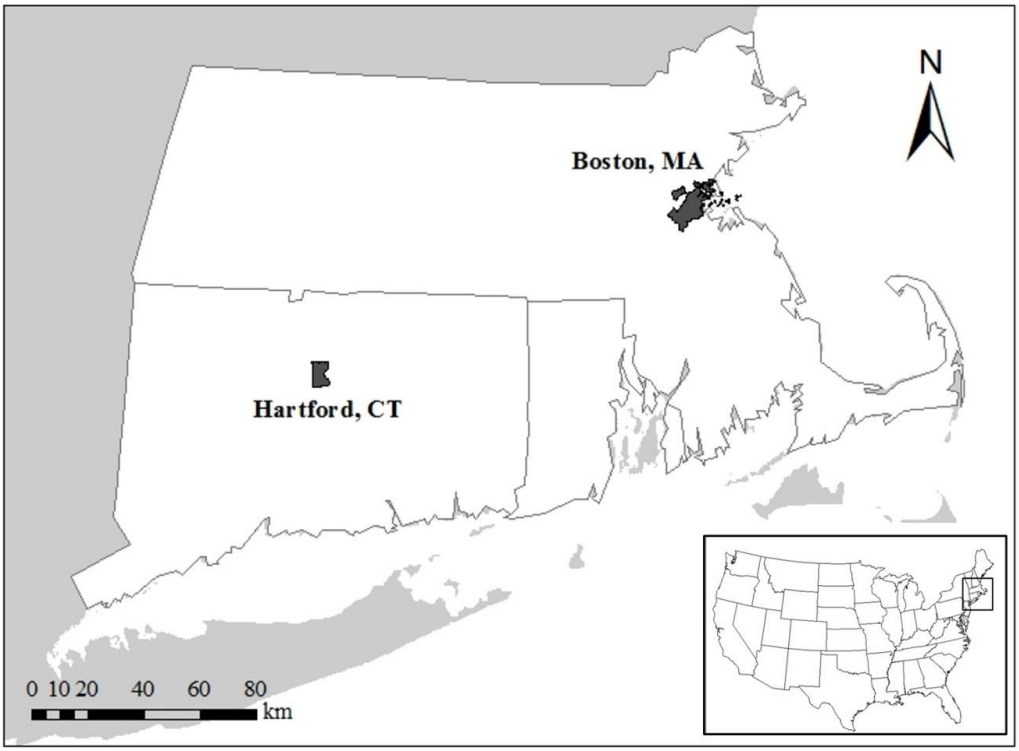
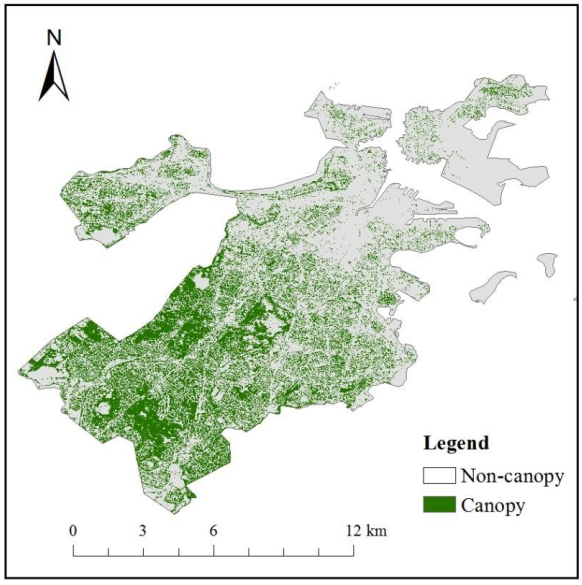


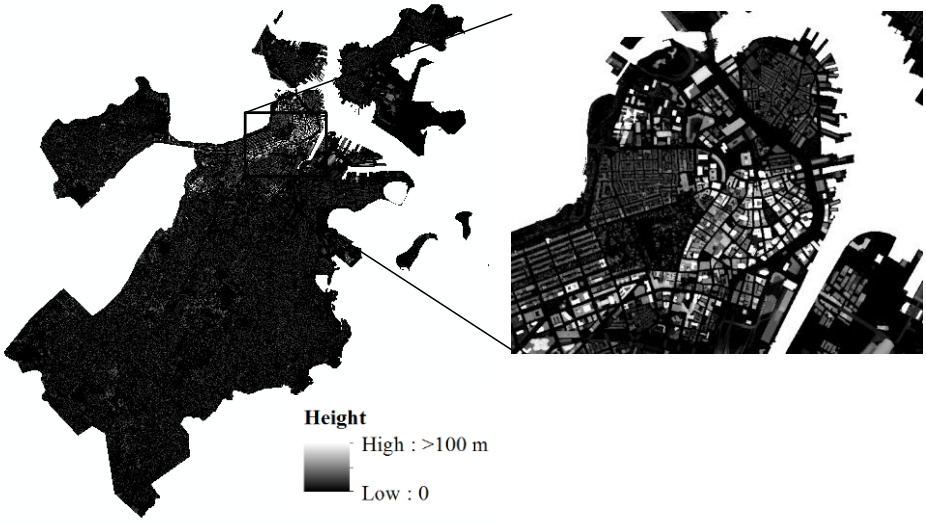
Fig.5.1. The location of Boston, MA and Hartford, CT.

The datasets of Boston include a vegetation cover map (**Fig. 5.2** (a)) and a DSM (Digital Surface Model) derived from LiDAR data (**Fig. 5.2** (b)). The LiDAR data products of Boston, in the form of pre-processed x, y, z points cloud files, were obtained from NOAA Digital Coast (http://coast.noaa.gov/dataviewer/index.html?#). The horizontal accuracy is 50 cm, and vertical accuracy is reported as 15cm. The LiDAR point cloud data includes two geo-spatial layers representing the first returns and the ground. The point cloud file was converted to a raster file using ArcGIS 10.2. The DSM was then created by subtracting the ground model from the first returns layer. The vegetation cover map was obtained from Raciti et al (2014), which was

derived from 2.4 resolution QuickBird images and 1m resolution LiDAR data. The vegetation cover map represents the urban canopy cover, because those vegetation areas less than 1m in height were removed (Raciti et al, 2014). Accuracy assessment result shows that the canopy map has an overall accuracy of 95%.



(a) Canopy cover map of Boston



(b) The DSM model of Boston derived from LiDAR data.

Fig.5.2. The canopy cover and DSM in Boston.

The land cover map of Hartford (http://gis.w3.uvm.edu/utc/) was derived from 1-meter resolution remotely sensed data of 2008 (**Fig. 5.3**). The land cover classification was performed by Spatial Analysis Laboratory (SAL) at the University of Vermont, in consultation with the USDA Forest Service's Northern Research Station.

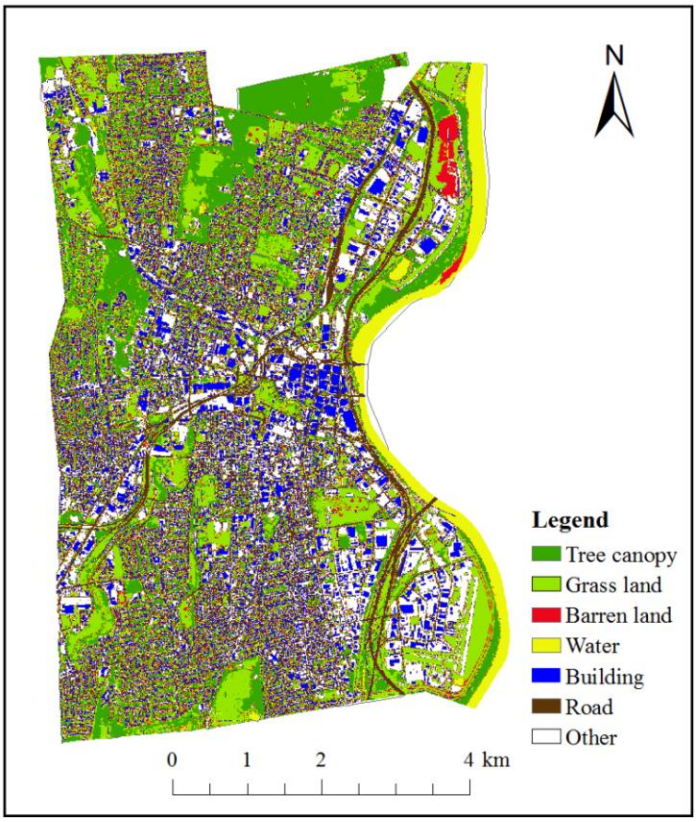


Fig.5.3. The land cover map of Hartford, Connecticut.

5.2 Modified Green View index (MGVI)

Sample sites were first created along residential streets in Hartford and Boston respectively. The created sample sites were then used as inputs to the *GSV_TIMING* tool (see Section 4.3.2) to request the meta-data of GSV panoramas for all selected sample sites. **Fig. 5.4** shows the spatial

distribution of the time information for all chosen sample sites in Boston and Hartford, respectively. In Hartford, among the total 2,838 GSV sites, most of these GSV sites (2,042) have their static images taken in June 2011 (**Fig. 5.5(a)**). Other time points include July 2015 (390 sites), August 2012 (216 sites), July 2011 (79 sites), and October 2011 (84 sites). In Boston, the GSV images were updated very frequently. Therefore, only the year information is listed here (**Fig. 5. 5(b)**), although all available GSV panoramas have time information at month level. Most GSV panoramas in Boston were taken in 2013 (2,777) and 2014 (2,907). There are few panoramas taken in 2007 (55), 2009 (30), and 2012 (8).

In order to keep the time consistency of all chosen sample sites, those sites that have GSV images taken in non-green seasons were excluded in this study, because it is difficult to extract the vegetation from those GSV images. Based on the spectral analysis of GSV images captured in different seasons, May, June, July, August, and September were defined as green seasons in Boston and Hartford. Other months were categorized as non-green seasons and those GSV images taken in non-green seasons were removed from further analysis. **Fig. 5.6** shows the spatial distributions of sample sites have GSV panoramas taken in green seasons and non-green seasons in Hartford (**Fig.5.6 (a)**) and Boston (**Fig.5.6 (b)**). By checking the spatial distribution of sample sites having GSV images taken in non-green seasons, it can be seen clearly that those non-green sites are not aggregated and the relatively small number of non-green GSV sites have no much influence on the further analysis.

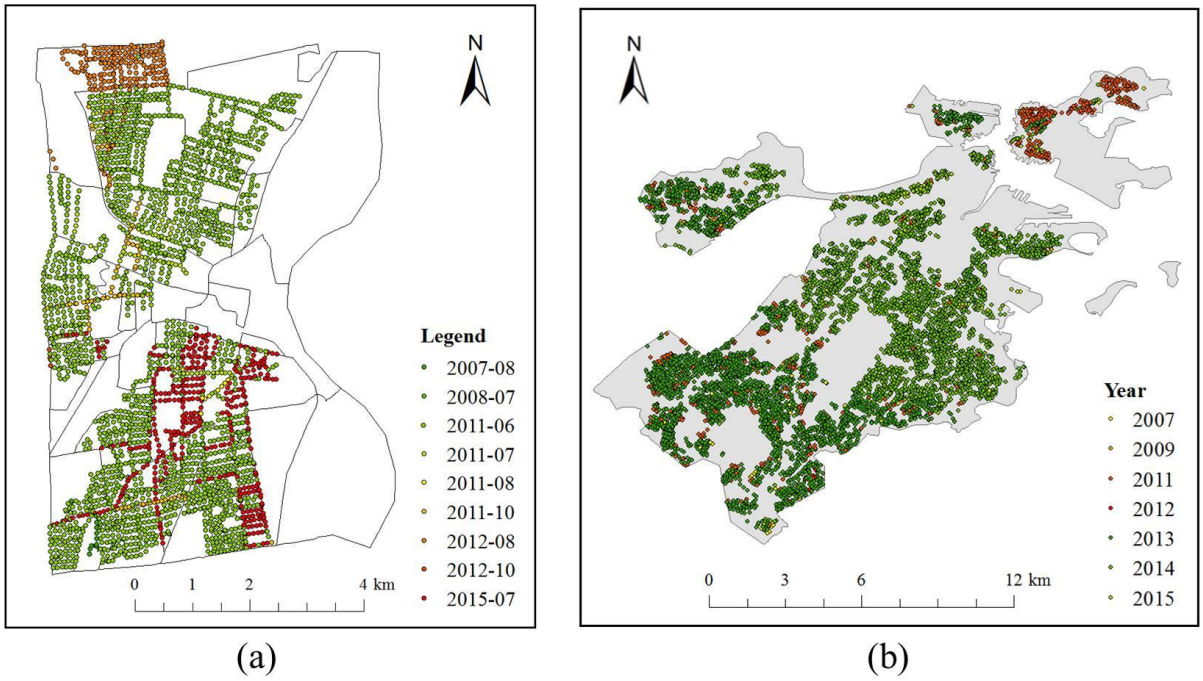


Fig.5.4. The spatial distributions of captured dates of GSV sample sites in Hartford (a) and Boston (b).

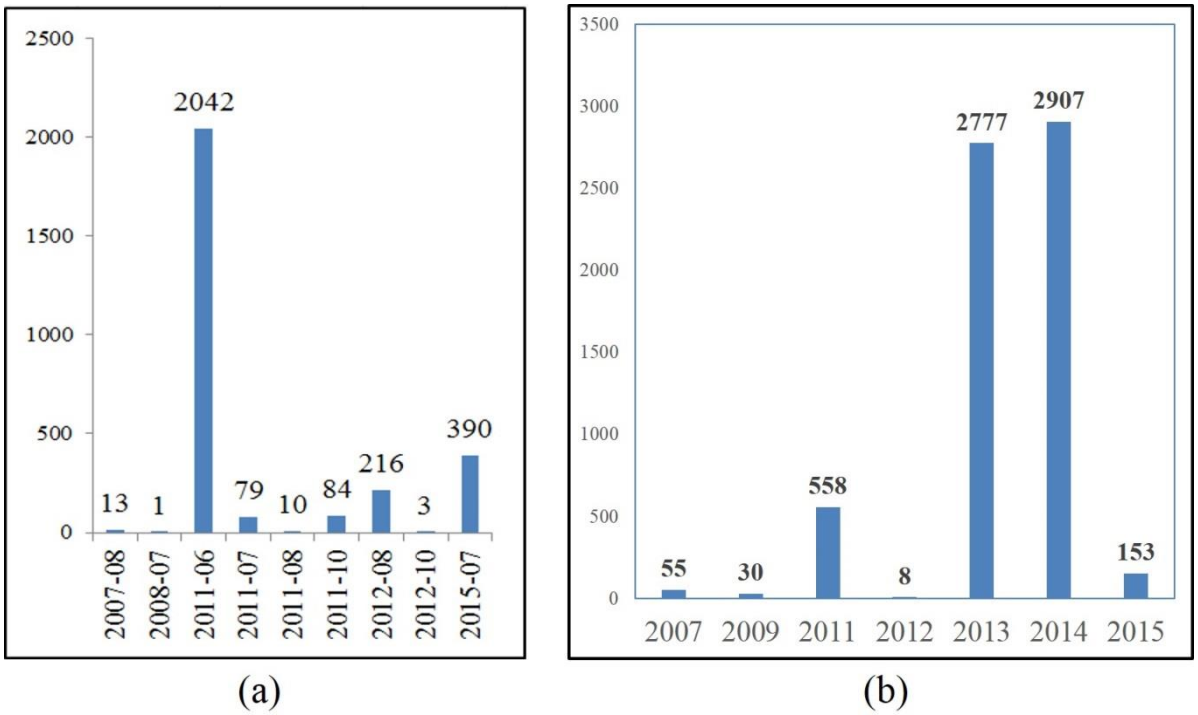


Fig.5.5. The statistics of captured date in the images in Hartford (a) and Boston (b).

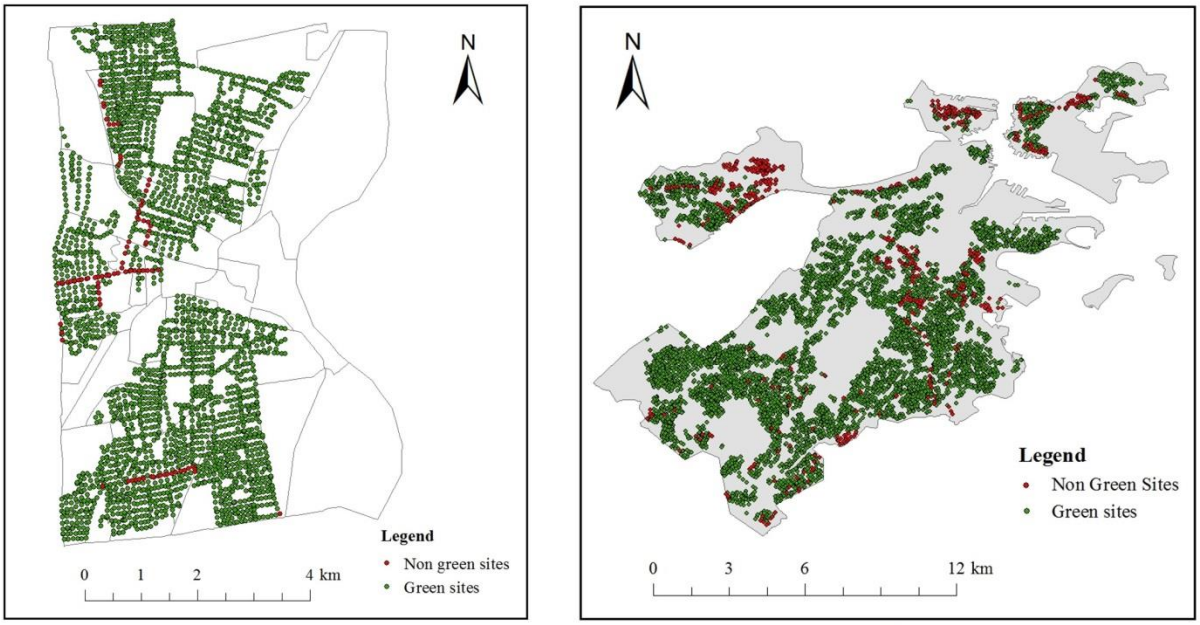
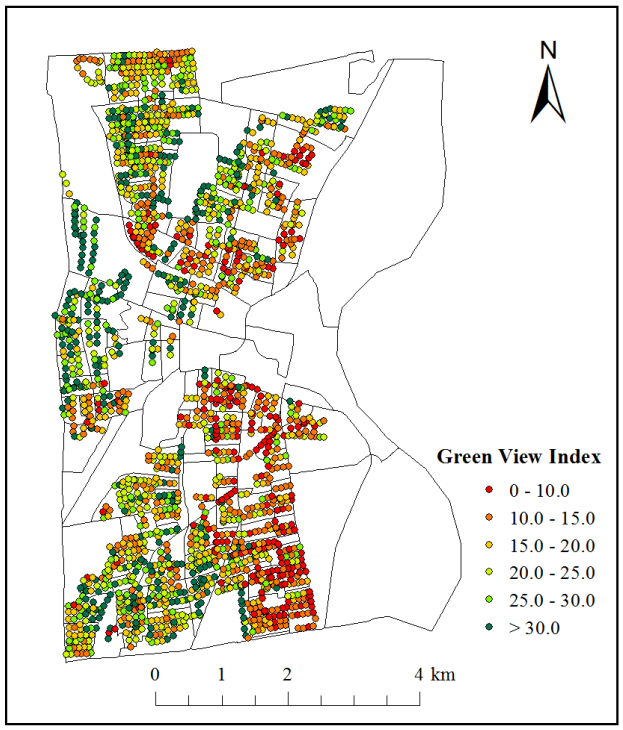


Fig.5.6. The spatial distributions of non-green GSV sites in Hartford, CT (a) and Boston, MA (b).

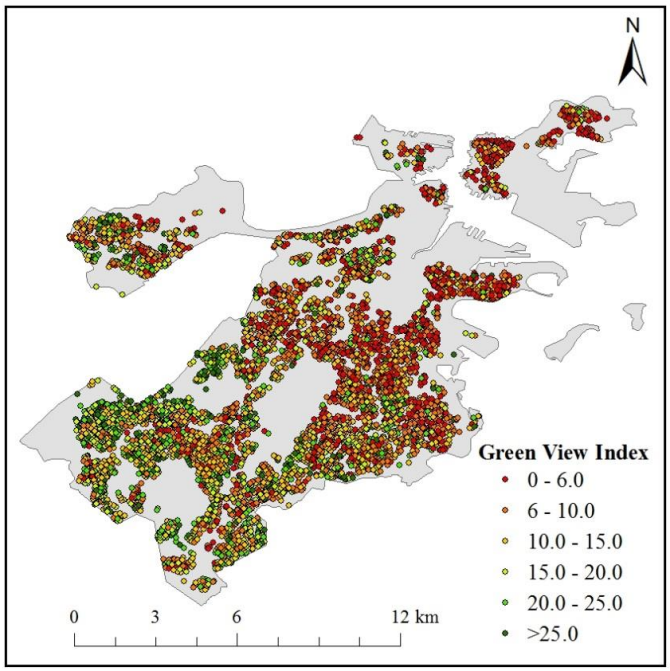
The spatial distributions of the MGVI in Hartford and Boston are shown in **Fig. 5.7**. Among all chosen sample sites in Hartford (**Fig. 5.7 (a)**), most sites with low MGVI values are located in the east and south of the study area. Sites with high MGVI values are mainly distributed in the west, north, and southwest subareas. In Boston (**Fig. 5.7 (b)**), most sites with high MGVI values are located in the southwest and west regions. Most sites with low MGVI values are located in the eastern region.

Fig. 5.8 presents several sites in Hartford and Boston with different MGVI values. In general, those sites with more street trees tend to have higher MGVI values, and those sites with large-size street trees, multi-layer vegetation, and lawns along streets usually have higher MGVI values. This is not difficult to understand because the MGVI was calculated by 18 GSV images captured in six horizontal and three vertical angles and the lawns and large street trees were

counted.



(a) The spatial distribution of the MGVI in Hartford, CT



(b)The spatial distribution of the MGVI in Boston, MA

Fig.5.7. The spatial distributions of the MGVI in Hartford, CT and Boston, MA.

(a) MGVI values for different sites in Hartford, CT (Li *et al*., 2015c)

(b) MGVI values for different sites in Boston, MA

Fig.5.8. GSV sites in Hartford and Boston with different MGVI values.

5.3 Correlation analysis of MGVI and vegetation characteristics

Different from the previous green metrics, the MGVI indicates the visibility of greenery on the ground. Previous studies show that the green view index values are influenced by the patterns of trees, the size of trees, and the distances between trees and viewers (Yang et al., 2009). In order to better illustrate the difference of MGVI with the previous green metrics, correlation analysis between MGVI and the previous green metrics was conducted in this study. For Hartford, the LiDAR data is not available, therefore, the chosen green metrics were based on a land cover map, which was derived from multispectral remotely sensed imagery. In Boston, because of the availability of the LiDAR data, the canopy height was also considered in the correlation analysis.

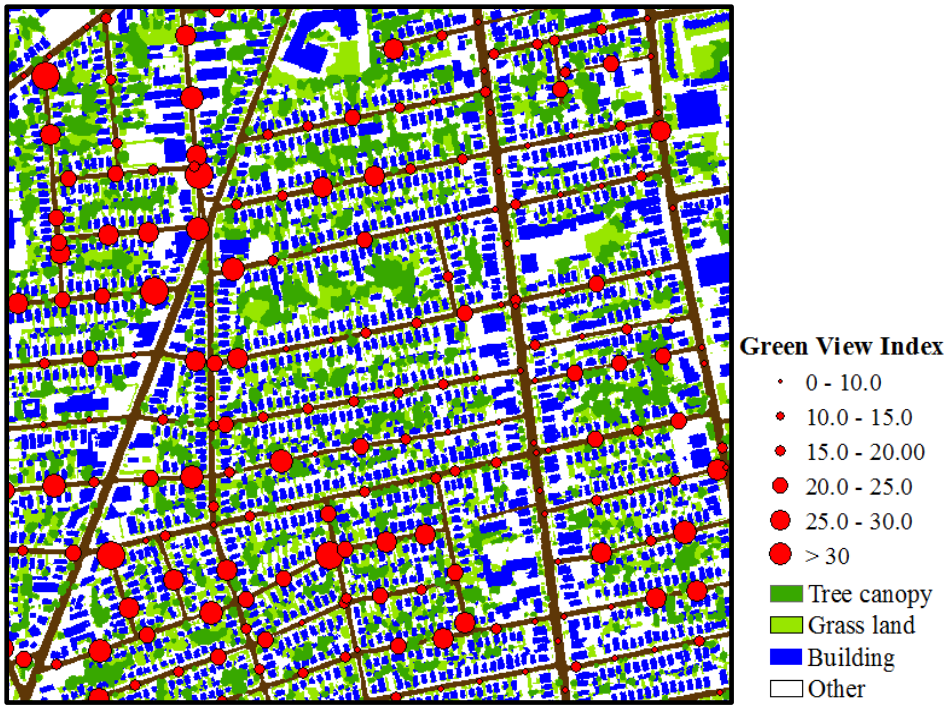


Fig.5.9. The MGVI map and land cover map in a small portion of Hartford. The sizes of the solid dots represent the magnitudes of MGVI values.

Fig. 5.9 shows the distribution of MGVI values and land cover map in a small portion of Hartford, CT. **Table 5.1** provides the Pearson's correlation coefficients between MGVI values and vegetation coverage at different buffer distances around the sample points. From **Table 5.1**, it can be seen that the MGVI has a significantly positive correlation with the canopy coverage in the buffered zones. However, the correlation coefficients decrease with increase of the buffer distances. Compared with the significant correlation between MGVI and canopy coverage, the correlation between the MGVI and lawn coverage is not statistically significant. The correlation analysis results show that the MGVI is significantly influenced by nearby canopy covers, but the correlations decrease with the increase of buffer distances. It is not difficult to understand this phenomenon, because the trees in faraway look smaller than those closer trees. The non-significant correlation between the MGVI shows that the nearby lawns do not influence the MGVI values significantly. This may be because the MGVI used in this study covers 6 horizontal directions and 3 vertical directs (see **section 4.4** for more details). Some lawns are invisible in those GSV images with vertical angles of 0 and 45**.** In addition, most lawns in residential areas are located in the backyard and blocked by the building blocks. This to some extent may further explain the non-significant correlation between the MGVI and the lawn coverage.

Table 5.1. The Pearson's correlation coefficients between MGVI and the canopy characteristics. ACH: Average Canopy Height, PCC: Percentage of Canopy Cover.

Buffer distances	Pearson's correlation coefficients		*N*
	Canopy Coverage	Lawn coverage	
20 m	0.735**	-0.005	2752
40 m	0.694**	0.006	
60 m	0.637**	0.028	
80 m	0.598**	0.052	
100 m	0.575**	0.061	

** Correlation is significant at the 0.01 level (2-tailed).

In Boston, the canopy height was further considered in the correlation analysis with the availability of LiDAR data. **Fig. 5.10** shows the overlap of the MGVI values with the canopy cover map and the canopy height map in a small portion of Boston. For simplicity, the averaged canopy height was calculated as the indicator of height of nearby trees. **Table 5.2** shows the Pearson's correlation coeffiencts between the MGVI values and the canopy characteristics at different buffer distances around the sample points in whole Boston.

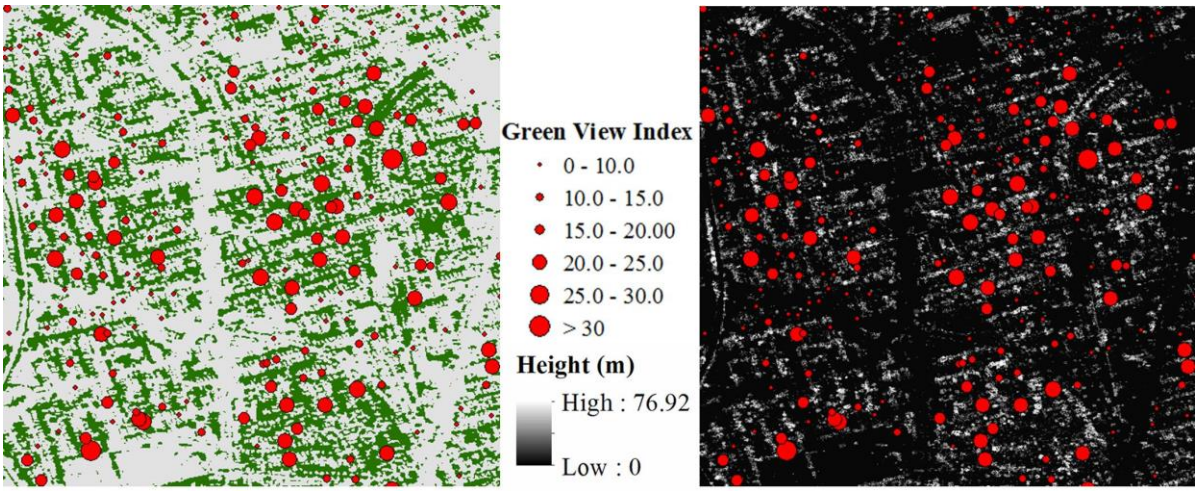


Fig.5.10. The MGVI map and canopy cover map in a small portion of Boston (a) and a Digital Surface Model (DSM) of canopy (b).

Table 5.2. The Pearson's correlation coefficients between the MGVI and the canopy characteristics in Boston, MA. ACH: Averaged Canopy Height, PCC: Percentage of Canopy Cover.

Buffer distances	Pearson's correlation coefficients		*N*
	PCC	ACH	
20 m	0.701**	0.594**	5925
40 m	0.621**	0.555**	
60 m	0.547**	0.507**	
80 m	0.512**	0.482**	
100 m	0.489**	0.466**	

** Correlation is significant at the 0.01 level (2-tailed).

In Boston, the correlation analysis results show that the correlation between the canopy coverage and the MGVI has a similar pattern with the correlation in Hartford. The canopy coverage is strongly correlated with the MGVI, and the correlation becomes weaker as the increase of the buffer distances. The MGVI also has a very significant correlation with the ACH (Averaged Canopy Height) of nearby canopy. And the MGVI has a stronger correlation with the closer canopy covers than those canopies far away.

The MGVI is proposed to measure the neighborhood greenness in terms of the visibility of street greenery. The correlation analysis results in Boston and Hartford show that the MGVI can reflect the amount of tree canopy coverages of the street greenery. The MGVI cannot directly represent the coverages of lawns. Therefore, the MGVI is more suitable for measurement of street trees rather than lawns. Different from the 2D green metrics based on remote sensing data,

the MGVI can also reflect the 3D structural information of street greenery. Those sites with large canopy coverages or higher canopies tend to have large MGVI values. Therefore, the MGVI is a new indicator, which reflect both the canopy coverage and the 3D vertical structure information of the street trees.

5.4 Distribution of different vegetation types of urban greenery

Different types of urban greenery, which are managed differently, provide different kinds of nature experiences to urban residents. The GSV based MGVI is more suitable for the assessement of street greenery rather than other types of urban greenery, such as backyard vegetation and urban parks. This section compares the spatial distributions of the MGVI and different types of urban greenery in Hartford, CT.

5.4.1 Percentage of private yard vegetation coverage

The private yard vegetation mainly locates on private residential parcels and provides benefits directly to property owners. In this study, the percentage of total vegetation coverage (*PerVeg*) and the percentage of tree/shrub coverage (*PerTree*) in residential property parcels were used to represent the distribution of private yard greenery. These two green metrics – *PerVeg* and *PerTree* were calculated by intersecting the residential parcel map and the vegetation cover map for each residential parcel using the following formulas,

$$PerVeg = \frac{Area_{veg}}{Area_{parcel}} \times 100\% \qquad (5.1)$$

$$PerTree = \frac{Area_{tree/shrub}}{Area_{parcel}} \times 100\% \qquad (5.2)$$

where $Area_{veg}$ is the area of total vegetation coverage in a residential parcel, $Area_{tree/shrub}$ is the area of tree/shrub coverage in a residential parcel, and $Area_{parcel}$ is the area of the parcel. These parcel-level green metrics were then aggregated at block group level using their median values.

The residential parcel map was delineated manually based on a parcel map and the land cover map by checking Google Map and Google Street View. **Fig. 5.11** shows the satellite image, the land cover map, and the residential parcel map of a small portion of the study area.

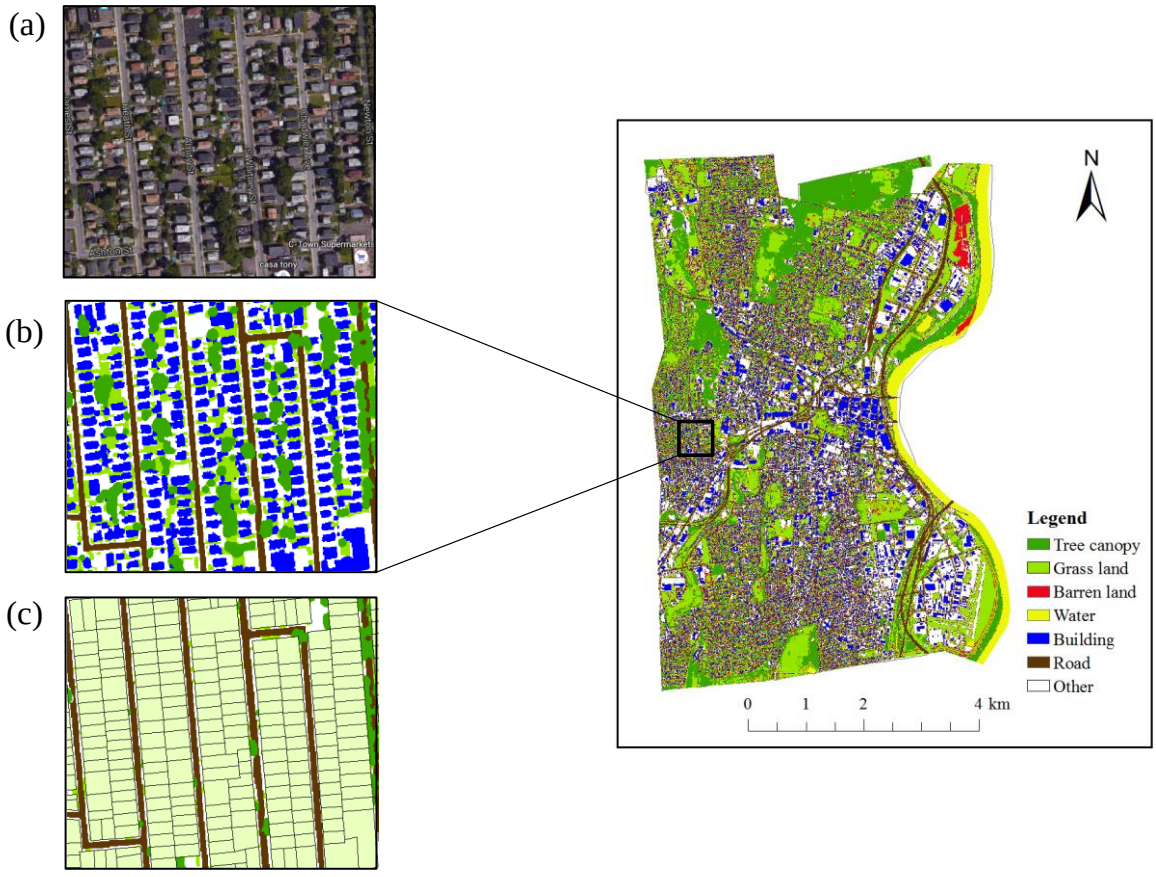


Fig.5. 11. The vegetation in private yards, shown by (a) a satellite image from Google Map, (b) a land cover map derived from the remotely sensed data, (c) a residential property parcel map (Li et al., 2016b).

5.4.2 Proximity to urban parks

Based on previous studies, I used a buffer analysis method to measure park's accessibility in this study. I used a 400-meter buffer zone around each park, a distance most people are willing to walk to an urban park (Boone et al., 2009; Lindsey et al., 2001; Wolch et al., 2005; Leslie et al., 2010). The proportion of residential parcels in the buffer zones for all block groups was further used as the indicator of proximity to urban parks at the block group level. If the centroid of a residential parcel is in the buffer zone of an urban park, then this parcel will be treated as in the service area of the park.

The urban park map in Hartford was obtained from the Hartford Open Data website (https://data.hartford.gov/). The small cemeteries, monuments, and other small green spaces that should not be defined as parks were removed by checking Google Map and Google Street View, since those parks are too small to offer significant benefits to residents.

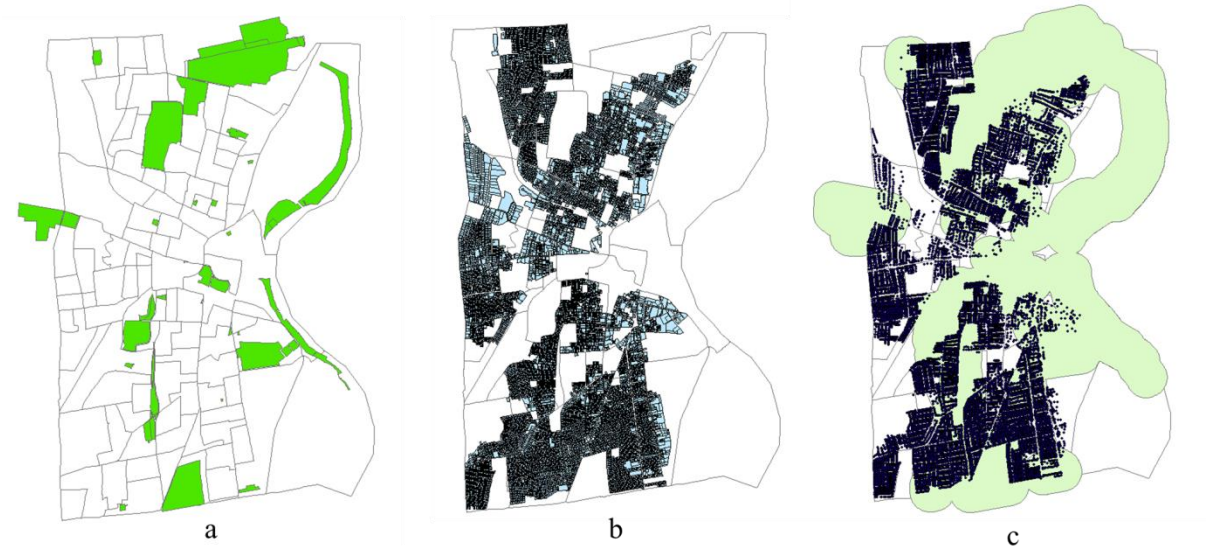


Fig.5.12. The overlap of the urban park map and the residential parcel map in Hartford, CT, (a) the urban park map, (b) and the residential parcel map, (c) the overlap of the buffer map of urban parks and the residential parcel map.

5.4.3 Distributions of different types of vegetation in Hartford, CT

Fig. 5.13 shows the spatial distributions of the aggregated green metrics at block group level, which were used to indicate the street greenery (**Fig. 5.13a**), proximity to urban parks (**Fig. 5.13b**), percentage of total yard vegetation coverage (**Fig. 5.13c**), and percentage of yard tree/shrub coverage (**Fig. 5.13d**), respectively. The street greenery (**Fig. 5.13a**) has a very different distribution compared with other types of urban greenery in Hartford. In the street greenery map, block groups in the west, northwest, and southwest of the study area have higher MGVI values than block groups in the east. That means neighborhoods in west, southwest and northwest have more street greenery than the eastern region. The eastern and middle regions have closer proximity to urban parks than the other regions (**Fig. 5.13b**). Yard vegetation and yard trees/shrubs (**Figs. 5.13c** and **d**) are more abundant in the north than in the south.

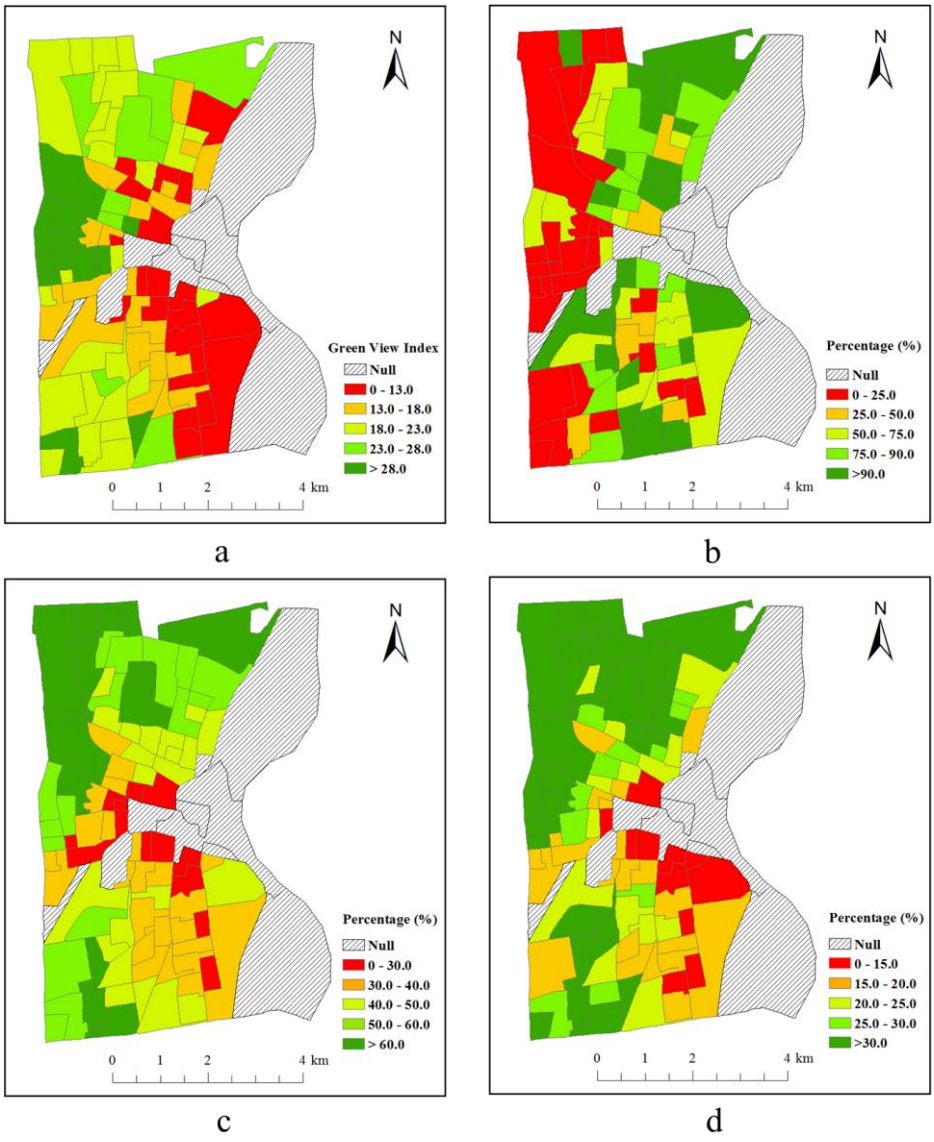


Fig.5.13. Green metrics mapped at the block group level: (a) the MGVI values, (b) proximity to urban parks (proportion of residential parcels in 400m buffer zones of urban parks), (c) percentage of yard vegetation coverage, and (d) percentage of yard tree/shrub coverage (Li et al., 2016b).

5.5 Summary

The comparision analysis of the MGVI with nearby canopy coverage, lawn coverage, and canopy height proves that the MGVI measures street tree canopy coverage and the vertical structural information of street trees. The size and height of street trees influence the MGVI positively. Those neighborhoods with more street trees, larger canopy covers, and higher

canopies tend to have larger MGVI values. In addition, the GSV based MGVI is more suitable to indicate the distribution of street greenery rather than backyard vegetation or vegetation in urban parks.

Chapter 6 Environmental inequity analyses

Different from previous studies using the canopy cover or vegetation indices derived from remotely sensed data (Grove et al., 2006; Mennis, 2006; Landry and Chakraborty, 2009; Leslie et al., 2010; Jenerette et al., 2013), in this study, the MGVI was used to investigate the distribution of urban street greenery among different socio-economic and racial/ethnic groups. The objective of this study is to test the hypothesis that environmental disparities in terms of street greenery are linked to the racial/ethnic makeup and socioeconomic status of residents. Chapter 4 gives details about the workflow for calculating the MGVI based on Google Street View images. The calculated MGVI values represent the spatial distribution of street greenery at site level. In order to make the MGVI maps and census data comparable, the site level MGVI maps were further aggregated at census tract level. The median MGVI values were summarized by census tract. Sites on the borders of census tract were not counted in order to eliminate the neighborhood effect. Bivariate correlation analysis and regression models were then used to investigate environmental inequity in terms of street greenery in different cities.

6.1 Extraction of social variables from census data

Based on previous studies (Landry and Chakraborty, 2009; Huang et al., 2011; Pham et al., 2012), eight social variables at the census tract level were selected to represent the racial/ethnic and socio-economic status of residents in this study (**Table 6.1**). These eight chosen variables include per capita income, proportion of African Americans, proportion of Asians, proportion of non-Hispanic Whites, proportion of Hispanics, proportion of owner-occupied units, proportion of people with bachelor or higher degrees, and proportion of people without high school degree.

Economic status affects people's capability to improve their living space, which may further influence their physical living environments. There are many developed variables to indicate the economic status of residents, such as household income (Landry and Chakraborty, 2009; Heynen, 2006), per capita income (Li et al, 2015b), and proportion of households with income below the poverty line (Huang et al., 2011). In this study, the per capita income rather than household income was chosen as the indicator of a resident's economic status, considering the fact that household income does not consider the household size. The race/ethnicity variables include proportion of African Americans, proportion of Asians, proportion of Hispanics, and the proportion of non-Hispanic Whites. Previous studies have showed that vegetation coverage is associated with people's education levels. Therefore, in this study, two educational variables (the proportion of people with bachelor or higher degrees and the proportion of people without high school degree) were included in the analysis. In order to control the impact of built environment on the distribution of street greenery, a built environment variable (median building age) was considered (Grove et al., 2006; Pham et al., 2012; Landry and Chakraborty, 2009). In addition,

two age variables (proportion of people under 18 years of age and proportion of people older than 65 years of age) were also included in the analysis.

Table 6.1.

The chosen social variables from American Community Survey (ACS) census data

Category	Variables
Economic status	Per capita income
Education	Proportion of people without high school degree Proportion of people with bachelor or higher degrees
Lifestyle	Proportion of owner-occupied units
Built environment	Median building age
Race/Ethnicity	Proportion of non-Hispanic Whites Proportion of Hispanics Proportion of Asians Proportion of African Americans
Age	Proportion of people under 18 years of age Proportion of people older than 65 years of age

In 2010, the American Community Survey (ACS) replaced the decennial census as the sole national source of demographic and economic data for small areas like block group and census tract (Spielman and Singleton, 2015). Although the questions in the survey for ACS data collecting are very similar to those for the 2000 decennial census, there are some important differences between the collections of these two datasets. While the decennial census has provided a snapshot of the U.S. population once every 10 years, the ACS is kind of “rolling survey” (Spielman and Singleton, 2015), and has been described as a "moving video image, continually updated to provide much needed data about our nation in today's fast-moving world”

(Cooper, 2005). Compared with the decennial data, the ACS is based on a relatively smaller sample size.

The Google Street View, which was firstly launched in 2007, keeps updating to represent what the most recent streetscapes look like. To some extent, the ACS data and Google Street View use the same method to represent the real world, which both can be described as "moving video image, continually updated to provide much needed data". Therefore, the ACS data was used as the data source for extraction of social variables in this study. Although the ACS data is criticized by many users for its uncertainty and large margin of error in small areas, the ACS data is still used in this study since it is the only national scale census data, which includes all social variables in this study.

Although block group is the finest-level areal unit for most social variables used in ACS data, in order to increase the accuracy of both MGVI map and the census data, the census tract was used as the common boundary for all geospatial operations in this study. This is because choosing a large geographical unit means more GSV sample sites in each unit, which may further help to increase the reliability of estimation at areal units. The census data in census tract usually has smaller margin of error compared with block group level. In addition, since this study focused on residential housing units, census tracts located in non-residential areas were excluded from the analysis.

In order to keep the time consistency between the census data and the GSV images, ACS census data of different years was chosen for different cities correspondingly based on the time information of the GSV images in different cities.

6.2 Statistical Analysis

To determine whether the street greenery has an unequal distribution across different racial/ethnic and socio-economic groups in the study areas, statistical analyses were performed between the MGVI and the selected social variables. The statistical analyses were conducted in three steps. Bivariate correlation analysis was first used to explore the correlations between the MGVI and each of the social variables at census tract level for each of the selected cities. Second, ordinary least square (OLS) multivariate regression analyses were conducted to model the associations between dependent variable (the MGVI) and independent variables (social variables) in each of the selected city. Since this study focuses on investigating the environmental inequities among different socio-economic groups, the confounded variables of racial/ethnic variables and economic variables were excluded from the further multivariate regression analyses. The finalist variables in the regression models include, proportion of African American, proportion of Asians, proportion of Hispanics, per capita income, proportion of owner-occupied units, and median building age.

Finally, global Moran's *I*-statistics were used to analyze the spatial autocorrelation of residuals in regression models to determine whether the regression results were spatially biased (Landry and Chakraborty, 2009). In Moran's *I* analyses, Euclidean distances among centroids of census tracts were used to generate the spatial weight matrix, which were used to represent the strength of spatial interactions. When spatial dependence was detected in the residuals of an OLS model, a spatial regression model was then conducted to include an additional term to account

for the spatial autocorrelation.

In geography, the spatial autocorrelation has been a frequently faced problem for the analysis of socioeconomic data (Talen and Anselin, 1998). Geographical features usually are spatially auto-correlated, because geographical features near each other are more similar than those features further away (Tobler, 1970). The spatial dependence could cause spatial dependence of residuals in regression models, which would further violate the assumption of independent observations in traditional OLS multivariate regression models (Landry and Chakraborty, 2009). The simultaneous autoregressive (SAR) model is a spatial regression model to augment the standard linear regression model by incorporating spatial interrelationship structure of the units of analysis (Anselin and Bera, 1998). In SAR models, the spatial weight matrices are used to represent the strength of the interaction of neighboring sites. The spatial weight matrices can be defined in various ways, in this study, due to irregularity of census tract sizes and shapes in different cities, Euclidean distances among census tract centroids were used to generate the spatial weight matrices (Pastor et al, 2005; Landry and Chakraborty, 2009). There are two major SAR models, spatial lag regression model and spatial error regression model, which incorporate the spatial dependence into the dependent variable and error terms, respectively. The spatial lag regression model (SAR_{lag}) assumes that the spatial dependence effect exists in the dependent variables. A pure spatial lag regression model simply consists of a spatially lagged version of the dependent variable, and takes the matrix algebra model form of:

$$y = \rho Wy + \varepsilon \tag{6.1}$$

Where W is the predefined $n \times n$ spatial weighting matrix, y is the observed variable, and ρ is a

spatial autoregression parameter, which typically has to be estimated from the data. SAR_{lag} mixes spatial autoregressive model with standard regression. A first order SAR_{lag} model can be written as,

$$y = \rho Wy + X\beta + \varepsilon \tag{6.2}$$

Where ρWy is the pure spatial autoregression part, and X is the vector of explanatory variables, β is the vector of regression coefficients, and ε is the error term. The mixed regressive-spatial autoregressive model incorporates spatial autocorrelation together with the influence of predictor variables. The mixed model improved the standard OLS model, and the level of improvement is dependent on how well the weighting matrix represents the spatial relation between geographic features of different locations and how well the mixed model represents or explains the source data (de Smith, 2015).

SAR_{err} model is another kind of mixed regressive-spatial autoregressive model. Different from SAR_{lag} model, SAR_{err} model incorporates the spatial effect into the error term,

$$y = X\beta + \varepsilon,$$

$$\text{where} \quad \varepsilon = \lambda W\varepsilon + u \tag{6.3}$$

In this study, the spatial regression models for different cities were implemented in GeoDA (Anselin, 2005). Here, the y is the MGVI, and X represents the selected independent variables. The spatial weight matrix W was calculated as the Euclidean distances among census tract centroids. The spatial models of SAR_{lag} and SAR_{err} were compared and the choice of either spatial lag or error model was determined based on the Lagrange Multiplier and Robust

Lagrange Multiplier tests for each city, which indicate whether the spatial dependence occurs at the error term or the dependent variable (Anselin, 2005).

6.3 Results

6.3.1 Descriptive statistics of the MGVI in major cities

Based on the spectral analysis of the GSV images in different seasons (see section 4.3.1) and the time information of GSV sites in different cities, several cities were excluded from the study. This is because the image analysis in this study is based on the assumption that street greenery will be shown as green in GSV images. The method used in this study is not suitable for those cities with too many sample sites having their GSV images captured in non-green seasons. As an example, **Fig. 6.1** shows the spatial distribution of green GSV sample sites and non-green GSV sample sites in cities of Chicago, IL Milwaukee, WI, Atlanta, GA, and Louisville, KY based on the previous spectral analysis in Chapter 4. It can be seen clearly that these four cities have large percentages of non-green sample sites. In addition, those non-green sample sites are distributed in clusters. Excluding those aggregated non-green sites will decrease the representation of GSV sample sites for the street greenery. Therefore, those cities having too many clustered non-green GSV sample sites may not be suitable in the GSV based street greenery analysis and should be removed from the study areas. By checking the spatial distribution of non-green GSV sites in all study areas, cities of Chicago, Indianapolis, Minneapolis, Cincinnati, Louisville, Milwaukee, Atlanta, Houston, San Antonio, Dallas, Los Angeles, San Francisco, and Seattle were removed from the city list of the study areas (**Table 3.1.1**). In the finally selected cities (**Table 6.2**), most

of the sample sites having their corresponding GSV images captured in green seasons and those non-green GSV sample sites are distributed randomly and evenly in cities. By controlling the spatial distribution of non-green GSV sample sites, the GSV based analysis helps to represent the street greenery in different cities more reasonably without much influence by seasons. In addition, cities without land use maps available were also removed.

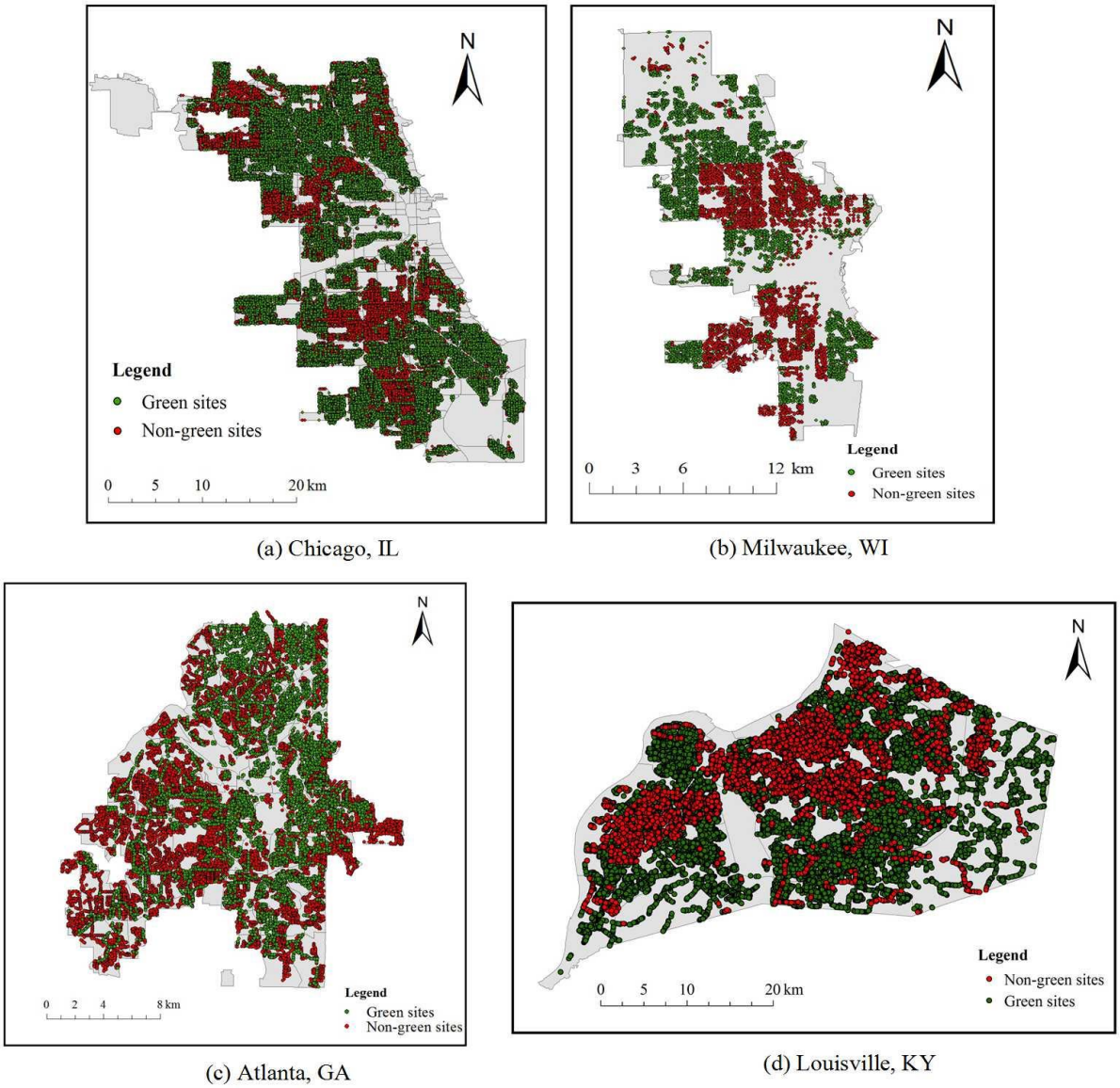


Fig.6.1. The spatial distributions of the green GSV sites and non-green GSV sites in Chicago, Milwaukee, Atlanta, and Louisville**.**

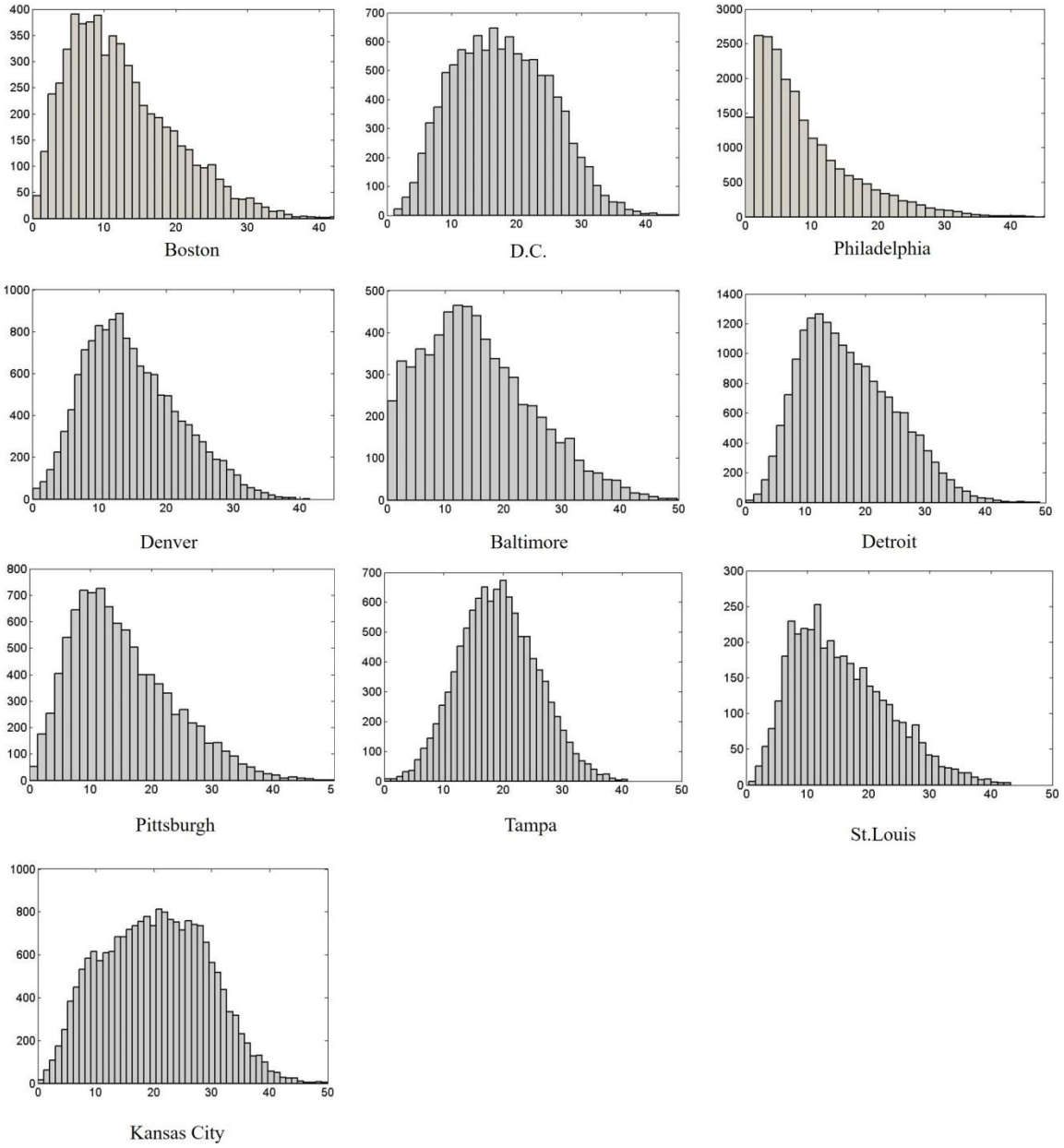


Fig.6.2. The histograms of MGVI values at site level for all studied cities.

The **Fig. 6.2** shows the histograms of MGVI values of the finally selected cities. These MGVI histograms give us an overview of the street greenery level in different cities. For example, Philadelphia has a very skewed MGVI histogram and most GSV sample sites have their MGVI values less than 10. Most of these cities have relatively skewed MGVI histograms,

such as, Boston, Denver, Baltimore, Detroit, Pittsburgh, and St. Louis. Most of the GSV sample sites in these cities have the MGVI values in the range of 5 to 20. However, Cities, like Washington D.C, Tampa, and Kansas City have perfect normal histograms and most GSV sample sites have their MGVI values in the range of 10 to 30 in these three cities. These three cities have greener neighborhoods compared with other cities in term of street greenery. This can be further proved by the relatively larger mean and median site level MGVI values in Washington D.C, Tampa, and Kansas City (**Table 6.2**). On the contrary, with the most skewed MGVI histogram and the lowest site level mean and median MGVI values, the City of Philadelphia has the least street greenery compared with other cities.

Table 6.2
The description of MGVI values in different cities at site level and census tract level.

Cities	**Site level**			**Census tract level**		
	Mean	Median	Std	Min	Max	Std
Boston	12.45	11.14	7.32	2.79	25.47	3.94
Washington D.C	17.88	17.53	7.38	7.01	25.92	4.08
Philadelphia	8.73	6.67	7.11	1.13	31.84	4.34
Denver	14.90	13.87	6.99	2.84	24.44	4.29
Baltimore	15.70	14.38	9.57	2.21	28.26	6.00
Detroit	17.18	16.08	7.75	3.64	26.60	3.68
Pittsburgh	15.20	13.63	8.34	3.66	23.06	3.95
Tampa	19.10	18.97	6.65	7.65	23.82	2.92
St. Louis	15.58	14.28	7.82	7.75	34.70	4.04
Kansas City	20.25	20.28	8.91	3.30	33.57	5.35

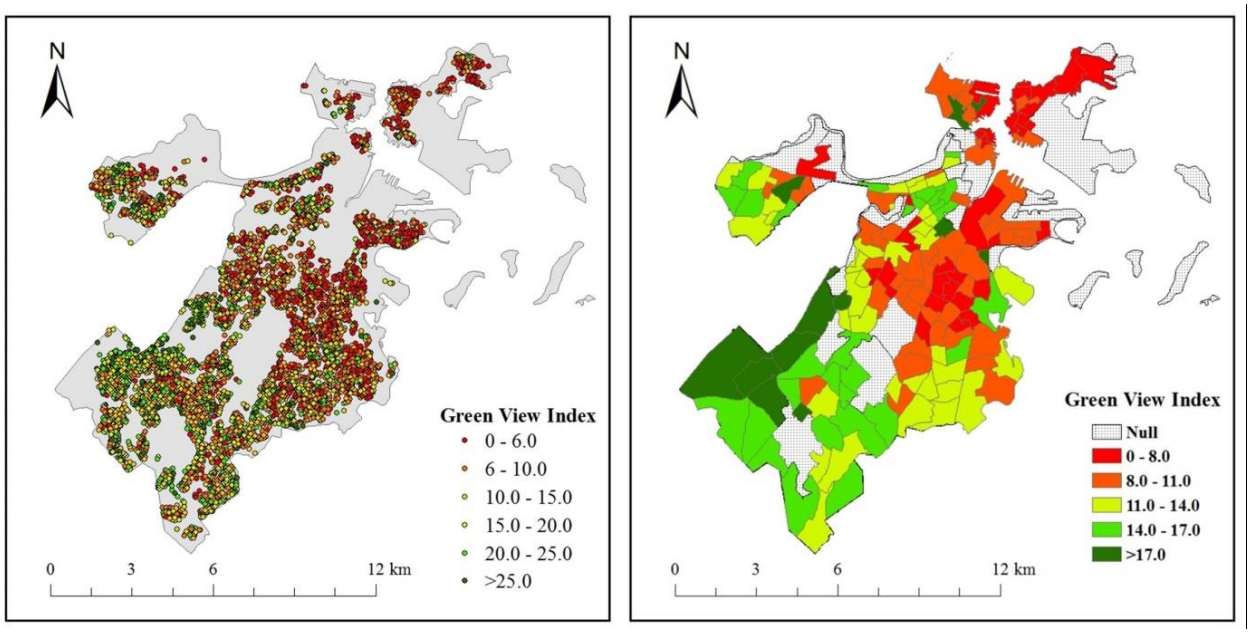
In order to make the MGVI maps and census data comparable, the point level MGVI maps were further aggregated to areal level. The census tract was used as the uniform geographic unit and the median value of the MGVI in each census tract was chosen to represent the MGVI values of the census tracts. Those census tracts with less than three GSV sample sites in it were excluded from further analysis.

Table 6.2 gives the descriptive statistics of MGVI values at census tract level for all chosen cities. In most census tracts of those selected cities, the MGVI values range from 1.13 to 33.57, with standard deviations in a range of 2.92 – 6.00.

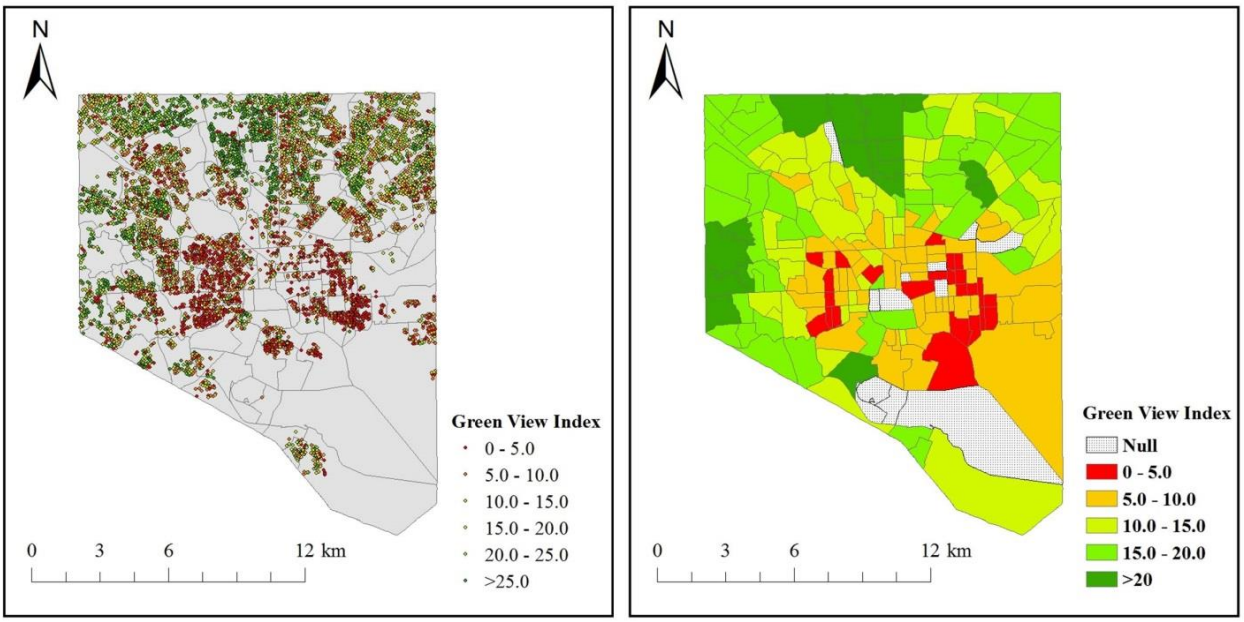
Fig. 6.3 presents the spatial distributions of MGVI values in different cities at both the site level and census tract level. Most cities share a similar pattern that those sites with high MGVI values are located in the periphery regions and those sites with low MGVI values are located in the inner part of cities. The census tract level MGVI maps have similar patterns with the site level maps. This distribution pattern is especially obvious in cities of Boston, Baltimore, D. C, Philadelphia, Pittsburgh, Detroit, and St. Louis. This is not difficult to understand considering the fact that the inner part of a city has less open space for tree planting. Other than this general pattern, different cities also have specific spatial patterns in terms of the distribution of the MGVI values. In Boston (**Fig. 6.3**(a)), most sites with high MGVI values are mostly distributed in the southwest and southern areas. Those sites with low MGVI values are mostly located in the north and the northeastern regions. In Washington D.C and Detroit (**Fig. 6.3 (c, g)**), the western regions have obviously higher MGVI values than the eastern regions. In St. Louis and Kansas City (**Fig. 6.3 (i, j)**), GSV sample sites in south usually have higher MGVI values than sites in

north. However, in some cities, like Denver and Tampa (**Fig. 6.3 (f, h)**), there exists no obvious pattern in terms of the MGVI values.

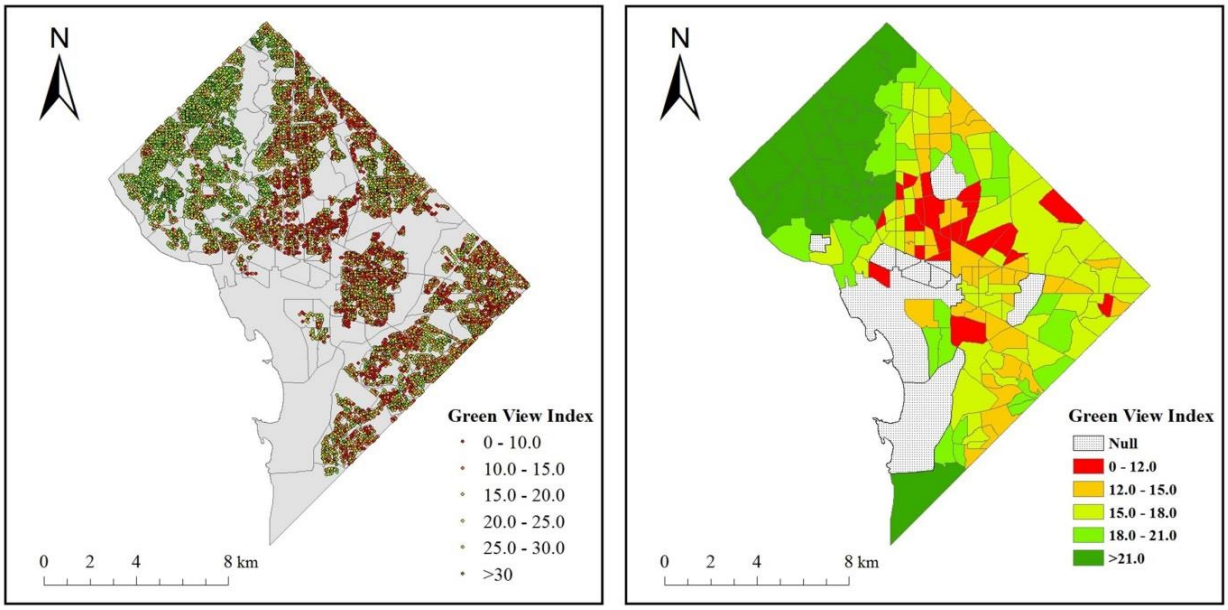
(a) Boston, MA



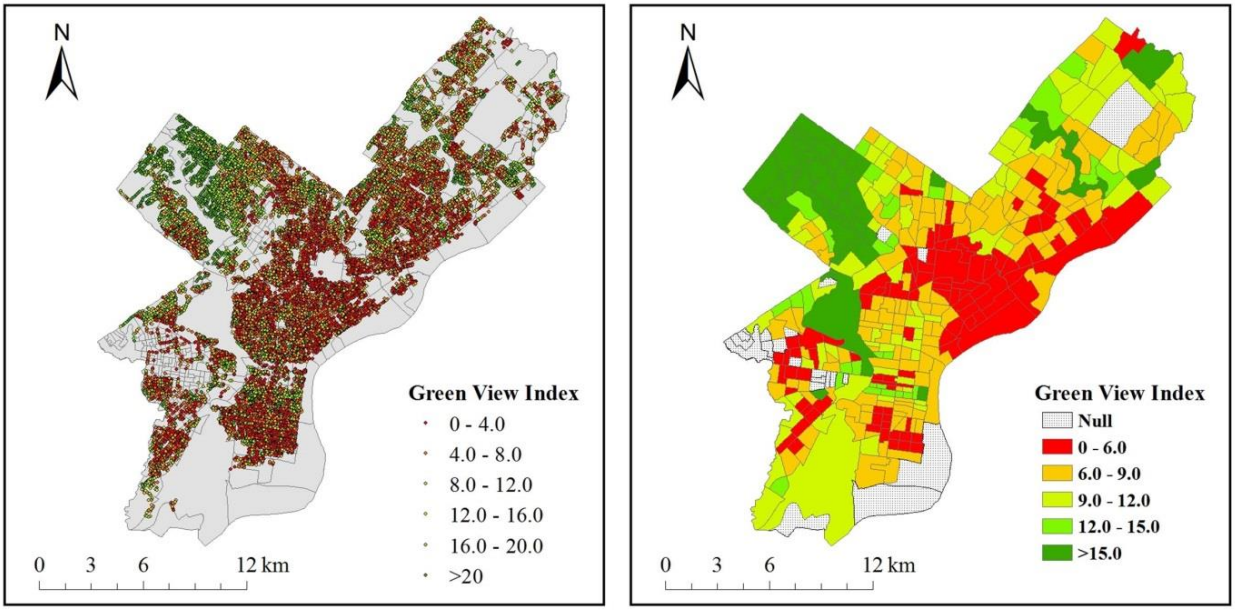
(b) Baltimore, MD



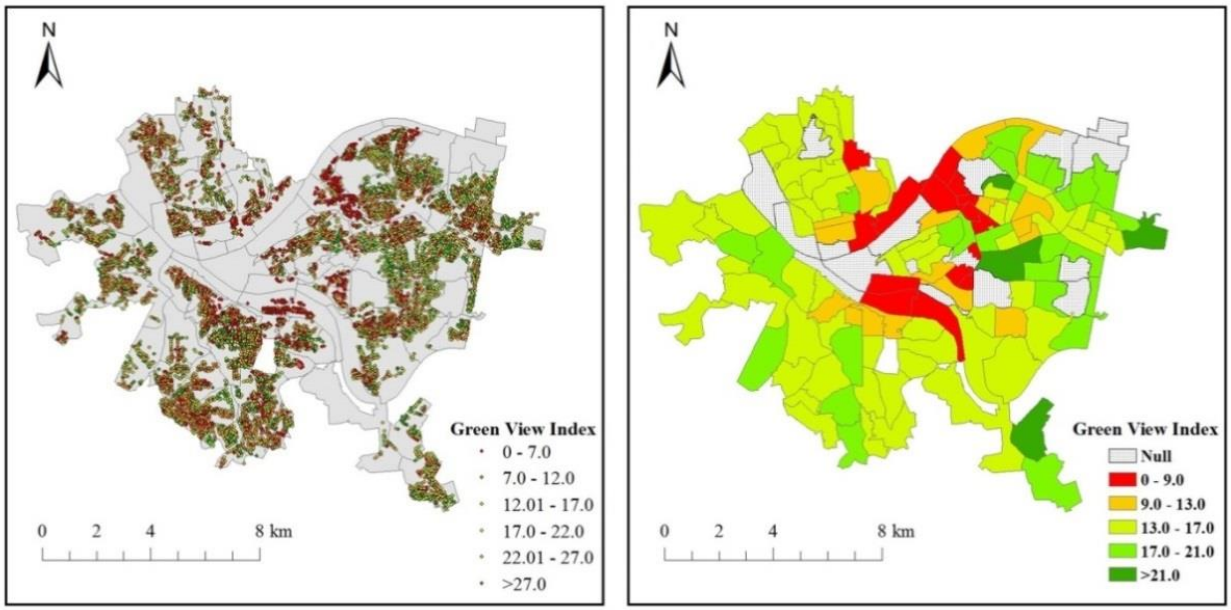
(c) Washington D.C



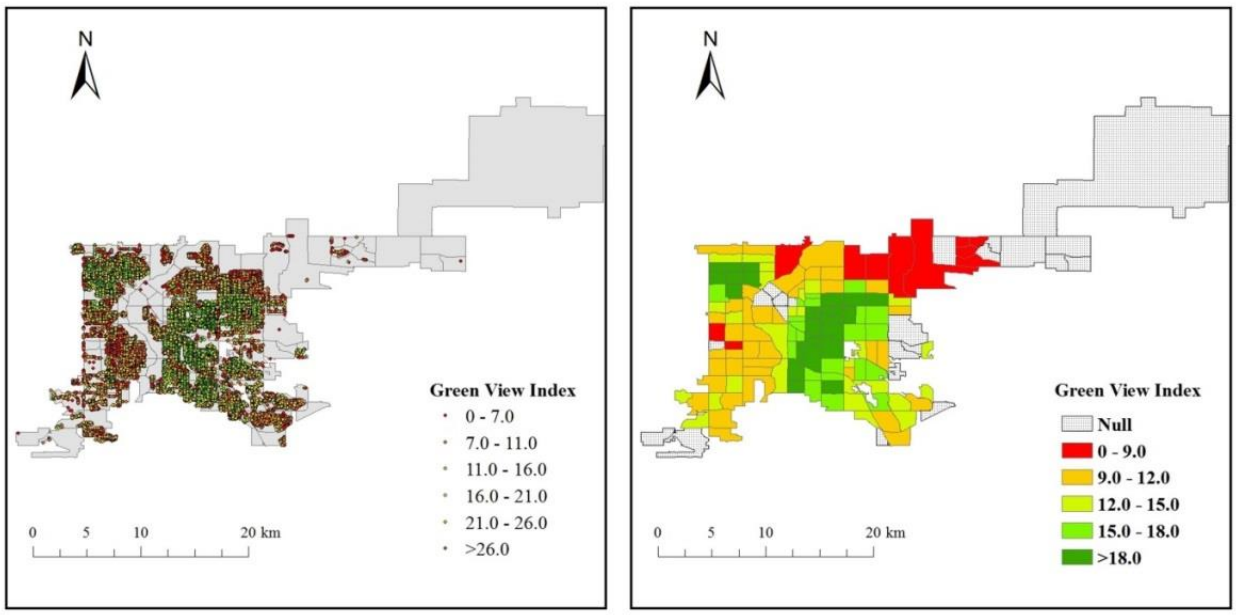
(d) Philadelphia, PA



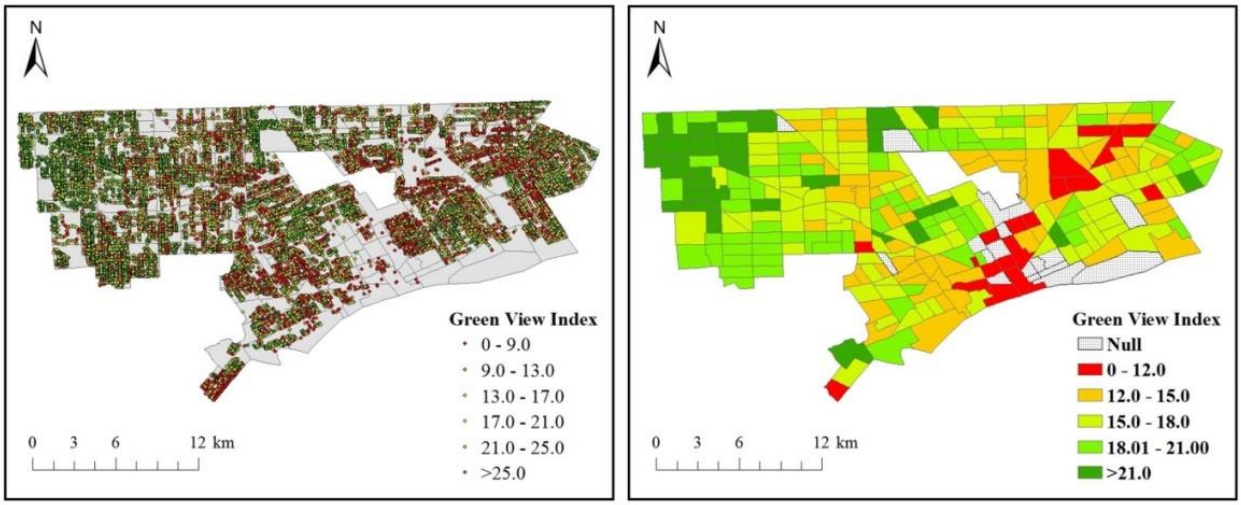
(e) Pittsburgh, PA



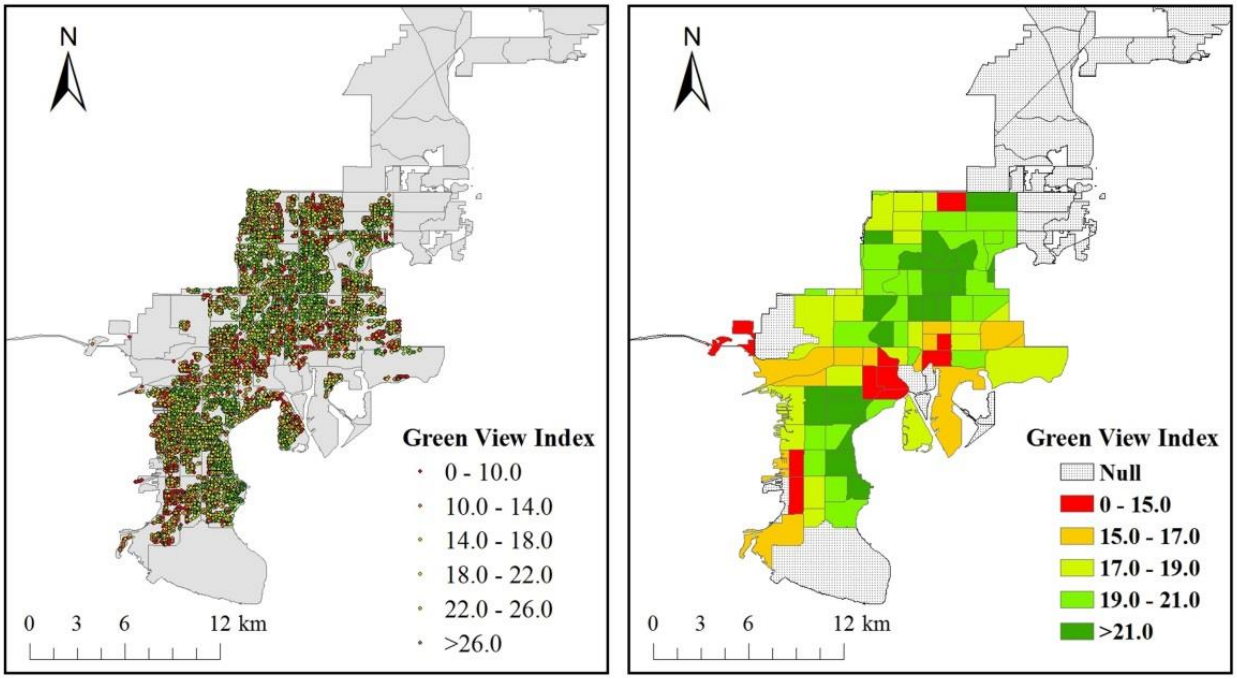
(f) Denver, CO



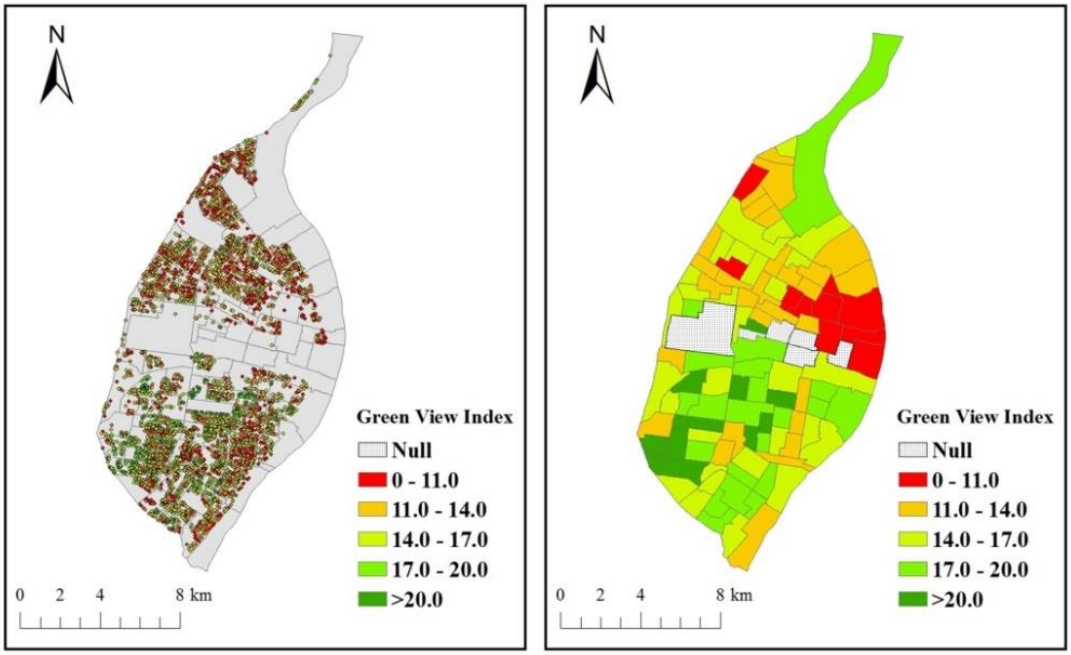
(g) Detroit, MI



(h) Tampa, FL



(i) St.Louis, MO



(j) Kansas City

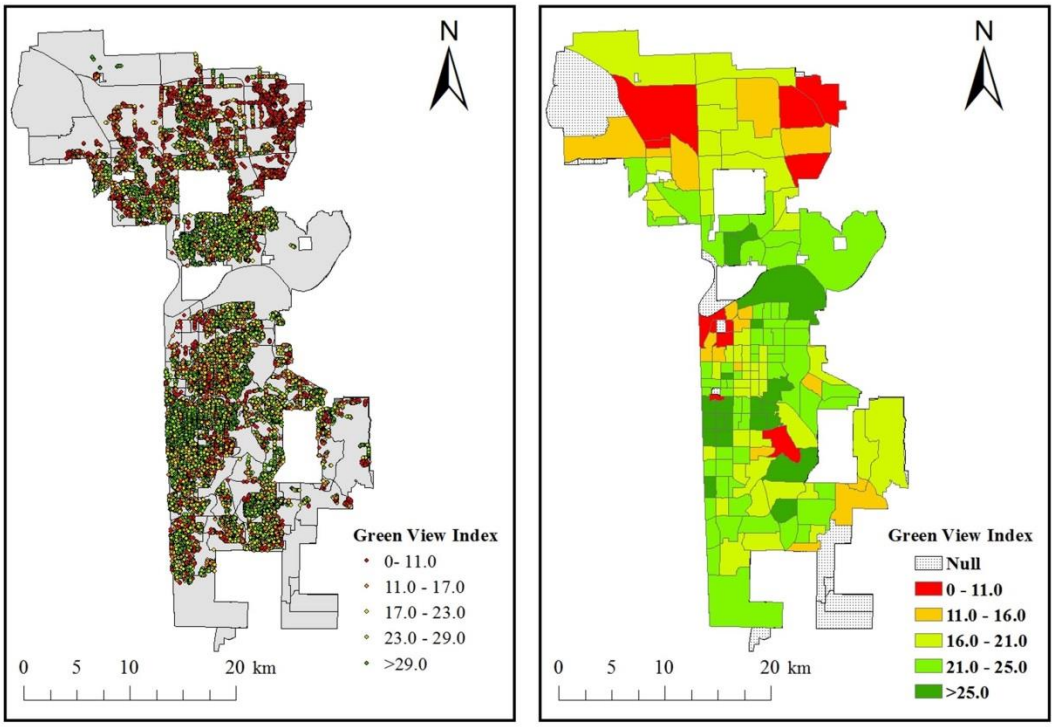


Fig.6.3. The spatial distribution of MGVI maps at site level and census tract level for the chosen cities.

6.3.2 Bivariate analysis results

Fig. 6.4 shows the histograms of GSV time information in different cities. All cities have most of their GSV images taken after 2009. Therefore, the 2010–2014 American Community Survey (ACS) census data was used to calculate the social variables for all cities in this study. Bivariate analyses reveal significant correlations between the MGVI social variables derived from census data (**Table 6.3**). However, the signs of the correlation coefficients and the significant levels were not consistent across different cities.

The correlation between the MGVI and per capita income varies in different cities, however, the sign of the correlation is consistently positive. In cities of Baltimore, Washington D.C, Philadelphia, St. Louis, Denver, the per capita income is positively and significantly correlated with the MGVI. However, in Boston, Pittsburgh, Detroit, and Tampa, there exists no significant correlation between per capita income and the MGVI.

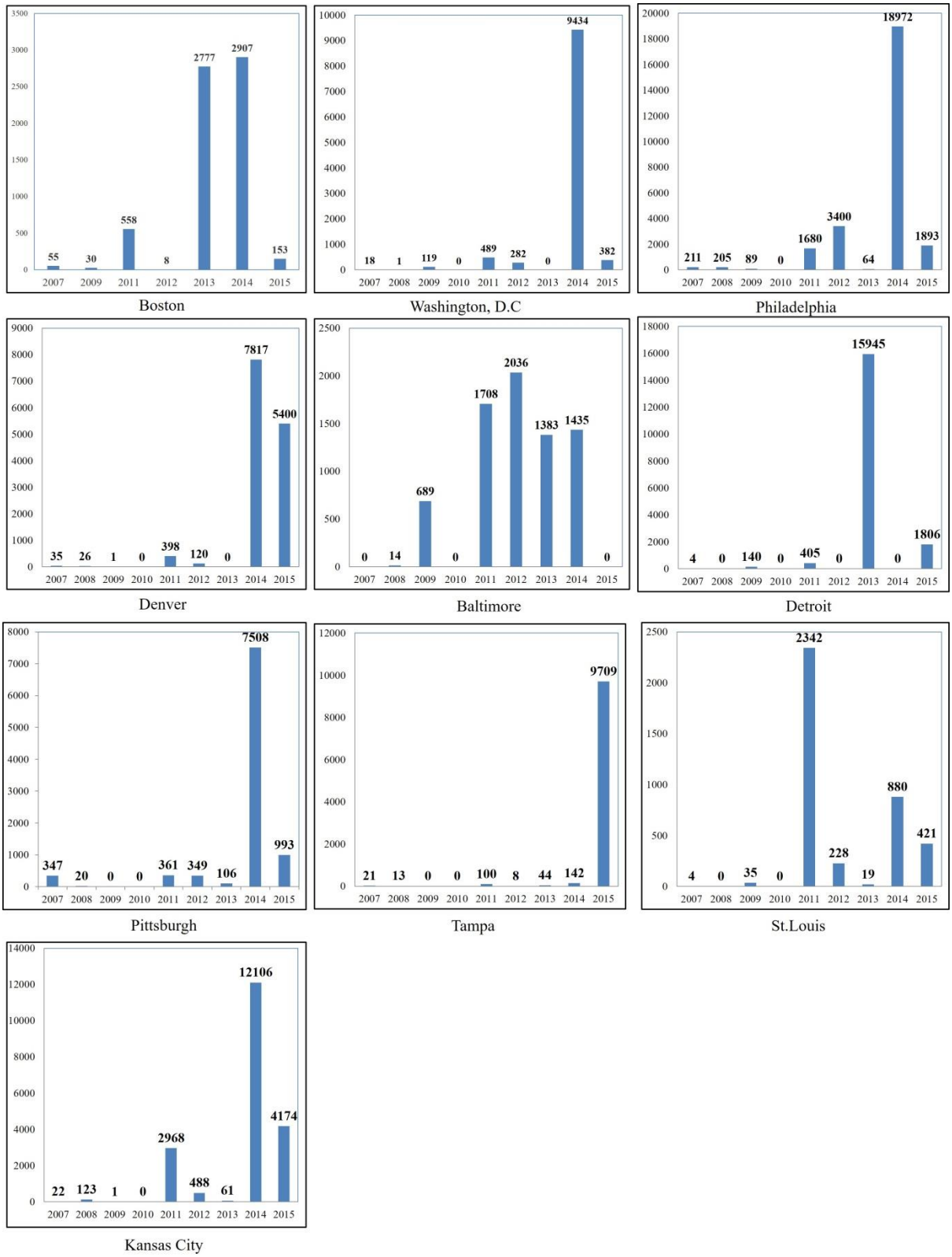


Fig.6.4. The statistics of the GSV image time information in different cities.

There exist relatively consistent correlations between the educational variables (proportion of people without high school degree and proportion of people with bachelor or higher degrees) and MGVI. Consistent across all cities except Pittsburgh, Tampa, and Kansas City are significantly positive correlations between the MGVI and the proportion of people with bachelor or higher degrees, and significantly negative correlations between the MGVI and the proportion of people without high school degree (**Table 6.3**). In Pittsburgh, Tampa, and Kansas City, the MGVI is not significantly correlated with both of the two educational variables (proportion of people without high school degree and proportion of people with bachelor or higher degrees).

Consistent relationships were also detected between the MGVI and the lifestyle variable (proportion of owner-occupied units) across different cities (**Table 6.3**). Only in Philadelphia and Denver, the correlation between MGVI and the proportion of owner-occupied units is not significant. In the rest of cities (Boston, Baltimore, Washington D.C, Pittsburgh, St. Louis, Detroit, Tampa, Kansas City), the MGVI is positively and significantly correlated with the proportion of owner-occupied units.

Different from the consistent correlations between the MGVI and economic variables, the correlations between the MGVI and racial/ethnic variables vary across different cities. In Boston, Baltimore, Detroit, Tampa, and Kansas City, the MGVI has no significant correlation with the proportion of non-Hispanic Whites. However, in Washington D.C, Philadelphia, St. Louis, and Denver, the proportion of non-Hispanic Whites has a significantly positive correlation with the MGVI at significance level of 0.01. Different from other cities**,** in Pittsburgh, the MGVI is significantly and negatively correlated with the proportion of non-Hispanic Whites at

significance level of 0.05 (r=-0.228, p=0.013).

There are no consistent correlations between the MGVI and the proportion of African Americans across different cities. In Boston, Baltimore, Tampa, and Kansas City, no significant correlation was found between the MGVI and the proportion of African Americans. In Washington D.C, Philadelphia, St. Louis, and Denver, the MGVI has a negative correlation with the proportion of African Americans at significance level of 0.01. However, in Pittsburgh and Detroit, the MGVI is positively correlated with the proportion of African Americans at significance level of 0.05.

No consistent correlation was detected between the MGVI and the proportion of Hispanics in different cities. In Boston, Baltimore, Philadelphia, Denver, the MGVI is significantly and negatively correlated with the proportion of Hispanics at 0.01 significance level. In St. Louis, the MGVI has a significant and positive correlation with the proportion of Hispanics at significance level of 0.05 (r=0.21, p=0.036). No significant correlation was detected between the MGVI and the proportion of Hispanics in Washington D.C, Pittsburgh, Detroit, Tampa, and Kansas City.

The proportion of Asians is not significantly correlated with the MGVI in all cities except St. Louis. In St. Louis, the proportion of Asians has a significant and positive correlation with the MGVI (r=0.36, p=0.000).

Bivariate correlation analyses results show mixed findings about the correlations between the MGVI and the two selected age variables (proportion of people under 18 years of age and proportion of people older than 65 years of age). In Boston, Baltimore, Washington D.C, Detroit, Tampa, and Kansas City, there exists no significant correlation between the proportion of people

under 18 years of age and the MGVI. The proportion of people under 18 years of age is negatively correlated with the MGVI in Philadelphia, St. Louis, and Denver. However, in Pittsburgh, the proportion of people under 18 years of age has a significant and positive correlation with the MGVI. Among the chosen cities in the study, the proportion of people older than 65 years of age generally has a positive correlation with the MGVI. In Boston, Baltimore, Washington D.C, Philadelphia, Pittsburgh, Denver, and Kansas City, the proportion of people older than 65 years of age has a significant and positive correlation with the MGVI. In Detroit, Tampa the MGVI has no significant correlation with the proportion of people under 18 years of age and the proportion of people older than 65 years of age.

As an indicator of the built environment of neighborhoods, the median building age is not consistently correlated with the MGVI in this study. In cities of Boston, Pittsburgh, St. Louis, and Detroit, there is no significant correlation between the MGVI and the median building age. In cities of Baltimore and Philadelphia, the MGVI is significantly and negatively correlated with the median building age at significance level of 0.01. The MGVI has a significant and positive correlation with the median building age in cities of Washington D.C, Denver, Tampa, and Kansas City.

Table 6.3. Pearson's correlation coefficients (r) between the MGVI and the selected social variables

City	Boston (N=156)	Baltimore (N=188)	D.C (N=169)	Philadelphia (N=360)	Pittsburgh (N=119)	St.Louis (N=100)	Denver (N=121)	Detroit (N=280)	Tampa (N=88)	Kansas City (N=147)
Per capita income	0.157 (0.050)	0.212** (0.004)	0.445** (0.000)	0.457** (0.000)	0.139 (0.132)	0.478** (0.000)	0.652** (0.000)	0.096 (0.111)	0.050 (0.643)	0.025 (0.765)
Proportion of people without high school degree	-0.355** (0.000)	-0.417** (0.000)	-0.392** (0.000)	-0.459** (0.000)	-0.024 (0.793)	-0.507** (0.000)	-0.636** (0.000)	-0.258** (0.000)	-0.128 (0.235)	0.035 (0.677)
Proportion of people with bachelor or higher degrees	0.234** (0.003)	0.201** (0.006)	0.333** (0.000)	0.515** (0.000)	0.051 (0.579)	0.544** (0.000)	0.734** (0.000)	0.145* (0.015)	0.034 (0.754)	-0.001 (0.990)
Proportion of owner-occupied units	0.321** (0.000)	0.385** (0.000)	0.243** (0.001)	0.041 (0.438)	0.384** (0.000)	0.203* (0.043)	0.114 (0.213)	0.274** (0.000)	0.349** (0.001)	0.176* (0.033)
Median building age	-0.020 (0.800)	-0.209** (0.004)	0.161* (0.037)	-0.262** (0.000)	-0.137 (0.139)	0.130 (0.198)	0.397** (0.000)	-0.030 (0.616)	0.261* (0.014)	0.449** (0.000)
Proportion of non-Hispanic Whites	0.156 (0.051)	0.085 (0.244)	0.369** (0.000)	0.332** (0.000)	-0.228* (0.013)	0.538** (0.000)	0.726** (0.000)	-0.030 (0.614)	0.084 (0.438)	-0.058 (0.489)
Proportion of Hispanics	-0.277** (0.000)	-0.196** (0.007)	-0.022 (0.775)	-0.292** (0.000)	-0.125 (0.177)	0.210* (0.036)	-0.604** (0.000)	-0.107 (0.074)	0.016 (0.884)	0.088 (0.290)
Proportion of Asians	0.085 (0.293)	0.129 (0.077)	0.116 (0.132)	-0.046 (0.379)	-0.008 (0.934)	0.362** (0.000)	0.087 (0.342)	-0.109 (0.069)	-0.161 (0.135)	-0.012 (0.890)
Proportion of African Americans	-0.022 (0.783)	-0.061 (0.407)	-0.329** (0.000)	-0.176** (0.001)	0.222* (0.015)	-0.567** (0.000)	-0.375** (0.000)	0.119* (0.046)	-0.066 (0.539)	0.025 (0.764)
Proportion of people under 18 years of age	-0.004 (0.960)	-0.030 (0.685)	-0.014 (0.854)	-0.368** (0.000)	0.432** (0.000)	-0.440** (0.000)	-0.579** (0.000)	-0.077 (0.200)	0.110 (0.309)	0.038 (0.646)
Proportion of people older than 65 years of age	0.323** (0.000)	0.278** (0.000)	0.353** (0.000)	0.328** (0.000)	0.272** (0.003)	-0.093 (0.357)	0.216* (0.017)	-0.052 (0.386)	0.004 (0.973)	0.221** (0.007)

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

6.3.3 Multivariate Analyses

In general, the multivariable regression models investigate environmental inequities in terms of street greenery by controlling socio-economic variables other than environmental justice variables (race/ethnicity, and income). In regression models, the environmental justice variables were selected as the independent variables, and other socio-economic variables were selected as the cofounding variables. Results of regression models in different cities, such as coefficients, *z*-values, and significance levels are presented in **Table 6.4**.

In Boston, the OLS regression model shows that there is no significant association between the MGVI and the per capita income. The MGVI increases significantly with the proportion of Asians. No significant association was found between the MGVI and other minority variables (proportion of African Americans and proportion of Hispanics). The proportion of owner-occupied units is significantly and positively associated with the MGVI. Significant Moran's *I* value (Moran's $I = 0.23$, z score $=5.54$) shows that residuals of the OLS regression model suffer from a significant spatial autocorrelation. Therefore, the spatial lag regression model (SAR_{lag}) was further used to investigate the associations between the MGVI and independent variables. Same as the OLS regression model, the SAR_{lag} regression model results show that the MGVI has no significant association with the per capita income, proportion of Hispanics, and proportion of African Americans. In the SAR_{lag} model, the proportion of owner-occupied units is still significantly and positively associated with the MGVI, which further proves that the census tracts with higher proportion of owner-occupied units tend to have

higher MGVI values or more street greenery. However, different from the significant association between the MGVI and proportion of Asians in the OLS model, in the SAR_{lag} model, the MGVI is not significantly associated with the proportion of Asians after controlling the spatial autocorrelation.

In Washington D.C, the OLS regression model explains 20% of the variation in MGVI changes in 169 census tracts. Diagnostics for spatial dependence of the residuals show that the Moran's *I* value was significant (Moran's $I = 0.46$, *z* score =17.37). This means that the OLS regression suffers from spatial autocorrelation in terms of the residuals. The SAR_{lag} was then deployed to conduct a further analysis of the relationship between the MGVI and the chosen independent variables. The higher R^2 for SAR_{lag} than that for OLS model suggests the improved goodness of fit for the SAR_{lag}. The significant value of the spatial lag coefficient, rho (coeff=0.84, *z* score=15.42), for SAR_{lag} indicates there is a strong spatial dependence in the MGVI map. For both the OLS and SAR_{lag} models, the MGVI values increase significantly with the per capita income ($p<0.01$). The significantly positive coefficient of the per capita income indicates that those census tracts with higher per capita income have more street greenery. Although in the bivariate analysis the MGVI is negatively correlated with the proportion of African Americans, there is no significant association between the MGVI and the proportion of African Americans in both the OLS model and SAR_{lag} model after controlling other social variables and spatial autocorrelation. Similar with the bivariate analysis results, both the proportion of Hispanics and the proportion of Asians are not significantly associated with the MGVI in regression models.

Philadelphia is located in the similar climate zone with Boston and Washington D.C. The

OLS regression model shows that the MGVI is significantly associated with the per capita income, proportion of Asians, and proportion of Hispanics. The MGVI increases significantly with the per capita income, and decreases with the proportion of Asians and proportion of Hispanics. No significant association between the MGVI and the proportion of African Americans was detected in the OLS regression model. In addition, the proportion of owner-occupied units is not significantly associated with the MGVI. The SAR_{lag} model was further applied to investigate the associations between the MGVI and independent variables because of the significant spatial autocorrelation in terms of residuals in the OLS regression model. After incorporating the spatial dependence, the association between the MGVI and the proportion of Hispanics becomes insignificant. The MGVI is still positively and significantly associated with the per capita income, and negatively associated with the proportion of Asians.

The situation in Baltimore is very similar with Boston. The OLS regression result shows that the MGVI is significantly associated with the proportion Asians and proportion of owner-occupied units. The MGVI increases significantly with increases of the proportion of Asians and proportion of owner-occupied units. Since there is a significant spatial dependence of residuals in OLS regression model (Moran's $I = 0.26$, z score = 16.68), the SAR_{lag} model was also applied. After controlling the spatial dependence, the MGVI is still significantly associated with the proportion of Asians and the proportion of owner-occupied units and the MGVI increases significantly with increases of the proportion of Asians and the proportion of owner-occupied units. Both the proportion of Hispanics and proportion of African Americans are not significantly associated with the MGVI in both OLS and SAR regression models. The per

capita income is not a significant contributor of the MGVI in Baltimore, since there is non-significant association between the MGVI and per capita income in both OLS and SAR regression models.

Denver is in the semi-arid climate zone, and maintaining street greenery is more expensive than humid regions. The associations between the MGVI and social variables are different compared with other cities in humid climate zones. The MGVI is significantly associated with the per capita income, median building age, proportion of Hispanics, and proportion of African American in both OLS and SAR_{lag} models. The MGVI increases significantly with the per capita income and median building age, and decreases significantly with the proportion of African Americans and the proportion of Hispanics. This means that neighborhoods with a higher proportion of African Americans or higher proportion of Hispanics tend to have significantly less street greenery. There is no significant association between the MGVI and the proportion of owner-occupied units in Denver. The MGVI may have less to do with the home ownership in Denver than it has with income and proportion of minorities. The proportion of Asians is not significantly associated with the MGVI in both OLS regression model and SAR_{lag} regression model.

In Detroit, a significant spatial dependence of residual (Moran's $I = 0.51$, z score =13.53) was detected in OLS model. Thus, the spatial regression model was used to investigate the associations between the MGVI and independent variables. The high pseudo R^2 (0.51) for the SAR_{lag} model suggests a better model fit compared with the OLS model. Bivariate correlation analysis result shows that the MGVI has a significantly positive correlation with the proportion

of African Americans (p<0.05), and has no significant correlation with the proportion of Asians and Proportion of Hispanics. However, in both the OLS model and SAR_{lag} model, the MGVI is not significantly associated with all racial/ethnic variables (proportion of African Americans, proportion of Asians, and proportion of Hispanics). The regression results show that the significant correlation between the MGVI and the proportion of African Americans may be spurious, and the apparent negative relationship disappears when per capita income, proportion of owner-occupied units and spatial dependence were taken into account. The regression results also show that there exists no significant association between the MGVI and per capita income. In both OLS and SAR_{lag} regression models, the significant and positive coefficients of proportion of owner-occupied units show that neighborhoods with higher owner-occupied units tend to have more street greenery in Detroit. This means that the MGVI may have less to do with the issue of race/ethnicity and per capita income than it does with the issue of home ownership in Detroit.

Non-significant spatial dependence of residual (Moran's I = 0.04, z score =1.68) was detected in the OLS model in Pittsburgh, therefore, only the OLS model was used to investigate the association between the MGVI and independent variables. The OLS regression model explains 44% of the variation in MGVI changes in 120 census tracts. The MGVI is significantly and positively associated with per capita income, proportion of African Americans, proportion of Asians, and proportion of owner-occupied units. There is no significant association between the MGVI and proportion of Hispanics. The built environment variable has a weakly significant association with the MGVI.

Tampa is in the tropical and humid climate zone, tree planting and growth are supposed to be

much easier. Since a significant spatial autocorrelation of the error term was detected in the residuals of OLS regression model, the spatial regression analysis was also conducted in Tampa. The results of the ordinary linear regression model and the spatial regression model are presented in **Table 6.4**. The coefficients, *z*-values, and significance levels are shown in the table. The ordinary linear regression model and the spatial regression model (SAR_{lag}) show similar results. In both models, the MGVI is significantly associated with the proportion of owner-occupied units, and increases significantly with increases of the proportion of owner-occupied units. However, no significant association was found between the MGVI and racial/ethnic variables (proportion of African Americans, proportion of Asians, and proportion of Hispanics). There exists no significant association between the MGVI and the per capita income. Regression analysis results prove that in Tampa, the MGVI may have less to do with the issue of race/ethnicity and per capita income than it does with the issue of home ownership.

The OLS regression results show that in St. Louis, the MGVI increases significantly with the proportion of Asians, and declines significantly with the proportion of African Americans. However, in the SAR model, there exists no significant association between the MGVI and the proportion of African Americans after controlling the spatial dependence effects. In both the OLS model and the SAR_{lag} model, the MGVI is not significantly associated with the per capita income, proportion of owner-occupied units, and proportion of Hispanics. Regression results show that in St. Louis, the environmental inequity is not serious in terms of the street greenery among different racial/ethnic groups and social classes.

In Kansas City, the SAR regression model was used because of the significant spatial

dependence of the residuals in the OLS regression model (Moran's I = 0.20, z score=7.25). The SAR_{err} was used in Kansas City based on the results of Lagrange Multiplier and Robust Lagrange Multiplier tests. The regression results show that in Kansas City, the MGVI is significantly associated with the proportion of owner-occupied units and the median building age. The MGVI increases significantly with the proportion of owner-occupied units and the median building age. The proportion of Asians, proportion of Hispanics, and proportion of African Americans have no significant association with the MGVI. No significant association between the per capita income and the MGVI was detected. The regression analysis results prove that in Kansas City, the MGVI may have less to do with the issue of race/ethnicity and per capita income than it does with the issue of home ownership.

Table 6.4

The ordinary Least Squares (OLS) regression models and SAR regression models of MGVI and environmental justice variables for different cities.

Cities	Variables	OLS		SAR	
		Coefficient	z-values	Coefficient	z-values
Boston	Constant	9.21	4.84**	4.16	2.26*
	Per capita income	0.01	0.78	0.003	0.24
	Proportion of African Americans	2.04	1.45	1.09	0.91
	Proportion of Asians	9.81	2.66**	5.00	1.59
	Proportion of Hispanics	21.36	0.85	22.88	1.07
	Proportion of owner-occupied units	8.14	4.40**	4.41	2.73**
	Median building age	-0.04	-1.60	-0.03	-1.22
	Rho			0.55	6.74**
	R^2	0.17		0.37	
	Adjusted R^2	0.13			
	F-statistic	4.98**			
	Akaike info criterion			823.94	

	Moran's *I* of residuals	0.23 (5.54**)			
D.C.	Constant	12.29	3.86**	-3.24	-1.35
	Per capita income	0.09	3.37**	0.06	3.33**
	Proportion of African Americans	0.58	0.26	2.50	1.61
	Proportion of Asians	-21.29	-1.63	-7.71	-0.87
	Proportion of Hispanics	-1.74	-0.46	-2.26	-0.86
	Proportion of owner-occupied units	-0.04	-0.02	-1.57	-1.30
	Median building age	0.01	0.42	0.05	2.50
	Rho			0.84	15.42**
	R^2	0.23		0.62	
	Adjusted R^2	0.20			
	F-statistic	7.93**			
	Akaike info criterion	924.47		820.67	
	Moran's *I* of residuals	0.46 (17.37**)			
Philadelphia	Constant	13.00	8.45	5.00	3.74**
	Per capita income	0.09	6.00**	0.05	3.66**
	Proportion of African Americans	-0.934	-1.17	-0.75	-1.20
	Proportion of Asians	-8.21	-2.90**	-5.00	-2.27*
	Proportion of Hispanics	-4.72	-3.37**	-2.00	-1.78
	Proportion of owner-occupied units	-1.10	-0.98	-0.69	-0.79
	Median building age	-0.07	-3.87**	-0.03	-2.01*
	Rho			0.66	13.85**
	R^2	0.28		0.56	
	Adjusted R^2	0.27			
	F-statistic	23.08			
	Akaike info criterion	1974.04		1834.71	
	Moran's *I* of residuals	0.42 (14.39**)			
Baltimore	Constant	11.84	3.01**	-0.66	-0.22
	Per capita income	0.01	0.29	0.04	1.05
	Proportion of African Americans	2.03	0.91	-0.07	-0.04
	Proportion of Asians	33.90	2.54*	21.30	2.15*
	Proportion of Hispanics	-9.97	-1.42	0.36	0.07
	Proportion of owner-occupied units	14.02	6.15**	6.34	3.73**
	Median building age	-0.12	-3.43**	-0.04	-1.40
	Rho			0.94	26.28**
	R^2	0.27			
	Adjusted R^2	0.25		0.58	
	F-statistic	11.15**			

	Akaike info criterion	1161.14		1068.64	
	Moran's *I* of residuals	0.26 (16.68**)			
Denver	Constant	10.72	7.08**	6.48	3.08**
	Per capita income	0.06	3.20**	0.05	2.85**
	Proportion of African Americans	-12.60	-5.24**	-11.30	-4.79**
	Proportion of Asians	0.21	0.02	0.58	0.07
	Proportion of Hispanics	-8.00	-5.70**	-5.93	-3.83**
	Proportion of owner-occupied units	-3.78	-3.25**	-2.26	-1.86
	Median building age	0.12	7.90**	0.09	5.52**
	Rho			0.31	2.88**
	R2	0.72		0.74	
	Adjusted R2	0.71			
	F-statistic	48.85**			
	Akaike info criterion	556.01		552.45	
	Moran's I of residuals	0.064 (2.74**)			
Detroit	Constant	15.16	5.77**	4.30	2.09*
	Per capita income	0.03	0.08	0.06	0.21
	Proportion of African Americans	-1.23	-0.61	-0.94	-0.64
	Proportion of Asians	-10.62	-1.94	-5.83	-1.47
	Proportion of Hispanics	-3.32	-1.18	-1.19	-0.58
	Proportion of owner-occupied units	6.05	4.22**	3.28	3.10**
	Median building age	-0.003	-0.13	0.003	0.14
	Rho			0.68	14.22**
	R^2	0.10		0.51	
	Adjusted R^2	0.08			
	F-statistic	4.76**			
	Akaike info criterion	1509.54		1370.32	
	Moran's *I* of residuals	0.51 (13.53**)			
Pittsburgh	Constant	7.31	3.08**		
	Per capita income	0.12	3.94**		
	Proportion of African Americans	8.41	7.09**		
	Proportion of Asians	12.49	2.36*		
	Proportion of Hispanics	0.11	0.01		
	Proportion of owner-occupied units	12.46	7.75**		
	Median building age	-0.07	-2.53*		
	Rho				
	R^2	0.47			

	Adjusted R^2	0.44			
	F-statistic	16.81**			
	Akaike info criterion	606.48			
	Moran's *I* of residuals	0.04 (1.68)			
Tampa	Constant	16.71	7.23**	1.81	0.64
	Per capita income	-0.02	-0.57	0.01	0.21
	Proportion of African Americans	-0.58	-0.30	-0.42	-0.24
	Proportion of Asians	-14.64	-1.26	-14.15	-1.41
	Proportion of Hispanics	-0.74	-0.31	-0.35	-0.17
	Proportion of owner-occupied units	5.18	2.58*	3.81	2.19*
	Median building age	0.01	0.47	0.02	0.70
	Rho			0.77	6.82**
	R^2	0.15		0.32	
	Adjusted R^2	0.10			
	F-statistic	2.52*			
	Akaike info criterion	436.30		424.61	
	Moran's *I* of residuals	0.10 (4.51**)			
St. Louis	Constant	18.03	5.22**	11.38	2.72**
	Per capita income	0.004	0.08	0.01	-0.34
	Proportion of African Americans	-5.67	-3.13**	-3.60	-1.95
	Proportion of Asians	26.17	2.46*	26.38	2.69**
	Proportion of Hispanics	-15.00	-1.14	-19.16	-1.58
	Proportion of owner-occupied units	0.000	0.02	0.42	0.20
	Median building age	0.74	0.32	-0.01	-0.34
	Rho			0.42	2.94**
	R^2	0.37		0.43	
	Adjusted R^2	0.33			
	F-statistic	9.06			
	Akaike info criterion	529.99		524.63	
	Moran's *I* of residuals	0.10 (3.28**)			
Kansas City	Constant	8.57	3.76**	8.74	3.29**
	Per capita income	-0.03	-0.70	-0.01	-0.26
	Proportion of African Americans	-0.83	-0.42	-0.08	-0.04
	Proportion of Asians	5.51	0.34	4.58	0.32
	Proportion of Hispanics	0.30	0.07	3.72	0.94
	Proportion of owner-occupied units	8.49	4.18**	7.83	3.99**
	Median building age	0.15	6.76**	0.14	4.24**
	LAMBDA			0.69	6.66**

R^2	0.31	0.44
Adjusted R^2	0.28	
F-statistic	10.28**	
Akaike info criterion	857.86	834.17
Moran's I of residuals	0.20 (7.25**)	

**Significant at the 0.01 level (2-tailed).

*Significant at the 0.05 level (2-tailed).

Chapter 7 Discussion

In this study, the GSV based MGVI was used to explore the uneven distributions of urban street greenery in residential areas among different socioeconomic and racial/ethnic groups in ten U.S. major cities. The objective of this study is to test the hypothesis that the environmental inequity in terms of street greenery is linked to the racial/ethnic makeup or socioeconomic status of residents. The GSV based MGVI, which is calculated based on static GSV images captured in different directions, was used to indicate the distribution of the street greenery.

Different cities and different parts of cities have very different residential street greenery levels. Generally, the periphery parts of cities have more street greenery than the inner part of cities. The environmental inequities in terms of street greenery in different cities are different among different racial/ethnic and socio-economic groups. No consistent inequity was detected in different cities in terms of uneven distribution of the residential street greenery.

7.1 The GSV based MGVI

The MGVI method is more suitable for measuring the distribution of street greenery rather

than other types of urban greenery, like backyard greenery and urban parks. The MGVI was calculated based on street-level images at different horizontal and vertical view angles, thus representing the amount of street greenery people can see from the ground. Compared with the canopy cover and vegetation indices, the MGVI may be more suitable for quantifying the amount of street greenery in residential areas (Yang et al., 2009; Leslie et al., 2010). Although the canopy coverage and vegetation indices can be used to map the overall level of greenness in a specific area, it is difficult to quantify the amount of street greenery, which is distributed along linear features. Chapter 5 compared the MGVI with surrounding canopy coverage and averaged canopy height. Results show that the MGVI has a strong and positive correlation with the surrounding canopy coverage and canopy height. Therefore, the MGVI not just represents the amount the street tree canopies but also reflects the vertical structure information of street tree canopies. More street trees, larger canopy cover, and higher canopy height all contribute to large MGVI values.

A map of the MGVI can effectively provide urban planners with detailed information on the spatial distribution of street greenery at the site level. Based on the site level MGVI map, the potential street greening sites can be easily delineated, which seems difficult using the canopy coverage indicators for this purpose. In urban settings, the MGVI is also affected by the layout of buildings and vegetation and the distance between trees and viewers. Therefore, in urban greening projects, planting trees close to pedestrians, choosing tree species with large canopies, or using large-size trees along streets all help to augment the MGVI.

As a free online service provided by Google, Google Street View (GSV) covers cityscapes of

most U.S. cities. The GSV static images can be accessed and downloaded through the Google Street View static image API (Google, 2015) for any place with GSV available. Compared with Yang's (2009) *in situ* image collection and manual green vegetation delineation method for calculating the green view index, the GSV-based MGVI in this study is calculated by downloading and processing online GSV images automatically. Therefore, the GSV-based MGVI method is time-efficient and suitable for street greenery assessment at large-scale level. In addition, the MGVI considers GSV images at three vertical angles and six horizontal angles, which could better represent what people on the ground really see or feel. What more important is that the GSV-based method can be used for street greenery assessment for any place where GSV images are available. This is very important for some areas, where high resolution remotely sensed data are not available. GSV would provide a free data source for mapping the spatial distribution of street greenery. Even for those areas with inventory data or high-resolution remotely sensed data available, the MGVI derived from GSV images would also provide an additional information about the distribution of street greenery, which is different from vegetation indicators derived from a land cover map or remotely sensed imagery. GSV images may be seen as an additional data source for geographers and urban planners in future urban studies and urban greening practices.

7.2 Distributions of street greenery

The MGVI maps show that different cities and different parts of cities have very different distributions of street greenery. Since this study focuses on the residential area, the MGVI maps

represent the residential street greenery level in living neighborhoods of different cities. **Fig. 6.2** presents the histograms of MGVI values at site level for all studied cities. Histograms show that different cities have very different amounts of residential street greenery. Cities of Washington D.C, Tampa, and Kansas City have relative normal histograms, and are greener than other cities in terms of street greenery. However, cities like Philadelphia and Boston have quite skewed histograms, and have less street greenery compared with other cities. In future, Philadelphia and Boston may consider taking measures to increase the street greenery level to develop more environmental friendly neighborhoods. Aoki (1991) suggested that most people would have a favorable impression of a street landscape if more than 30% of the view includes greenery. Based on Aoki's criterion, many streets in the residential areas in these cities, should be given a higher priority in future urban greening projects. Considering the fact that street greenery is managed and maintained by public agencies, municipal governments can make some differences to the amount of street greenery by different policies and initiatives.

Fig. 6.3 shows the spatial distribution of the MGVI at both site level and census tract level for all cities. Generally, the periphery parts of cities have larger MGVI values than inner parts in both of these two types of maps. This is not difficult to understand, because planting trees in densely urban areas is not as easy as the suburban areas. The physical environment in inner parts of city is usually hash for tree growth. In addition, there is no much space in the inner parts of cities for tree planting.

7.3 Environmental inequities in terms of street greenery

The statistical results show that there is no consistent relationship between the MGVI and residents' racial/ethnic and socio-economic statuses across different cities. The per capita income is not consistently correlated with the MGVI (**Table 6.3**). For example, in cities of Baltimore, Washington D.C, Philadelphia, St. Louis, Denver, significant positive correlations are detected between the MGVI and per capita income. However, in cities of Boston, Detroit, Pittsburgh, Tampa, and Kansas City, the correlations between per capita income and MGVI are not significant. The regression results show that there is no consistent relationship between the MGVI and per capita income after controlling other social variables. Generally, residents with higher per capita incomes tend to live in areas with more street greenery compared with those with lower per capita incomes (Jesdale et al., 2013; Schwarz et al., 2014; Pedlowski et al., 2002; Landry and Chakraborty, 2009). This trend could be explained by the fact that people with higher incomes tend to spend more money to choose or improve their living environments with more greenery for a series of benefits. Those areas with less street greenery may be more affordable for low-income people (Pham et al., 2012), and low-income people have less budget to maintain or increase the greenery around their properties. However, the relationships between the street greenery and residents' socio-economic statuses are complex issues, and are affected by many social and environmental factors. Different cities have different urban planning histories, urban patterns, climate environments, soil types, racial/ethnic compositions, etc. All of these factors could affect the people's incentives to greening their neighborhoods. For example, there exists a strong positive correlation between the per capita income and the MGVI in Denver (r = 0.652, sig < 0.01). After controlling the effect of building age, racial/ethnic variables, and lifestyle

variable, the association between the MGVI and per capita income is still very significant (coeff = 0.06, sig < 0.01). Denver is located in the semi-arid region, planting or maintaining street greenery is supposed to be much more difficult that other humid regions. This to some extent could explain the significant and positive coefficients for the per capita income in the regression models of Denver. However, in Tampa, results of both the bivariate correlation analysis and regression models show that there exists no significant correlation and association between the MGVI and per capita income. Different from the semiarid Denver, Tampa is located in a humid tropical climate zone. Its physical environment is much better than that of Denver for tree planting and growth. The different climates could explain the different associations between the per capita income and the MGVI in different cities.

The relationships between the MGVI and the racial/ethnic variables vary across cities. Both bivariate analysis and regression models results show that there are no significant relationships between the MGVI and proportion of Hispanics in Washington D.C, Pittsburgh, Detroit, Tampa, and Kansas City. There is a weakly and significantly positive correlation between the MGVI and proportion of Hispanics in St. Louis. However, regression models show that the weak significance disappears in regression models after controlling other social variables (**Table 6.4**, St. Louis). In Denver and northeastern cities (Boston, Baltimore, Philadelphia, with exception of Washington D.C), the MGVI is significantly and negatively correlated with the proportion of Hispanics (**Table 6.3**). However, after controlling other social variables and spatial dependence, these associations in Boston, Baltimore, and Philadelphia become insignificant (**Table 6.4**). These statistical analyses show that there is no significantly less street greenery in neighborhoods

with Hispanics in these cities. In Denver, the association between the MGVI and the proportion of Hispanics is still very significant and negative in OLS and SAR_{lag} regression models. The negative coefficient (OLS model: coeff = -8.0, sig<0.01; SAR_{lag} model: coeff = -5.93, sig<0.01) of proportion of Hispanics shows that the MGVI decreases significantly with increases of the proportion of Hispanics even after controlling the effect of spatial dependence and other social variables. This clearly shows that Hispanics tend to live in neighborhoods with less street greenery in Denver.

Bivariate analysis results show that the correlations between the MGVI and the proportion of African Americans vary in different cities. Regression results show that, generally, after controlling other social variables and spatial dependence, the significances disappear in all cities except Denver, Colorado. The relationships between the MGVI and the proportion of Asians are mixed across different cities. Bivariate correlation analyses show that the MGVI is significantly and positively correlated only in cities of Baltimore and St. Louis. The correlations are not significant in the rest cities. Regression results show that, after controlling the effect of other social variables and spatial dependence, the MGVI is negatively associated with proportion of Asians in Philadelphia, but positively associated with proportion of Asians in St. Louis, Pittsburgh, and Baltimore. In Philadelphia, the negative association between the proportion of Asians and the MGVI shows that Asians live in neighborhoods with less street greenery in Philadelphia. Considering the lowest street greenery level in Philadelphia compared with other cities, this situation is supposed to be even worse. Future urban greening projects and practices should put more efforts on neighborhoods with more Asians to decrease the environmental

inequities in terms of street greenery in Philadelphia.

In accordance with previous studies (Li et al, 2015), bivariate correlation analysis shows that in almost all cities in this study, the MGVI is significantly and positively correlated with the lifestyle variable – proportion of owner-occupied units. Regression models further proved the relatively consistent positive associations between the MGVI and proportion of owner-occupied units in those chosen cities in this study. It seems that the homeowners are more willing to live in neighborhoods with more street greenery. The homeowners can also help to maintain the street greenery, which may also contribute to more street greenery.

7.4 Limitations and future studies

While this study demonstrates that GSV is feasible for assessing street greenery and may deliver useful street greenery information that was unavailable previously, there are still some issues that need to be resolved in future studies. The first issue concerns the time consistency of GSV images. Google provides the acquisition date of street view images, which provides the information for researchers and practitioners to better match environmental conditions with their data analysis and study outcome. Since this study analyzed the street greenery over a period of five years, the neglect of accurate image dates may not have a great effect on the analysis. However, for some studies focusing at a specific time point, the time consistency could matter. Therefore, how to keep the time consistency is an important issue for future GSV applications in assessing urban green spaces.

Secondly, too many GSV images captured in non-green seasons may affect the applications of GSV for urban greenery studies. Originally, this study selected 26 major U.S. cities; however, more than half of these cities were removed from the further analyses because of the clustered non-green sample sites in those cities. Actually, Google is still updating the Google Street View images, and there could have GSV images captured in different seasons or different years for one site. Although it is possible to collect the time information of GSV images, it is still impossible or difficult to use Google Street View to investigate the multi-temporal changes of street greenery, because the current method is still not able to access the GSV images for any specific time point. In this study, many major cities were not included in the analyses because there are too many sites have GSV images captured in non-green seasons in these cities. It could be possible to access GSV images at different time points in future, since Google has already collected historical GSV data and presented it online. In future, the new version of the Google Street View Image API may let users to access to GSV images for any available time point. Therefore, future studies should also focus on the temporal changes of the MGVI in different cities.

Thirdly, different types of urban greenery show very different distributions across different neighborhoods. Different types of urban greenery are managed and maintained in different ways, therefore, different measures should be taken to reduce the environmental inequities in terms of different types of urban greenery. The proposed GSV based MGVI measures the amount of street greenery from the perspective of the visibility of urban greenery. The visibility of urban greenery may have direct connections with the sensory functions of urban greenery, since the

green view would influence human perception of the environment and human activities. However, the visibility of urban greenery may have no direct connection with other physical benefits provided by the urban greenery, such as mitigation of urban heat island, air pollution, and urban floods. Therefore, different methods should be used to thoroughly evaluate the benefits of urban greenery from different perspectives. In addition, the GSV images were all collected along streets, and not all cityscapes are covered. Therefore, the proposed GSV based MGVI is more suitable for the assessment of street greenery but not suitable for other types of urban greenery, although the MGVI may also cover some backyard greenery.

This study used the fixed census tract as the geographic unit for the statistical analysis. There would have the modifiable areal unit problem (MAUP). This is because the MGVI was calculated at site level along streets. However, aggregating site level maps to census tract level will inevitably cause some loss of information or bias. In this study, those GSV sites on the border of two neighboring census tracts were removed to solve the MAUP problem to some extent. Future studies should investigate the effect of the MAUP on the reliability of the environmental inequity analysis results thoroughly.

The relationship between the street greenery and residents' socio-economic status is a complex issue. Climate, demographics, policy, density of city, and city size could all be the potential factors for influencing the relationships between MGVI and social variables (Schwarz et al., 2015). In this study, only the social variables from census data were considered on investigating the environmental inequities among different racial/ethnic and socio-economic groups. In future studies, more factors should be considered, such as the different residential land

use types (downtown, multifamily residential land, single-family residential land), residents' cultural backgrounds, climate zones, etc. Trees need many years to grow, and tree-planting projects cannot get immediate effects on the local greenness. The mismatch between the physical environment and the social status could also make the relationships between street greenery and social variables vary. In addition, the ACS data has large margin-of-errors, which will also increase the uncertainty in investigating the relations between street greenery and social statuses of local residents.

In addition, urban greenery is not always an environmental amenity. It may increase the budget for cleaning the dead leaves and branches. The root of the street greenery could break the road conditions along the streets, especially the walkways. It could enhance the damage of storms. The dead branches could cut the wire, which further bring inconvenience to local residents. In rural areas, there have enough space for tree planning and there have no as much impervious surface as the densely populated cities, the importance of street greenery of other kinds of green spaces may not be highlighted. Planting trees in the dense urban areas is not easy, and it could be a big financial burden to the municipal governments. In addition, people from different cultural backgrounds may have very different opinions about the urban greenery. Future study should also investigate public opinions on different types of urban greenery.

Chapter 8 Conclusions and Contributions

The study used millions of static Google Street View (GSV) images to investigate the environmental inequities in terms of street greenery in ten U. S. major cities. As a novel data driven geographical study based on GSV, this study provides an example to geographers to utilize geo-tagged GSV images for geographical analyses. Different from the traditional remotely sensed data, which usually represent the overhead view of geographic features, the GSV images capture the profile view of cityscapes. The profile view of cityscapes is closer to what people can see or feel on the ground than the overhead view. Therefore, the GSV based MGVI may be more suitable to represent the greenness of neighborhoods in terms of what people see or feel on the ground. In addition, the Google Street View cars have taken pictures of streetscapes all over the world. Google keeps updating the GSV images on Google Maps periodically. It is a great data source for urban studies in future.

Different cities and different parts of cities have very different residential street greenery levels. Generally, the periphery parts of cities have more street greenery than the inner part of cities. The cities of Washington D.C, Tampa, and Kansas City generally have more residential street greenery than other cities. Philadelphia and Boston have the lowest residential street

greenery levels compared with other cities in this study.

The environmental inequities in terms of street greenery in different cities are different among different racial/ethnic and socio-economic groups. No consistent inequity was detected in different cities in terms of uneven distribution of residential street greenery. In cities of Washington D.C, Philadelphia, Pittsburgh, and Denver, people with lower income levels tend to live in areas with less greenery while those with higher incomes live in greener areas. In cities of Boston, Baltimore, Pittsburgh, Detroit, Tampa, and Kansas City, the proportion of owner-occupied units is positive associated with street greenery levels. The proportion of Hispanics and the proportion of African Americans are not significantly associated with the amount of street greenery in almost all cities except in Denver. In Denver, the environmental inequity problem among different racial/ethnic groups is significant. The Hispanics and African Americans tend to live in neighborhoods with less street greenery.

The street greenery provides many benefits to urban residents meeting various and overlapping goals. The existence of the street greenery is an important factor of life quality in densely urban areas. The street greenery is a kind of publicly financed facilities, and it relies heavily on the public investment. In other way, the public agencies or municipal governments can do something to modify the distribution of street greenery to reduce the environmental inequities. In the future, more attention needs to be paid on increasing the residential street greenery in those critical areas where environmental inequities exist. This could help to balance the living greenness among different socioeconomic groups in different cities.

Reference

Akbari, H., Pomerantz, M. & Taha, H. (2001). Cool surfaces and shade trees to reduce energy use and improve air quality in urban areas. *Solar energy* 70, 295–310.

Almeer. M. H. (2012). Vegetation extraction from Free Google Earth Images of deserts using a robust BPNN approach in HSV space. *International Journal of Emerging Technology and Advanced Engineering*, 2(5): 1-8.

Anselin, L. (2005). Exploring spatial data with GeoDa^{TM}: A workbook, Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics. University of Illinois., p. 244.

Anselin L, Bera A K. (1998). Spatial dependence in linear regression models with and introduction to spatial econometrics", in Handbook of Applied Economic Statistics Eds AUllah, DE AGiles (Marcel Dekker, New York) pp: 237 – 289.

Aoki, Y. (1991). Evaluation methods for landscapes with greenery. *Landscape Research*, 16(3), 3-6.

Asgarzadeh, M., Lusk, A., Koga, T., & Hirate, K. (2012). Measuring oppressiveness of streetscapes. *Landscape and Urban Planning*, 107(1), 1-11.

Bain, L., Gray, B., Rodgers, D. (2012). Living Streets: Strategies for Crafting Public Space. John Wiley & Sons.

Balram, S., & Dragićević, S. (2005). Attitudes toward urban green spaces: integrating questionnaire survey and collaborative GIS techniques to improve attitude measurements. *Landscape and Urban Planning*, *71*(2), 147-162.

Been, V. (1994). Locally undesirable land uses in minority neighborhoods: Disproportionate siting or market dynamics? *Yale law journal*, 1383-1422.

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

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS journal of photogrammetry and remote sensing*, *65*(1), 2-16.

Bolund, P., Hunhammar, S. (1999). Ecosystem services in urban areas, *Ecological Economics*, 29(2): 293-301.

Boone, C. G., Buckley, G. L., Grove, J. M., & Sister, C. (2009). Parks and people: An environmental justice inquiry in Baltimore, Maryland. *Annals of the Association of American Geographers*, 99(4), 767-787.

Brainard J S, Jones A P, Bateman I J, Lovett A A, Fallon P J. (2002). Modelling environmental equity: access to air quality in Birmingham, England, *Environment and Planning A* 34 695 – 716.

Camacho-Cervantes, M., Schondube, J. E., Castillo, A., & MacGregor-Fors, I. (2014). How do people perceive urban trees? Assessing likes and dislikes in relation to the trees of a city. *Urban ecosystems*, *17*(3), 761-773.

Charreire, H., Mackenbach, J.D., Ouasti, M., Lakerveld, J., Compernolle, S., Ben-Rebah, M., Oppert, J.M. (2014). Using remote sensing to define environmental characteristics related to physical activity and dietary behaviours: a systematic review (the SPOTLIGHT project). *Health Place*, 25, 1–9.

Chen, X. L., Zhao, H. M., Li, P. X., & Yin, Z. Y. (2006). Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. *Remote sensing of environment*, *104*(2), 133-146.

Cheng, Y. (1995). Mean shift, mode seeking, and clustering. *IEEE transactions on pattern analysis and machine intelligence*, *17*(8), 790-799.

Chen, Z., Xu, B., & Gao, B. (2015). Assessing visual green effects of individual urban trees using airborne Lidar data. *Science of the Total Environment*, *536*, 232-244.

Clark, L. P., Millet, D. B., & Marshall, J. D. (2014). National patterns in environmental injustice and inequality: outdoor NO 2 air pollution in the United States. *PloS one*, *9*(4), e94431.

Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5), 603-619.

Cooper, K. B. (2005). Halfway to the 2010 Census: The Countdown and Components to a Successful Decennial Census. presentation to the House Subcommittee on Federalism and the Census, US House of Representatives, Washington, DC.

Coutts, C. (2008). Greenway accessibility and physical-activity behavior. *Environment and Planning B: Planning and Design*, 35(3), 552-563.

Dai, D. (2011). Racial/ethnic and socioeconomic disparities in urban green space accessibility: Where to intervene? *Landscape and Urban Planning*, 102(4), 234-244.

Daniel, T.C., Boster, R.S. (1976). Measuring Landscape Esthetics: The Scenic Beauty Estimation Method. USDA Forest Service Research Paper, RM-167. Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO.

de Smith. (2015). Statistical Analysis handbook, A comprehensive handbook of statistical concepts, techniques and software tools, http://www.statsref.com/HTML/index.html?sar_models. html.

Donovan, G.H.; Prestemon, J.P. (2012). The effect of trees on crime in Portland, OR. *Environment and Behaviors*, 44, 3–30.

Dwivedi, P., Rathore, C.S. and Dubey, Y. (2009). Ecological benefits of urban forestry: the case of Kerwa Forest Area (KFA), Bhopal, India. *Applied Geography*, 29, 194–200.

Edwards, N., Hooper, P., Trapp, G. S., Bull, F., Boruff, B., & Giles-Corti, B. (2013). Development of a public open space desktop auditing tool (POSDAT): a remote sensing approach. *Applied Geography*, 38, 22-30.

Ellaway, A., S. Macintyre, and X. Bonnefoy. (2005). Graffiti, Greenery, and Obesity in Adults: Secondary Analysis of European Cross Sectional Survey. British Medical Journal (Clinical Research Edition) 331 (7517): 611–612. doi:10.1136/bmj.38575.664549.F7.

Faryadi, S., & Taheri, S. (2009). Interconnections of urban green spaces and environmental quality of Tehran.

Fisher, B. S., & Nasar, J. L. (1992). Fear of crime in relation to three exterior site features prospect, refuge, and escape. *Environment and Behavior*, 24(1), 35-65.

Fukunaga, K., & Hostetler, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on information theory*, *21*(1), 32-40.

Gidlow, C.J., Ellis, N.J., & Bostock, S. (2012). Development of the Neighbourhood Green Space Tool (NGST). *Landscape and Urban Planning* 106: 347-358.

Givoni, B. (1991). Impact of planted areas on urban environmental quality: a review. *Atmospheric Environment. Part B. Urban Atmosphere*, 25(3), 289-299.

Google. (2014). Google Street View Image API, https://developers.google.com/maps/documentation/streetview/, accessed October. 2014.

Google. (2015). Google Maps API, https://developers.google.com/maps/, accessed 2015.

Griew, P., Hillsdon, M., Foster, C., Coombes, E., Jones, A., & Wilkinson, P. (2013). Developing and testing a street audit tool using Google Street View to measure environmental supportiveness for physical activity. *International Journal of Behavioral Nutrition and Physical Activity*, 10(1), 1.

Grimm, N.B., Faeth, S.H., Golubiewski, N.E., Redman, C.L., Wu, J., Bai, X., & Briggs, J.M. (2008). Global change and the ecology of cities. *Science*, *319*: 756-760.

Grove, J.M., Troy, A.R., O'Neil-Dunne, J.P.M., Burch Jr., W.R., Cadenasso, M.L., Pickett, S.T.A. (2006). Characterization of households and its implications for the vegetation of urban ecosystems. *Ecosystems* 9 (4), 578–597.

Guijarro, M., Pajares, G., Riomoros, I., Herrera, P. J., Burgos-Artizzu, X. P., & Ribeiro, A. (2011). Automatic segmentation of relevant textures in agricultural images. *Computers and Electronics in Agriculture*, 75(1), 75-83.

Gupta, K., Kumar, P., Pathan, S.K., & Sharma, K.P. (2012). Urban Neighborhood Green Index- A measure of green spaces in urban areas. *Landscape and Urban Planning, 105*, 325-335.

Hewko J, Smoyer-Tomic K E, Hodgson M J. (2002), Measuring neighbourhood spatial accessibility to urban amenities: does aggregation error matter? *Environment and Planning A*, 34, 1185 – 1206.

Heynen, N., Perkins, H. A., & Roy, P. (2006). The political ecology of uneven urban green space the impact of political economy on race and ethnicity in producing environmental inequality in Milwaukee. *Urban Affairs Review*, 42(1), 3-25.

Hough, M. (1984). City form and natural processes. London: Croom Helm.

Huang, G., Zhou, W., & Cadenasso, M. L. (2011). Is everyone hot in the city? Spatial pattern of land surface temperatures, land cover and neighborhood socioeconomic characteristics

in Baltimore, MD. *Journal of environmental management*, 92(7), 1753-1759.

Jesdale, B.M., Morello-Frosch, R., Cushing, L. (2013). The racial/ethnic distribution of heat risk-related land cover in relation to residential segregation. *Environmental Health Perspectives,* 121 (7), 811.

Jenerette, G. D., Miller, G., Buyantuev, A., Pataki, D. E., Gillespie, T. W., & Pincetl, S. (2013). Urban vegetation and income segregation in drylands: a synthesis of seven metropolitan regions in the southwestern United States. *Environmental Research Letters*, 8(4), 044001.

Jennings, V., Johnson Gaither, C., & Gragg, R. S. (2012). Promoting environmental justice through urban green space access: A synopsis. *Environmental Justice*, 5(1), 1-7.

Jensen, R., Gatrell, J., Boulton, J., & Harper, B. (2004). Using remote sensing and geographic information systems to study urban quality of life and urban forest amenities. *Ecology and Society*, 9(5), 5.

Jim, C. Y., & Chen, W. Y. (2006). Impacts of urban environmental elements on residential housing prices in Guangzhou (China). *Landscape and Urban Planning*, *78*(4), 422-434.

Jim, C.Y., & Chen, W. (2008). Assessing the ecosystem service of air pollutant removal by urban trees in Guangzhou (China). *Journal of Environment Management* 88: 665-676.

Kardan, O., Gozdyra, P., Misic, B., Moola, F., Palmer, L. J., Paus, T., & Berman, M. G. (2015). Neighborhood greenspace and health in a large urban center. *Scientific reports*, 5.

Kaplan, R. (2001). The nature of the view from home: psychological benefits. *Environment and Behaviour*, 33, 507–542.

Kong, F., Yin, H., & Nakagoshi, N. (2007). Using GIS and landscape metrics in the hedonic price modeling of the amenity value of urban green space: A case study in Jinan City, China. *Landscape and Urban Planning*, *79*(3), 240-252.

Konijnendijk, C. C., Ricard, R. M., Kenney, A., & Randrup, T. B. (2006). Defining urban forestry–A comparative perspective of North America and Europe. *Urban Forestry & Urban Greening*, *4*(3), 93-103.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the

Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15(3), 259-263.

Kuo, F. E., & Sullivan, W. C. (2001). Environment and crime in the inner city: Does vegetation reduce crime? *Environment and Behavior,* 33, 343–367.

Lachowycz, K., Jones, A. (2013). Towards a better understanding of the relationship between greenspace and health: development of a theoretical framework. *Landscape Urban Planning*, 118, 62–69.

Landry, S. M., & Chakraborty, J. (2009). Street trees and equity: evaluating the spatial distribution of an urban amenity. *Environment and Planning A*, *41*(11), 2651-2670.

Lawrence, H.W. (1995). Changing forms and persistent values: historical perspectives on the urban forest. In: Bradley, G.A. (Ed.), Urban Forest Landscapes: Integrating Multidisciplinary Perspectives. University of Washington Press, Seattle, pp. 17–40.

Lee, A. C. K., & Maheswaran, R. (2011). The health benefits of urban green spaces: a review of the evidence. *Journal of Public Health*, 33(2), 212-222.

Leslie, E., Sugiyama, T., Ierodiaconou, D., Kremer, P. (2010). Perceived and objectively measured greenness of neighbourhoods: are they measuring the same thing? *Landscape and Urban Planning,* 95 (1), 28–33.

Li, X., Meng, Q., Li, W., Zhang, C., Jancso, T., & Mavromatis, S. (2014). An explorative study on the proximity of buildings to green spaces in urban areas using remotely sensed imagery. *Annals of GIS*, 20(3), 193-203.

Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., & Zhang, W. (2015a). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening*, 14(3), 675-685.

Li, X., Zhang, C., Li, W., Kuzovkina, Y. A., & Weiner, D. (2015b). Who lives in greener neighborhoods? The distribution of street greenery and its association with residents' socioeconomic conditions in Hartford, Connecticut, USA. *Urban Forestry & Urban Greening*, 14(4), 751-759.

Li, X., Li, W., Meng, Q., Zhang, C., Jancso, T., & Wu, K. (2016a). Modelling building proximity to greenery in a three-dimensional perspective using multi-source remotely

sensed data. *Journal of Spatial Science*, 1-15.

Li, X., Zhang, C., Li, W. and Kuzovkina, Y.A. (2016b). Environmental inequities in terms of different types of urban greenery in Hartford, Connecticut. *Urban Forestry & Urban Greening*.

Ling, P. P., & Ruzhitsky, V. N. (1996). Machine vision techniques for measuring the canopy of tomato seedling. *Journal of Agricultural Engineering Research*, *65*(2), 85-95.

Liu F. (1997). Dynamics and causation of environmental equity, locally unwanted land uses, and neighborhood changes, *Environmental Management*, 21, 643 – 656.

Liu, W., Chen, W., & Peng, C. (2014). Assessing the effectiveness of green infrastructures on urban flooding reduction: A community scale study. *Ecological Modeling*, 291, 6-14.

Lotfi, S., & Koohsari, M. J. (2011). Proximity to neighborhood public open space across different socio-economic status areas in metropolitan Tehran. *Environmental Justice*, 4(3), 179-184.

Mansfield, C., Pattanayak, S. K., McDow, W., McDonald, R., & Halpin, P. (2005). Shades of green: measuring the value of urban forests in the housing market. *Journal of forest economics*, *11*(3), 177-199.

Maas, J., R. A. Verheij, P. P. Groenewegen, S. de Vries, and P. Spreeuwenberg. (2006). Green Space, Urbanity, and Health: How Strong Is the Relation? *Journal of Epidemiology and Community Health*, 60: 587–592. doi:10.1136/ jech.2005.043125.

Meitner, M. J. (2004). Scenic beauty of river views in the Grand Canyon: relating perceptual judgments to locations. *Landscape and urban planning*, 68(1), 3-13.

Mennis, J. (2006). Socioeconomic-vegetation relationships in urban, residential land. *Photogrammetric Engineering & Remote Sensing*, 72(8), 911-921.

Meyers L S, Gamst G, Guarino A J. (2006). Applied Multivariate Research: Design and Interpretation (Sage, Thousand Oaks, CA).

Mičušík, Branislav, and Jana Košecká. "Piecewise planar city 3D modeling from street view panoramic sequences." In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pp. 2906-2912.

Miller, R.W. Urban forestry: Planning and managing urban green spaces. (1997). 2nd ed. New

Jersey: Prentice Hall. Inc. Upper Saddle River.

Mohai P, Bryant B. (1992). Environmental racism: reviewing the evidence", in Race and the Incidence of Environmental Hazards: A Time for Discourse Eds P Mohai, B Bryant (Westview Press, Boulder, CO) pp. 163 – 175.

Naik, N., S. D. Kominers, R. Raskar, E. L. Glaeser, and C. A. Hidalgo. (2015). Do people shape cities, or do cities shape people? The co-evolution of physical, social, and economic change in five major US cities. NBER Working Paper No. 21620.

Nasar, J., Fisher, B., Grannis, M. (1993). Proximate physical cues to fear of crime. *Landscape and Urban Planning*, 26, 161–178.

Neto, J.C. (2004). A Combined Statistical—Soft Computing Approach for Classification and Mapping Weed Species in Minimum Tillage Systems. University of Nebraska, Lincoln, NE.

Nichol, J., & Wong, M.S. (2005). Modeling urban environmental quality in a tropical city. *Landscape and Urban Planning*, 73: 49-58.

Onishi, A., Cao, X., Ito, T., Shi, F., & Imura, H. (2010). Evaluating the potential for urban heat-island mitigation by greening parking lots. *Urban Forestry & Urban Greening*, 9: 323-332.

Odgers, C.L., Caspi, A., Bates, C.J., Sampson, R.J., Moffitt, T.E. (2012). Systematic social observation of children's neighborhoods using Google Street View: a reliable and cost-effective method. J. Child Psychol. Psychiatry, 53 (10), 1009–1017.

Otsu, N. (1979). A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9: 62-66.

Pastor M, Morello-Frosch R, Sadd J L. (2005). The air is always cleaner on the other side: race, space, and ambient air toxics exposures in California. *Journal of Urban Affairs*, 27, 127-148.

Pazhouhanfar, M., Kamal, M. (2014). Effect of predictors of visual preference as characteristics of urban natural landscapes in increasing perceived restorative potential. *Urban Forestry and Urban Greening* 13 (1), 145–151.

Pearsall, H., & Christman, Z. (2012). Tree-lined lanes or vacant lots? Evaluating

non-stationarity between urban greenness and socio-economic conditions in Philadelphia, Pennsylvania, USA at multiple scales. *Applied Geography*, *35*(1), 257-264.

Pedlowski, M. A., Da Silva, V. A. C., Adell, J. J. C., & Heynen, N. C. (2002). Urban forest and environmental inequality in Campos dos Goytacazes, Rio de Janeiro, Brazil. *Urban Ecosystems*, *6*(1-2), 9-20.

Pham, T.-T.-H., Apparicio, P., Séguin, A.-M., Gagnon, M. (2011). Mapping the greenscape and environmental equity in Montreal: an application of remote sensing and GIS. In: Caquard, S., Vaughan, L., Cartwright, W. (Eds.), Springer Lecture Notes in Geoinformation and Cartography Mapping Environmental Issues in the City. Arts and Cartography Cross Perspectives. Springer, pp. 30–48.

Apparicio, P., Landry, S., Séguin, A. M., & Gagnon, M. (2013). Predictors of the distribution of street and backyard vegetation in Montreal, Canada. *Urban forestry & urban greening*, 12(1), 18-27.

Apparicio, P., Séguin, A. M., Landry, S., & Gagnon, M. (2012). Spatial distribution of vegetation in Montreal: an uneven distribution or environmental inequity? *Landscape and urban planning*, 107(3), 214-224.

Pickett, S.T., Cadenasso, M.L., Grove, J.M., Boone, C.G., Groffman, P.M., Irwin, E., Kaushal, S.S., Marshall, V., McGrath, B.P., Nilon, C.H., Pouyat, R.V., Szlavecz, K., Troy, A., & Warren, P. (2011). Urban ecological systems: scientific foundations and a decade of progress. *Journal of Environmental Management, 92*: 331-362.

Raciti, S. M., Hutyra, L. R., & Newell, J. D. (2014). Mapping carbon storage in urban trees with multi-source remote sensing data: Relationships between biomass, land use, and demographics in Boston neighborhoods. *Science of The Total Environment*, 500, 72-83.

Ribeiro, A., Fernández-Quintanilla, C., Barroso, J., García-Alegre, M.C. (2005). Development of an image analysis system for estimation of weed. In: Proceedings of the 5th European Conference On Precision Agriculture (5ECPA), pp. 169–174.

Ruangrit, V., & Sokhi, B. S. (2004). Remote sensing and GIS for urban green space analysis–A case study of Jaipur city, Rajasthan, India. *Journal of Institute of Town Planners India*, *1*(2), 55-67.

Rundle, A. G., Bader, M. D., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using Google Street View to audit neighborhood environments. *American journal of preventive medicine*, 40(1), 94-100.

Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The collaborative image of the city: mapping the inequality of urban perception. *PloS one*, 8(7), e68400.

Schroeder, H. W., & Cannon, W. N. (1983). The esthetic contribution of trees to residential streets in Ohio towns. *Journal of Arboriculture*, 9(9), 237-243.

Schwarz, K., Fragkias, M., Boone, C. G., Zhou, W., McHale, M., Grove, J. M., ... & Ogden, L. (2015). Trees grow on money: urban tree canopy cover and environmental justice. *PloS one*, 10(4), e0122051.

Seymour, M., Wolch, J., Reynolds, K. D., & Bradbury, H. (2010). Resident perceptions of urban alleys and alley greening. *Applied Geography*, 30(3), 380-393.

Shrestha, D.S., Steward, B.L., Birrell, S.J. (2004). Video processing for early stage maize plant detection. *Biosystems in Engineering*, 89 (2), 119–129.

Shuttleworth, S. (1980). The use of photographs as an environment presentation medium in landscape studies. *Journal of Environmental Management*, 11(1), 61-76.

Spielman, S. E., & Singleton, A. (2015). Studying Neighborhoods Using Uncertain Data from the American Community Survey: A Contextual Approach. *Annals of the Association of American Geographers*, 105(5), 1003-1025.

Stamps, Arthur E. (1990). Use of photographs to simulate environments: A meta-analysis. *Perceptual and Motor Skills* 71(3): 907-913.

Takano, T., Nakamura, K., & Watanabe, M. (2002). Urban residential environments and senior citizens' longevity in megacity areas: the importance of walkable green spaces. *Journal of epidemiology and community health*, 56(12), 913-918.

Talen, E., & Anselin, L. (1998). Assessing spatial equity: an evaluation of measures of accessibility to public playgrounds. *Environment and planning A*, 30(4), 595-613.

Taylor, B. T., Fernando, P., Bauman, A. E., Williamson, A., Craig, J. C., & Redman, S. (2011). Measuring the quality of public open space using Google Earth. *American journal of preventive medicine*, 40(2), 105-112.

Tellaeche, A., Burgos-Artizzu, X. P., Pajares, G., & Ribeiro, A. (2008). A vision-based method for weeds identification through the Bayesian decision theory. *Pattern Recognition*, 41(2), 521-530.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46(sup1), 234-240.

Torii, A., Havlena, M., & Pajdla, T. (2009). From google street view to 3d city models. In ICCV Workshops pp. 2188-2195.

Troy, A., Grove, J. M., & O'Neil-Dunne, J. (2012). The relationship between tree canopy and crime rates across an urban–rural gradient in the greater Baltimore region. *Landscape and Urban Planning*, 106(3), 262-270.

Tsai, V. J., & Chang, C. T. (2013). Three-dimensional positioning from Google street view panoramas. IET Image Processing, 7(3), 229-239.

Ulrich, R. (1984). View through a window may influence recovery. *Science* 224(4647), 224–225.

Van Dillen, S. M., de Vries, S., Groenewegen, P. P., & Spreeuwenberg, P. (2012). Greenspace in urban neighbourhoods and residents' health: adding quality to quantity. *Journal of Epidemiology and Community health*, 66(6), e8-e8.

Walker, H.K., Hall, W.D., Hurst, J.W. (1990). Claudication—Clinical Methods: The History, Physical, and Laboratory Examinations. Butterworths, Boston

Wells, J. E., Buckley, G. L., & Boone, C. G. (2008). Separate but equal? Desegregating Baltimore's golf courses. *Geographical Review*, 98(2), 151-170.

Wendel, H.E., Downs, J.A., & Mihelcic, J.R. (2011). Assessing equitable access to urban green space: the role of engineered water infrastructure. *Environment Science Technology*, 45: 6728-6734.

Woebbecke, D.M. et al. (1992). Lant species identification, size, and enumeration using machine vision techniques on near-binary images. SPIE Opt. Agric., Forestry, Biological Process. 1836, 208–219.

Wolf, K. L. (2005). Business district streetscapes, trees, and consumer response. *Journal of Forestry*, 103(8), 396-400.

Wolfe, M. K., & Mennis, J. (2012). Does vegetation encourage or suppress urban crime? Evidence from Philadelphia, PA. *Landscape and Urban Planning*, 108(2), 112-122.

Wolch, J., Wilson, J. P., & Fehrenbach, J. (2005). Parks and park funding in Los Angeles: An equity-mapping analysis. *Urban Geography*, 26(1), 4-35.

Yang, J., Zhao, L., Mcbride, J., & Gong, P. (2009). Can you see green? Assessing the visibility of urban forests in cities. *Landscape and Urban Planning*, 91(2), 97-104.

Yang, W., L. Mu, and Y. Shen. (2015). Effect of climate and seasonality on depressed mood among Twitter users. *Applied Geography*, 63, 184-191.

Yao, Y., Zhu, X., Xu, Y., Yang, H., Wu, X., Li, Y., & Zhang, Y. (2012). Assessing the visual quality of green landscaping in rural residential areas: the case of Changzhou, China. *Environmental monitoring and assessment*, 184(2), 951-967.

Zamir, A.R., Darino,A., Patrick, R., Shah, M. (2011). Street View challenge: Identification of commercial entities in Street View imagery. Tenth International Conference on Machine Learning and Applications, Honolulu, Hawaii, Dec.2011, pp. 380–383.

Zhang, B., Xie, G., Zhang, C., & Zhang, J. (2012). The economic benefits of rainwater-runoff reduction by urban green spaces: A case study in Beijing, China. *Journal of environmental management*, 100, 65-71.

Zheng, L., Zhang, J., & Wang, Q. (2009). Mean-shift-based color segmentation of images containing green vegetation. *Computers and Electronics in Agriculture*, 65(1), 93-98.

Zheng, L., Shi, D., & Zhang, J. (2010). Segmentation of green vegetation of crop canopy images based on mean shift and fisher linear discriminant. *Pattern Recognition Letters*, 31(9), 920-925.

Zhu, G., Bian, F., Zhang, M., 2003. A flexible method for urban vegetation cover measurement based on remote sensing images. In: ISPRS WG I/5 Workshop, High Resolution Mapping from Space.

Zhou, X., & Kim, J. (2013). Social disparities in tree canopy and park accessibility: A case study of six cities in Illinois using GIS and remote sensing. *Urban forestry & urban greening*, 12(1), 88-97.

Zhou, W., & Troy, A. (2008). An object-oriented approach for analysing and characterizing

urban landscape at the parcel level. *International Journal of Remote Sensing*, 29(11), 3119-3135.