The distribution of greenspace quantity and quality and their association with neighbourhood socioeconomic conditions in Guangzhou, China: A new approach using deep learning method and street view images

Citation for published version:
Wang, R, Feng, Z, Pearce, J, Yao, Y, Li, X & Liu, Y 2021, 'The distribution of greenspace quantity and quality and their association with neighbourhood socioeconomic conditions in Guangzhou, China: A new approach using deep learning method and street view images', *Sustainable Cities and Society*, vol. 66, 102664. https://doi.org/10.1016/j.scs.2020.102664

Digital Object Identifier (DOI):
10.1016/j.scs.2020.102664

Sustainable Cities and Society

The distribution of greenspace quantity and quality and their association with neighbourhood socioeconomic conditions in Guangzhou, China: A new approach using deep learning method and street view images

--Manuscript Draft--

Manuscript Number:	SCSI-D-20-03564R1
Article Type:	Full Length Article
Keywords:	Greenspace; Socioeconomic conditions; Street View; Machine Learning; Environmental disparity; China
Corresponding Author:	Ruoyu Wang UNITED KINGDOM
First Author:	Ruoyu Wang
Order of Authors:	Ruoyu Wang
	Zhiqiang Feng
	Jamie Pearce
	Yao Yao
	Xiaojiang Li
	Ye Liu
Abstract:	Awareness is mounting that urban greenspace is beneficial for residents' health. While a plethora of studies have focused on greenspace quantity, scant attention has been paid to greenspace quality. Existing methods for assessing greenspace quality is either highly labor-intensive and/or prohibitively time-consuming. This study develops a new machine learning method to assess greenspace quality based on street view images collected from Guangzhou, China. It also examines whether greenspace exposure disparities are linked to the neighbourhood socioeconomic status (SES). The validation process indicated that our scoring system achieved high accuracy for predicting street view-based greenspace quality outside the training data. Results also show that there were marked differences in spatial distribution between aggregated NDVI (Normalized Difference Vegetation Index), street view greenness quantity and quality. Regression models show that neighbourhood SES is not associated with NDVI. Although neighbourhood SES is associated with both street view greenness quantity and quality index value, street view greenness quality is more sensitive to the change of neighbourhood SES. Our work suggests that policymakers and planners are advised to pay more attention to greenspace quality and greenspace exposure disparities in urban area.
Suggested Reviewers:	Linchuan Yang Southwest Jiaotong University yanglc0125@swjtu.edu.cn An expert in this topic.
	Yi Lu City University of Hong Kong yilu24@cityu.edu.hk An expert in this topic.
	Wenjie Wu Jinan University wenjiewu@jnu.edu.cn An expert in this topic.
Response to Reviewers:	

December 11, 2020

Professor F. Haghighat
Editor-in-Chief
Sustainable Cities and Society

Dear Editor,

We would like to re-submit our manuscript entitled "The distribution of greenspace quantity and quality and their association with neighbourhood socioeconomic conditions in Guangzhou, China: A new approach using deep learning method and street view images" for publication as a full-length research paper in the journal, *Sustainable Cities and Society*.

We thank you and the reviewers for the insightful suggestions and comments, which have helped us to substantially improve the manuscript. The manuscript has been thoroughly revised in accordance with the reviewers' suggestions; each of these revisions is discussed in the following pages. We hope that our revised manuscript meets the high standards that you have for your journal.

Thank you for your reconsideration.

Sincerely,
Ruoyu Wang

Institute of Geography, School of GeoSciences, University of Edinburgh, Edinburgh, UK; R.Wang-54@sms.ed.ac.uk

The distribution of greenspace quantity and quality and their association with neighbourhood socioeconomic conditions in Guangzhou, China: A new approach using deep learning method and street view images

Highlights

- This study develops a new method to assess greenspace quality using deep learning method and street view images.
- Street view-based greenspace quality is validated through field audit methods and quality prediction of additional images.
- Street view greenness quality is more sensitive to the change of neighbourhood SES than street view greenness quantity and NDVI.

ABSTRACT

Awareness is mounting that urban greenspace is beneficial for residents' health. While a plethora of studies have focused on greenspace quantity, scant attention has been paid to greenspace quality. Existing methods for assessing greenspace quality is either highly labor-intensive and/or prohibitively time-consuming. This study develops a new machine learning method to assess greenspace quality based on street view images collected from Guangzhou, China. It also examines whether greenspace exposure disparities are linked to the neighbourhood socioeconomic status (SES). The validation process indicated that our scoring system achieved high accuracy for predicting street view-based greenspace quality outside the training data. Results also show that there were marked differences in spatial distribution between aggregated NDVI (Normalized Difference Vegetation Index), street view greenness quantity and quality. Regression models show that neighbourhood SES is not associated with NDVI. Although neighbourhood SES is associated with both street view greenness quantity and quality index value, street view greenness quality is more sensitive to the change of neighbourhood SES. Our work suggests that policymakers and planners are advised to pay more attention to greenspace quality and greenspace exposure disparities in urban area.

Keywords

Greenspace; Socioeconomic conditions; Street view; Machine learning; Environmental disparity; China

1. Introduction

Awareness is mounting that urban greenspace is beneficial for residents' health (Gascon et al., 2015; Hartig et al., 2014; Markevych et al., 2017; Nieuwenhuijseng et al.,, 2017; Wu et al., 2020). Several meta-reviews identify three main potential pathways through which greenness exposure promotes health (Gascon et al., 2015; Markevych et al., 2017). First, greenspace can reduce people's exposure to environmental stressors such as air pollution, noise and heat waves (Dadvand et al., 2015; Dzhambov et al.,, 2018a; Dzhambov et al., 2018b). Second, greenspace can restore people's capacities. Attention restoration theory (ART) (Kaplan, 1995) and stress reduction theory (SRT) (Ulrich et al., 1991) suggest that greenspace can restore people's attention by reducing stress and pressure. Last, greenspace can build people's capacities such as encouraging more physical activity and facilitating social cohesion (Su et al., 2016; Wang et al., 2019a; Yang et al., 2019; Yang et al., 2020).

Previous studies have reported environmental inequities in terms of urban greenery exposure in developed countries (Li et al., 2016; Xu et al., 2018; Zhou and Kim, 2013). Since greenspace can have potential health benefits, environmental disparities for greenspace such as unequal access or exposure to greenspace may result in disproportionate health benefits for different social groups (Jensen et al., 2004; Li et al., 2016). This kind of environmental disparities is also often associated with neighbourhood socioeconomic conditions (Apparicio et al., 2012, 2017; Barbosa et al., 2007; Jensen et al., 2004; Li et al., 2016; Landry and Chakraborty, 2009). Normally, neighbourhoods of high SES (socioeconomic status) often have greater financial resources, cultural and social capital, and political influence to maintain and enhance greenspace (Li et al., 2016; Li et al., 2015a), and potentially have more demand for greenspace quality (Jim and Shan, 2013). As a result, residents in high socioeconomic status neighbourhoods have better access to greenspace both in terms of quantity and quality. For example, Li et al. (2016) found that neighbourhoods in Hartford, Connecticut, USA with higher income have more street greenery than those with lower income. Similarly, in a study of six cities in Illinois, USA. However, Mears et al. (2020) found that although deprived areas in England had better access to greenspace, the greenspace was usually smaller in size, and worse in quality. Due to higher population density, deprived areas were disadvantaged with lower per capita greenspace. Therefore, the association between greenspace exposure and SES is complex and not always consistent. In China, although most of the greenspace is public greenspace and is provided by the government, it still distributes unequally across different neighbourhoods in terms of SES (Sun et al., 2019; You, 2016). First, local government finance usually is associated with neighbourhood socioeconomic conditions (e.g. taxes from property management fees or rents), so neighbourhoods with higher SES are more likely to support its local government to provide sufficient and better greenspace. Also, since greenspace may increase the land value in China, so local governments may follow land-based development process and are keen to provide more greenspace in neighbourhoods with higher SES (Chen and Hu, 2015). Second, neighbourhoods with more greenspace have higher housing price or rent in China, so disadvantaged social groups are less likely to afford the properties there (Xiao et al., 2017a). For example, You (2016) found inequalities in greenspace provision are associated with neighbourhood-level SES in Shenzhen while Shen et al. (2017) pointed out that disparities of greenspace provision exist for neighbourhoods with different levels of SES in Shanghai.

Despite the growing awareness of the importance greenspace quantity for population health, the role of greenspace quality has received less attention (Brindley et al., 2019). Compared with greenspace quantity which is an objective characteristic, greenspace quality reflects more about people's subjective attitudes towards surrounding greenness (Brindley et al., 2019). van Dillen et al. (2012) indicates that quality tends to be a marker for local people's eagerness to use the greenspace and the affordances they gain from this utilization. Furthermore, a high aesthetical value is likely to improve the restorative experience which leads to reduction of stress. Previous epidemiological studies mainly focus on the effect of greenspace availability, access or quantity on health (Gascon et al., 2015; Markevych et al., 2017), so many scholars argued that future research on neighborhood greenspace and health should focus more on its quality rather quantity (Van Dillen et al., 2012). A limited number of studies have compared the health benefit of both quantity and quality of greenspace and found that the quality of greenspace is more relevant to residents' health outcomes (Astell-Burt et al., 2014; Francis et al., 2012; Van Dillen et al., 2012). For example, Francis et al. (2012) found that residents living in neighbourhoods with high quality greenspace had lower odds of psychological distress, but this association was insignificant with greenspace quantity. The reason may be that quality reflects people's perception of the greenspace which directly influences the actual use of greenspace. However, Mears et al. (2020) found that whilst some quantity and quality indicators were not prominently associated with health outcomes, which highlights an urgent need for research including different measures of greenspace exposure. The omission of quality in research is not only due to the vagueness in definition but also due to methodological issues in operationalisation. (Brindley et al., 2019). Existing studies normally used one of two methods to evaluate greenspace quality including questionnaires and SSO (systematic social observation) (de Vries et al., 2013; Feng and Astell-Burt, 2017; Van Dillen et al., 2012). Both methods have obvious limits including being labor-intensive, time-consuming and difficult to apply across a large study area (Lu, 2018).

Given the challenges of collecting information about green space quality at large scale for epidemiological analyses, recently there has been interest in developing new methods for auto-extracting spatial data that provide indicators of quality. Most notably, the recent development of machine learning approaches combined with online mapping data has enabled the automated extraction of sentiments from social media text such as Flickr and Twitter data (Brindley et al., 2019), and ground objects (i.e., trees and grasses) from interactive panoramas such as street view images (Helbich et al., 2019; Labib et al., 2020; Larkin and Hystad, 2019; Li et al., 2018; Lu, 2018; Toikka et al., 2020; Wang et al., 2019a). For example, Brindley et al. (2019) used social media text to extract people's sentiments towards greenspace in order to better capture an indicator of urban greenspace quality. Hence, street view images have already been used for assessing eye-level greenspace quantity (Helbich et al., 2019; Labib et al., 2020; Larkin and Hystad, 2019; Li et al., 2018; Lu, 2018; Toikka et al., 2020; Wang et al., 2019a). For example, Larkin and Hystad (2019) used different exposure measures of visible greenspace and found weak relationships between street view quantity and other greenspace measures. Helbich et al.(2019) used both NDVI (Normalized Difference Vegetation Index) and street view images to assess greenspace quantity and found only greenspace evaluated by street view images is associated with mental health. Street view images have proven to be useful for field observation, since people can evaluate the local environment based on ground objects in street view images (Wang et al., 2019b; Wang et al., 2019c; Yao et al., 2019), an approach known as virtual systematic social observation (Plascak et al., 2020). For example, Ye et al. (2019) used street view data and machine learning methods to assess street quality, while Zhang et al. (2018) applied the similar approach in identifying different urban perceptions. People's perception of greenspace quality is also based on different ground objects (e.g. the absence of trash cans), so besides assessing greenspace quantity, street view images can also be applied for quality assessment. Lu (2019) used Google street view images to assess greenspace quality and used field observation to validate the results. Their findings showed that the results from street view images is highly correlated with the results from field observation and therefore may be a potentially more efficient way for assessing greenspaace quality.

This study addresses some of these research needs and develops a new method to assess greenspace quality based on street view images collected from Guangzhou, China and a machine learning approach. It also focuses on greenspace exposure disparities in terms of urban greenspace quantity and quality that are linked to the neighbourhood socioeconomic status, which enables us to examine whether socioeconomic disadvantaged populations are exposed to poorer quantity and quality of greenspace. This study extends previous research in several respects. First, instead of only focusing on greenspace quantity, it develops a new method to assess greenspace quality based on street view data which is important for epidemiological studies. Second, although previous studies have tried to use street view images to assess greenspace quality, they still evaluate manually based on comparatively few images, which means their methods cannot be applied readily in large scale studies and may lead to bias. Our new approach relies on machine learning approach which can assess millions of images over a large scale. Third, it further compares greenspace exposure disparities in terms of neighbourhood socioeconomic status between quantity and quality.

2. Method

2.1. Study area

Our analysis was conducted in Guangzhou which is located at the Pearl River in mainland China, part of a metropolitan area with a population of more than 13 million people. Our study area was restricted to seven old districts in the Guangzhou city (Liwan, Yuexiu, Haizhu, Tianhe, Baiyun, Panyu and Huangpu District) (Fig. 1) for two reasons. First, new districts (Huadu, Conghua, Zengcheng and Nansha District) are included in Guangzhou only recently due to administrative order, so they are economically and socially separated from seven districts in main urban zone. Second, new districts tend to have substantially lower population/housing density and have fewer built-up areas, and therefore much less street view data were collected there. The focus of this study is on neighbourhood-level (primary administrative unit) and there are 1677 residential neighbourhoods (*juweihui*) in our study area (average neighborhood size= 1 km^2; average population= 5660 persons). A neighbourhood (juwei) usually consists of several gated communities (*xiaoqu*) and non-gated communities (*xiaoqu*), but a community is often too small and residents' daily activity space is not limited by the boundary of community. , so neighborhoods (juwei) should be the better analytical unit. Hence, using community with a small area of greenspace may lead to, underestimation of residents' greenspace exposure.

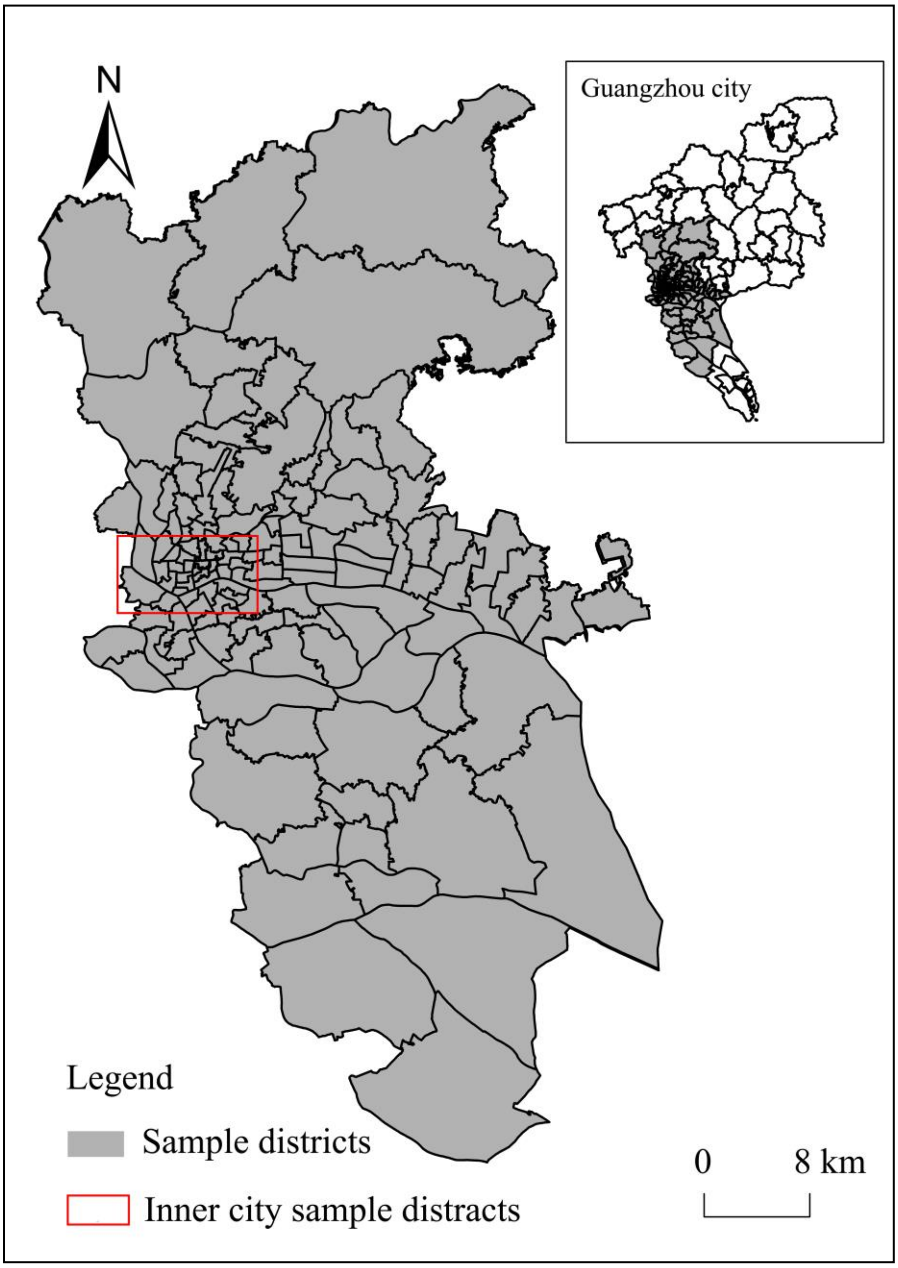


Fig. 1 Locations of the sampled districts in Guangzhou, China

2.2. Data collection

Street view images were collected from Tencent Online Map (downloaded through the Tencent Online Map API), the most comprehensive street view image database in China [https://map.qq.com/] (Helbich et al., 2019). Sampling points were created along each segment of the street network (obtained from OpenStreetMap (Haklay & Weber, 2008)) at 100m intervals following previous studies (Li et al., 2018). Street view images usually have a visual range of more than 50 meters, so 100m intervals can ensure that the street view images cover all the ground objects between any two sampling points. Following previous studies (Helbich et al., 2019; Wang et al., 2019a), we collected four images from four headings (i.e., 0, 90, 180, 270 degree) for each sampling point. In total, we defined 71,286 sampling points which led to 285,144 street view images. A calibration of image brightness was conduct to avoid side effects. In order to obtain a measure of over-head view greenspace, remote-sensing based greenspace was assessed in this study. The data were derived from the Landsat 8 Operational Land Imager and the Thermal Infrared Sensor at a spatial resolution of 30 meters in 2016 (USGS EarthExplorer: https://earthexplorer.usgs.gov/). Last, neighbourhood socioeconomic condition indicators was collected from the sixth census of Guangzhou in 2010, which is part of 2010 China's 10% population sample survey.

2.3. Machine-learning based image segmentation

In order to calculate an eye-level greenspace exposure, following previous studies (Helbich et al., 2019; Wang et al., 2019a), we extracted greenspace objects (e.g., grasses, trees) with a fully convolutional neural network for semantic image segmentation (FCN-8s) (Long et al., 2015) based on the ADE20K dataset of annotated images for training purposes (Zhou et al., 2017; Zhou et al., 2019). The accuracy of the FCN-8s was with 0.814 for the training data and 0.811 for the test data in this study.

2.4. Street view greenspace quantity and quality

Quantity

Following previous studies (Helbich et al., 2019; Wang et al., 2019a), street view greenspace quantity per sampling point was determined as the ratio of the number of greenspace pixels per image summed over the four cardinal directions to the total number of pixels per image summed over the four cardinal directions.

Quality

Fig 2 summarizes the workflow of assessing the street view greenspace quality. First, we constructed our training dataset. Specifically, 2000 images were randomly selected. Then, these images were scored (0 to 10) based on greenspace quality attributes by ten trained investigators who have resided in the research area for at least 3 years. As mentioned in literature review, there are various aspects of greenspace quality, so there are also many different operational items for assessing it. In order to get a robust greenspace quality indicator, we included a wide range of attributes of greenspace quality. The attributes (Cronbach's alpha=0.85) included accessibility (very bad-very good), maintenance (very bad-very good), variation (very monotonous-very varied), naturalness (very unnatural-very natural), colourfulness (very uncolourful-very colourful), clear arrangement (very unsurveyable-very surveyable), shelter (very enclosed-very open), absence of litter (very little trash-very much trash), safety (very unsafe-very safe) and general impression (very negative-very positive) (Lu, 2018; Van Dillen et al., 2012). This provided 10 attributes scores for 2000 images. In the next step, since people evaluate the neighbourhood environment based on ground objects, ground object elements within each street view images can be used to predict residents' perception of the local environment (Wang et al., 2019b; Wang et al., 2019c; Yao et al., 2019). Greenspace quality as one of people's perception of the local environment can also be evaluated through this way. After the image segmentation and the attributes scores of 2000 images, we calculated the proportion of each ground object elements. The random forest model (Breiman, 2001) for automatic rating was trained by fitting the inputted rating scores with the proportion of 151 elements (https://groups.csail.mit.edu/vision/datasets/ADE20K/) in the image segmentations. In this way, 151 ground elements within each image were automatically weighted based on the ten attributes scores. For example, with fewer trash can elements within an image, the absence of litter score would be higher for this image, and the trash can elements are given a weight based on ten attributes scores accordingly. Last, we used this automated scoring system to score all images in study area on these ten attributes.

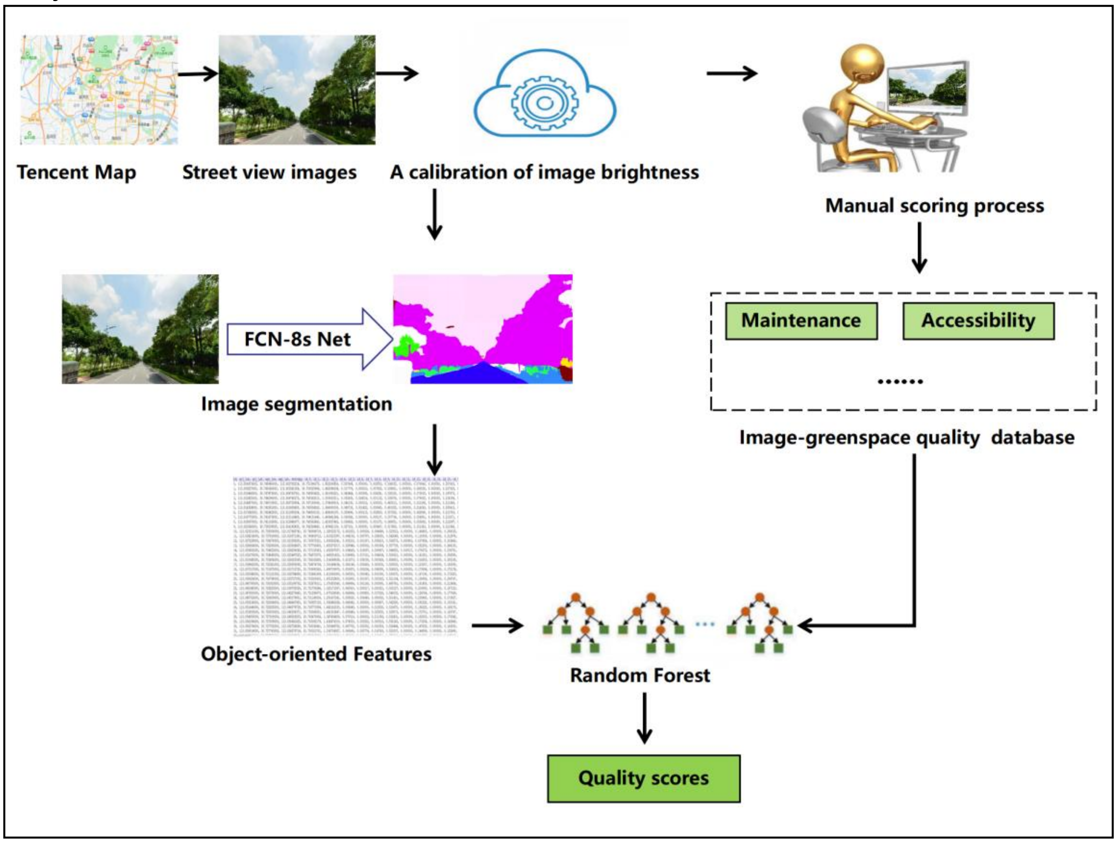


Fig 2 Workflow for assessing greenspace quality.

We took two steps to validate our results. First, one hundred Tencent view images were randomly selected, and attributes of greenspace quality of those images were again assessed manually. The scores from the automated scoring system were highly correlated with attributes of greenness quality of those manually assessed images: accessibility ($r=0.82$, $p<0.05$), maintenance ($r=0.87$, $p<0.05$), variation ($r=0.81$, $p<0.05$), naturalness ($r=0.95$, $p<0.05$), colourfulness ($r=0.92$, $p<0.05$), clear arrangement ($r=0.93$, $p<0.05$), shelter ($r=0.89$, $p<0.05$), absence of litter ($r=0.86$, $p<0.05$), safety ($r=0.98$, $p<0.05$) and general impression ($r=0.91$, $p<0.05$). This process indicates that our scoring system can achieve high accuracy for predicting ten attributes scores outside the training data.

Second, 26 residential neighbourhoods were randomly selected using a multi-stage stratified probability proportionate to population size (PPS) sampling technique and physically visited by three observers and audited with the same audit tool. The results showed reasonable inter-rater reliability (Pearson correlation $r > 0.90$; percentage agreement > 85%). We averaged the score from three observers and calculated the score based on our scoring system for these 26 residential neighbourhood using all the images within the neighbourhood. The correlation between the score from the field audit and scoring system was as follows: accessibility ($r=0.71$, $p<0.05$), maintenance ($r=0.73$, $p<0.05$), variation ($r=0.69$, $p<0.05$), naturalness ($r=0.72$, $p<0.05$), colourfulness ($r=0.68$, $p<0.05$), clear arrangement ($r=0.63$, $p<0.05$), shelter ($r=0.72$, $p<0.05$), absence of litter ($r=0.66$, $p<0.05$), safety ($r=0.88$, $p<0.05$) and general impression ($r=0.81$, $p<0.05$). This process indicates further that the score from our automatic scoring system was correlated with the results from field audit.

The above validation process suggested that our proposed method was suitable for measuring greenspace quality, so we collected scores of attributes of greenspace quality of those images for all images through the proposed automatic scoring system for all sampled neighbourhoods (1677 residential neighbourhoods). Ten attributes for all images achieved excellent internal consistency (Cronbach's alpha=0.88). Following previous studies (Lu, 2018; Van Dillen et al., 2012), the quality of greenspace in each image is the mean value of all 10 attributes. Thus, street view greenspace quality per sampling point was determined as the average greenspace quality score of four image from different cardinal directions. For each neighbourhood, the street view greenspace quality was calculated by the average score of all sampling point within the neighbourhood boundary.

2.5. Remote sensing greenspace quantity

Remote sensing greenspace quantity was assessed by NDVI (Normalized Difference Vegetation Index) (Tucker, 1979). We collected cloud-free images in the greenest season (i.e., June-August) to avoid assessment bias. The value of NDVI was calculated from the following formula: $(NIR - VIS)/(NIR + VIS)$, where NIR stood for reflectance in the near-infrared band and VIS stood for reflectance in the visible region. We aggregated the value of each pixel within the neighbourhood. The value of NDVI varies between −1 and 1and more positive values refer to greener vegetation within the neighbourhood.

2.6. Neighbourhood-level socioeconomic indicators

Following previous studies (Li et al., 2016; Li et al., 2015a; Li et al., 2015b), five socio-economic variables at the neighbourhood-level were selected to represent area-level SES from the 2010 population census data in Guangzhou. These include the proportion of residents with local hukou (registered permanent resident vs registered temporary resident), the proportion of residents with education attainment above high school, unemployment rate and the proportion of residents working in low status occupation, and per capita housing area. Due to the multicollinearity, we used Principal Component Analysis to combine all SES variables into a single indicator. We generated the correlation matrix and then calculated eigenvectors and eigenvalues. After that, we followed Kaiser-Guttman rule and chose one principal component with the largest eigenvalue, which accounted for 86 % of variance explained. The neighbourhood SES index ranged from 0.280 to 9.826. Higher scores mean higher levels of neighbourhood SES. In order to explore the nonlinear relationship between neighbourhood SES and greenness index, we treated neighbourhood SES index as quartile variable.

2.7. Covariates

We also controlled for a series of demographic and built environment variables. First, the proportion of residents aged 0-18, the proportion of residents aged above 65 and the proportion of married residents were controlled. Kabisch and Haase (2014) pointed out that age structure of the neighbourhood may have an influence on greenspace provision. Adolescents and elders have lower mobility than young adults, so they are more likely to benefit more from greenspace within neighbourhood which may be considered in the policy of local greenspace provision (Barbosa et al., 2007). Also, residents aged 0-18 are unemployed and residents aged above 65 are retired, so tend to spend a greater proportion of their day in their local neighbourhood may also influence neighbourhood SES. Second, population density is associated with supply of public facilities and dense neighbourhood may have lower SES since poorer residents are more likely to reside in higher density areas (Liu and Wu, 2006), so we controlled for the average score of this variable for each neighbourhood. Last, the proportion of residents living in houses built before 1979 was included. Old neighbourhoods in China have lower Floor Area Ratio which indicates they have more open space for greenspace (Xiao et al., 2017a). In China, residents living in house built before 1979 usually work in the state sectors which also influence their SES (He et al., 2010; Wu, 2007).

Table 1
Summary statistics for all variables.

Variables (number of neighbourhood=1677)	Mean (Standard Deviation)
Dependent variable	
NDVI	0.105(0.050)
SVG-quantity	0.201(0.089)
SVG-quality	5.560(0.572)
Independent variable	
Neighbourhoods SES quartile (Q1)	0.986(0.210)
Neighbourhoods SES quartile 2 (Q2)	1.416(0.088)
Neighbourhoods SES quartile 3 (Q3)	1.767(0.127)
Neighbourhoods SES quartile 4 (Q4)	2.482(0.577)
Covariates	
Population density (person/km^2)	30704.401(31974.252)
The proportion of residents aged 0-18	0.147(0.046)

The proportion of residents aged above 65	0.082(0.051)
The proportion of residents living in houses built before 1979	0.011(0.014)

2.8. Analysis

2.8.1 Global Moran's I

In order to identify the spatial distribution of greenspace characteristics, we examine the spatial autocorrelation of neighbourhood greenspace quantity and quality . Global Moran's I (Moran, 1950) was used to reflect the overall level of spatial autocorrelation of neighbourhood greenspace quantity and quality.

It was calculated as follow:

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \cdot \sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (1)$$

In the equation, x_i、 x_j are the level of greenspace quantity or quality in neighbourhood i 、 j ,

w_{ij} is the spatial weight matrix (queen contiguity)， n is the number of neighbourhood in study area. The higher absolute value of Global Moran's I indicates higher level of spatial autocorrelation of neighbourhood greenspace. If Global Moran's I is positive, neighbourhoods with higher level of greenspace cluster with each other and neighbourhoods with lower level of greenspace also cluster with each other. However, if Global Moran's I is negative, neighbourhoods with higher level of greenspace cluster with neighbourhoods with lower level of greenspace while lower level of greenspace cluster with neighbourhoods with higher level of greenspace.

2.8.2 Local Moran's I

Global Moran's I only reflects the overall level of spatial autocorrelation of neighbourhood greenspace, but further we need to identify the spatial autocorrelation level of greenspace for each neighbourhood, so that we can map it and examine the regions where spatial autocorrelation of greenspace is significant for most of the neighbourhoods. We used Local Moran's I (Anselin, 1995) to reflect the spatial relevance of greenspace quantity or quality in each neighbourhood to its neighbors. Local Moran's I reveals the degree of spatial difference and significance between the greenspace quantity or quality of each neighbourhood and its surrounding neighbourhood. It was calculated as follow:

$$I_i = z_i \sum_i W_{ij} z_j \quad (2)$$

In the equation: z_i and z_j are the standardized value of greenspace quantity or quality in neighbourhood i and j ; W_{ij} is the spatial weight matrix($\sum_j W_{ij} = 1$). If $I_i > 0$ and $z_i > 0$, then neighbourhood will be defined as high-high (H-H) region (significant cluster of high values); If $I_i < 0$ and $z_i < 0$, then neighbourhood will be defined as low-low (L-L) region (significant cluster of low values); If $I_i < 0$ but $z_i > 0$, then neighbourhood will be defined as high-low (H-L) region (significant cluster of outliers in which a high value is surrounded primarily by low values); If $I_i > 0$ but $z_i < 0$, then neighbourhood will be defined as low-high (L-H) region (significant cluster of outliers in which a low value is surrounded primarily by high values).

2.8.3 Spatial regression model

In order to link greenspace quantity and quality to neighbourhood socioeconomic conditions spatial regression model was used. It includes spatial lag model (SLM) and spatial error model (SEM) (Cliff and Ord, 1972). If spatial dependence of greenspace exists, OLS (ordinary least squares) models may cause bias. SLM. SLM has nested spatial dependence in dependent variables and the parameter estimation of independent variables while SEM has the parameter estimation of independent variables and error terms, so besides OLS we also adopted SLM and SEM to estimate the relationship between neighbourhood-level socioeconomic and demographic variables on neighbourhood greenspace quantity or quality. It was calculated as follow:

$$y_i = \rho \sum_{j=1}^{n} w_{ij} y_j + \beta x_i + \varepsilon_i \quad (3)$$

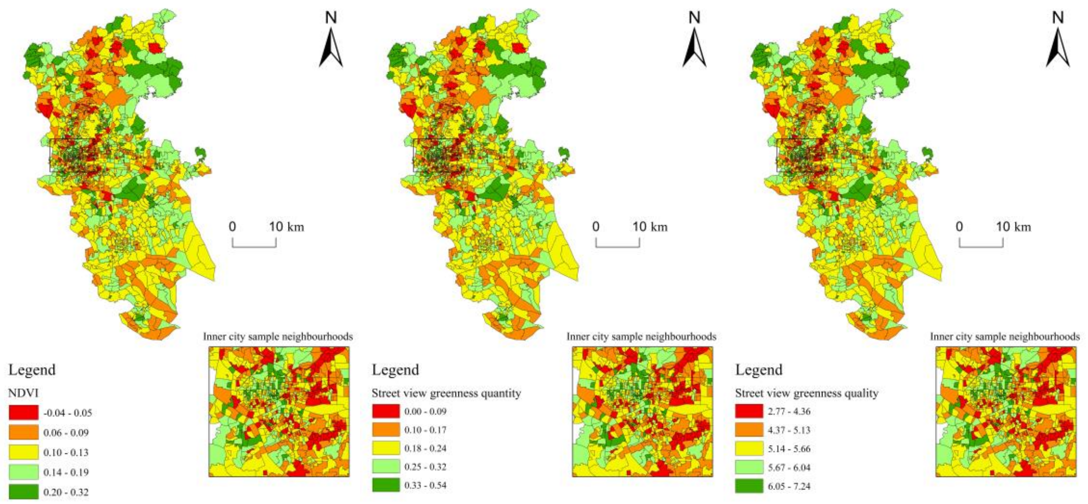
$$y_i = \beta x_i + \delta \sum_{j=1}^{n} w_{ij} \varepsilon_i \quad (4)$$

In the equation: y_i is the level of greenspace quantity or quality in neighbourhood i ; y_j is the level of greenspace quantity or quality in neighbourhood j ; ρ is the spatial autocorrelation coefficient; w_{ij} is the spatial weight matrix; β is the coefficient of independent variables. x_{it} is the value of independent variables in neighbourhood i ; δ is the coefficient of spatial lag explanatory variable, ε_i is the error term. n is the number of neighbourhood in study area.

3. Results

3.1. The distribution of greenspace quantity and quality

Fig. 3 shows the spatial distribution of the aggregated greenspace metrics at the neighbourhood level using NDVI (Fig. 3a), street view greenness quantity (Fig. 3b) and street view greenness quality (Fig. 3c), respectively. Compared with aggregated street view greenness, aggregated NDVI was relatively low in inner-city neighbourhoods. Fig. 4 shows the details of three greenery measurements in study area. In Fig. 4a, the sampled neighbourhoods were low in NDVI, but high in street view greenness. This is likely because neighbourhoods in the inner-city have a large number of street greenery which can be viewed by pedestrians but which are difficult to identify remotely overhead. Also, although street view greenness quantity and quality were both high in inner-city neighbourhoods, there was still a spatial mismatch between them. For example, in Fig. 4b, the sampled neighbourhoods were low in street view greenness quality, but high in street view greenness quantity. In such neighbourhoods, although street greenery was adequate, the green space in the surrounding environment tended to be of low quality. Hence, in Fig. 4c, the sampled neighbourhoods were low in street view greenness quantity, but high in street view greenness quality. Such neighbourhoods are usually in suburb or wealthy downtown area where people can afford villa (larger housing). In such neighbourhoods, street greenery is usually less common but well maintained.



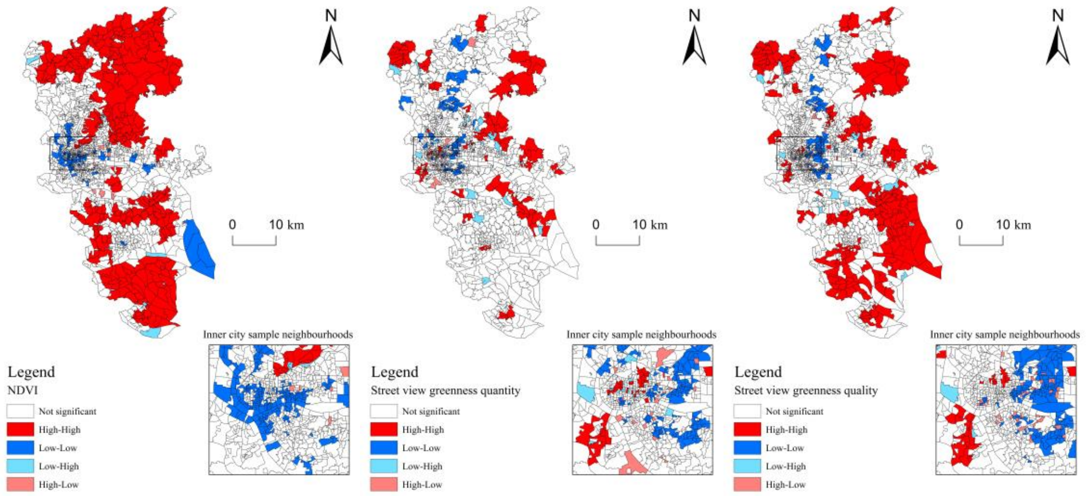
(a) (b) (c)

Fig 3. The distribution of the aggregated green view index values at the neighbourhood level (Natural Breaks): (a)NDVI; (B)Street view greenness quantity; (C)Street view greenness quality

(a) (b) (c)

Fig 4. Comparing three greenery measurements (A) Low level of NDVI and high level of street view greenness; (B) Low level of street view greenness quality and high level of street view greenness quantity; (C) High level of street view greenness quality and low level of street view greenness quantity.

Global Moran's I for distribution of the three aggregated green view index values at the neighbourhood level were all positive and significant at 5% significance level which indicates that distribution of the three aggregated green view index values had positive spatial dependence and space gathering. Figs. 5 displays local Moran's I values in relation to the three types of aggregated green view index values at the neighbourhood level. We only focused on HH and LL clusters, since HL and LH clusters only make up only a small part. Fig. 5a shows that similar to the spatial distribution of the aggregated of NDVI, HH clusters of NDVI were in suburb while LL clusters were in inner-city. Fig. 5b shows that HH, LL clusters of street view greenness quantity all could be observed in inner-city while HH clusters could also be observed in suburb. Fig. 5c shows that similar to street view greenness quantity, HH, LL clusters of street view greenness quality could also be observed in inner-city while HH clusters could be observed in suburb. However, unlike street view greenness quantity, local Moran's I values in relation to the street view greenness quality in inner-city showed spatial characteristic that HH clusters were in the west of inner-city while LL clusters were in the east.



(a) (b) (c)

Fig 5 LISA (Local Indicators of Spatial Association) cluster map of distribution of the aggregated green view index values at the neighbourhood level: (a)NDVI; (B)Street view greenness quantity; (C)Street view greenness quality

3.2. The relationship between three types of aggregated green view index values and neighbourhoods' socioeconomic conditions.

Table 2 shows the Spearman coefficients among three types of aggregated green view index values. The Spearman coefficients of street view greenness quantity versus street view greenness quality and the NDVI showed associations of 0.25 ($p<0.1$) and 0.31 ($p>0.1$). The Spearman coefficients of street view greenness quantity and quality showed associations of 0.78 ($p<0.1$). None of the Spearman coefficients among three types of aggregated green view index values was significant at 5% significance level which suggests they measure different aspects of urban greenness.

Table 2

Results of correlation test for different greenspace measures.

	NDVI	Street view greenness quantity	Street view greenness quality
NDVI	1		
Street view greenness quantity	0.25*	1	
Street view greenness quality	0.31	0.78*	1

*p < 0.10, **p < 0.05, ***p < 0.01.

Table 3, 4 and 5 show the relationship between three types of aggregated green view index values and neighbourhoods' socioeconomic conditions using OLS, SLM and SEM. Compared with OLS, SLM and SEM had higher Adjusted R^2. Also, robust LMLAG values were all significant for three aggregated green view index at 5% significance level while robust LMERR values were not, so we only focused on SLM in Table 3, 4 and 5. SLM in Table 3 showed that neighbourhood SES was not associated with aggregated NDVI values while population density and the proportion of residents aged above 65 were negatively associated with aggregated NDVI values. The significance of Lag Coeff(Rho) indicated that the distribution of aggregated NDVI had positive spatial dependence.

SLM in Table 4 showed that neighbourhood in Q4 SES status had higher aggregated street view greenness quantity values than those in Q1. Also, population density was positively associated with aggregated street view greenness quantity values. The significance of Lag Coeff(Rho) indicated that the distribution of aggregated street view greenness quantity had positive spatial dependence.

SLM in Table 5 showed that neighbourhood in Q2, Q3 and Q4 SES status had higher aggregated street view greenness quality values than those in Q1. Also, population density was positively associated with aggregated street view greenness quality values. The significance of Lag Coeff(Rho) indicated that the distribution of aggregated street view greenness quality had positive spatial dependence.

Table 3

Regression models of NDVI for neighbourhoods in inner-city area, Guangzhou, China.

	OLS		SLM		SEM	
	Coef.(SE)		Coef.(SE)		Coef.(SE)	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Population density		-0.040***(0.002)		-0.017***(0.001)		-0.026***(0.002)
The proportion of residents aged 0-18		0.063*(0.036)		0.024(0.020)		0.039*(0.020)
The proportion of residents aged above 65		-0.076***(0.025)		-0.052***(0.019)		-0.087***(0.021)
The proportion of residents living in house built before 1979		-0.101(0.086)		0.099(0.066)		0.101(0.073)
SES status (referenced: Q1)						
Q2	0.002*(0.001)	0.006*(0.003)	0.002(0.002)	0.003(0.002)	0.002(0.002)	0.002(0.002)
Q3	0.008(0.005)	0.011(0.013)	0.004(0.003)	0.008(0.012)	0.007(0.005)	0.007(0.013)
Q4	0.002(0.003)	0.005(0.003)	0.003(0.003)	0.005(0.003)	0.005(0.003)	0.005(0.003)
Constant	0.099***(0.002)	0.288***(0.009)	0.022***(0.002)	0.119***(0.008)	0.096***(0.003)	0.217***(0.010)
Lag Coeff(Rho)			0.754***(0.016)	0.633***(0.019)		
Lag Coeff(Lambda)					0.757***(0.015)	0.680***(0.019)
R2	0.017	0.363	0.602	0.623	0.602	0.622
AIC	-5297.35	-6011.8	-6508.71	-6700.35	-6508.69	-6660.33
Robust LMLAG			15.327***	75.726***		
Robust LMERR					0.386	11.122***

Coeff. = coefficient; SE = standard error. *p < 0.10, **p < 0.05, ***p < 0.01.

Table 4

Regression models of street view greenness quantity for neighbourhoods in inner-city area, Guangzhou, China.

	OLS Coef.(SE)		SLM Coef.(SE)		SEM Coef.(SE)	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Population density		0.017***(0.003)		0.009***(0.003)		0.010**(0.004)
The proportion of residents aged 0-18		0.038(0.057)		0.040(0.051)		0.053(0.052)
The proportion of residents aged above 65		0.046(0.055)		0.034(0.049)		0.032(0.053)
The proportion of residents living in house built before 1979		0.090(0.189)		0.093(0.169)		0.043(0.184)
SES status (referenced: Q1)						
Q2	0.008(0.006)	0.009(0.006)	0.004(0.005)	0.004(0.006)	0.001(0.005)	0.001(0.006)
Q3	0.016(0.016)	0.016(0.011)	0.015(0.015)	0.014(0.016)	0.017(0.017)	0.016(0.017)
Q4	0.027***(0.006)	0.022***(0.008)	0.021***(0.005)	0.018***(0.007)	0.023***(0.006)	0.020***(0.007)
Constant	0.188***(0.004)	0.260***(0.020)	0.093***(0.006)	0.131***(0.019)	0.1905***(0.005)	0.233***(0.022)
Lag Coeff(Rho)			0.484***(0.025)	0.475***(0.026)		
Lag Coeff(Lambda)					0.488***(0.031)	0.479***(0.026)
R2	0.010	0.022	0.217	0.219	0.219	0.221
AIC	-3349.14	-3365.47	-3635.91	-3635.75	-3640.43	-3637.73
Robust LMLAG			3.691**	3.593**		
Robust LMERR					0.044	0.068

Coeff. = coefficient; SE = standard error. *p < 0.10, **p < 0.05, ***p < 0.01.

Table 5

Regression models of street view greenness quality for neighbourhoods in inner-city area, Guangzhou, China.

	OLS Coef.(SE)		SLM Coef.(SE)		SEM Coef.(SE)	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Population density		0.274***(0.020)		0.130***(0.018)		0.176***(0.026)
The proportion of residents aged 0-18		0.592*(0.347)		0.336(0.289)		0.383(0.297)
The proportion of residents aged above 65		0.145(0.331)		0.150(0.276)		-0.037(0.305)
The proportion of residents living in house built before 1979		0.540(1.149)		0.534(0.959)		0.342(1.055)
SES status (referenced: Q1)						
Q2	0.052**(0.024)	0.061**(0.029)	0.026**(0.011)	0.030**(0.012)	0.035**(0.016)	0.031**(0.012)
Q3	0.074**(0.036)	0.056**(0.022)	0.073**(0.031)	0.064***(0.025)	0.101***(0.033)	0.088**(0.037)
Q4	0.226***(0.039)	0.106**(0.047)	0.169***(0.031)	0.116***(0.039)	0.204***(0.035)	0.166***(0.043)
Constant	5.472***(0.027)	6.547***(0.119)	2.111***(0.119)	2.887***(0.177)	5.473***(0.035)	6.227***(0.134)
Lag Coeff(Rho)			0.608***(0.021)	0.558***(0.023)		
Lag Coeff(Lambda)					0.614***(0.021)	0.575***(0.023)
R2	0.019	0.122	0.379	0.388	0.383	0.39
AIC	2861.71	2680.54	2273.17	2225.79	2263.29	2228.21
Robust LMLAG			8.979***	21.663***		
Robust LMERR					0.195	1.065

Coeff. = coefficient; SE = standard error. *p < 0.10, **p < 0.05, ***p < 0.01.

4. Discussion

This study extends previous research on the distribution of different aspects of greenspace and the association between greenspace and neighbourhood socioeconomic conditions in several respects. First, it develops a new method assessing greenspace quality based on street view data and machine learning approach. Compared with traditional measures for assessing greenspace quality, the proposed method is less labor-intensive and more efficient, so it could offer a major step forward in terms of understanding the role of greenspace quality for public health which is important for large scale epidemiological studies. Second, it compares the distribution of greenspace quantity and quality in Guangzhou, China using both street view images (street view greenness) and remote sensing data (NDVI). Results show that there is a difference between the distribution of greenspace quantity and quality. Third, it further investigates the association between greenspace quantity, quality and neighbourhoods' socioeconomic conditions. Compared with NDVI, street view greenness shows greater inequities associated with neighbourhoods' socioeconomic conditions.

4.1. The spatial mismatch between greenspace quantity and quality

We found that NDVI was relatively more abundant in the suburbs, but less abundant in the inner-city. Similar to the spatial distribution of the aggregated of NDVI, LISA cluster map shows that HH clusters of NDVI were in suburb while LL clusters were in inner-city. These findings are consistent with previous studies in the USA (Li et al., 2016; Li et al., 2015a; Li et al., 2015b) and Singapore (Ye et al., 2018). This may be because inner-city areas of Chinese cities tend to be dominated by commercial and residential amenities with few large green facilities. Also, NDVI is more accurate in measuring large green facilities such as park which are more common in the suburbs. As for street view greenness quantity and quality, they are both abundant in the suburb and inner-city. Also, HH, LL, HL and LH clusters of street view greenness all could be observed in inner-city while HH clusters could also be observed in suburb. Ye et al. (2018) and Li et al. (2015, 2016) drew similar conclusions for street view greenness quantity in developed countries. This may be because there is also a demand for greenspace for residents living in inner-city, but since there is little space there, greenspace in inner-city may be more likely to be street plants which can be perceived by pedestrians (Ye et al., 2018). However, street view greenness quality in inner-city shows spatial characteristic that HH clusters were in the west of inner-city while LL clusters were in the east in Guangzhou. A possible explanation is that the west of inner-city in Guangzhou is mainly old town while the east is new town. Residents living in old town in Guangzhou are more likely to own their house while residents living in new town in Guangzhou are more likely to rent (Chen, 2016), and it has been showed by previous studies that owners tend to spend more on maintaining greenspace amenities than renters (Heynen et al. 2006; Perkins et al., 2004).

This study also shows the details of difference among three greenery measurements in inner-city sampled neighbourhoods (Fig. 4). First, inner-city sampled neighbourhoods which are low in

NDVI, but high in street view greenness usually distribute in inner-city which is consistent with the finding in Singapore (Ye et al., 2018). This is due to the high density of street plants and low density of large green facilities in inner-city area. Second, inner-city sampled neighbourhoods which are low in street view greenness quality, but high in street view greenness quantity also usually distribute in inner-city. In such neighbourhoods, street greenery is not well maintained and perceived to be low quality. Third, inner-city sampled neighbourhoods which are low in street view greenness quantity, but high in street view greenness quality usually distribute in both inner-city and suburb. A possible explanation is that these two areas both have villa districts where street plants are not flourish but maintained well. Thus, the spatial mismatch between eye-level and over-head view greenspace as well as between greenspace quantity and quality, indicates that it is important to precisely identify *which* aspect of greenspace should be measured.

4.2. The association between aggregated green view index values and neighbourhood socioeconomic conditions

Consistent with previous studies (Helbich et al., 2019; Wang et al., 2019a), there was no evidence to suggest that NDVI was associated with street view greenness. Nor was there evidence that street view greenness quantity was associated with street view greenness quality. These results indicated that three types of aggregated greenspace index values measure different aspects of greenspace. NDVI measures top-down greenspace, so it does not reflect the cover of street plants (Helbich et al., 2019; Ye et al., 2018). Also, it is a global measurement and covers areas further away from roads, which are less likely to be green. Another explanation is that the resolution of the NDVI data is too coarse to reflect actual top-down greenspace. Therefore, it is important to acknowledge that this study cannot provide definitive evidence that SES disparities are evident, but rather emphasise that NDVI and street view measures identify different components of greenspace. While street view greenness quantity measures human-scale greenspace, it mainly measures eye-level street plants (Helbich et al., 2019; Ye et al., 2018). Compared with greenspace quantity, greenspace quality is more subjective which reflects people's evaluation of greenspace (Brindley et al., 2019). Although few studies have examined the association between greenspace quantity and greenspace quality, a limited number of studies have found that they may not be significantly correlated with each other, so they may measure different aspects of greenness (Lu, 2018; Van Dillen et al., 2012). Therefore, the above three measures capture slightly different aspects of greenspace and the 'best' measure would depend on the question being posed.

Regression results show that the proportion of residents aged above 65 and the proportion of married residents are negatively associated with the aggregated NDVI index values. However, none of the demographic variables are associated with aggregated street view greenness index values. For built environment variables, only population density is associated with three aggregated greenness index values. Population density is negatively associated with aggregated NDVI index values while it is positively associated with aggregated street view greenness index values. This may be because NDVI is more abundant in the suburbs than the inner-city where population density is higher, while street view greenness is more abundant in the inner-city, but less abundant in suburbs where population density is lower. Most importantly, there was no evidence to suggest SES status has association with the aggregated NDVI index values. However, neighbourhood with higher SES had signifificantly street view greenness quantity values (i.e., 4^{th} quartile) and quality values (i.e., 2^{nd}, 3^{rd} and 4^{th} quartile). These findings indicate that neighbourhoods with higher SES are more likely to have more street view greenspace which is consistent with previous studies (Li et al., 2016; Li et al., 2015a; Li et al., 2015b). This may be because people living in neighbourhood with higher SES are more likely to be able to afford the costs associated with maintaining greenspace (Heynen et al., 2006; Perkins et al., 2004).

In terms of three measurements of greenery, street view greenness quality is more sensitive to the change of SES status (2^{nd}, 3^{rd} and 4^{th} quartile) than quantity (only 4^{th} quartile). On the contrary, NDVI is not associated with SES status. This further indicates the environmental inequity for street view greenness especially street view greenness quality can be explained by neighbourhood SES inequity. A possible explanation is that NDVI can measure large green facilities such as park and their distribution can be planed more equally through urban planning process in China (Xiao et al., 2016; Xiao et al., 2017a). However, street view greenness captures street plants which are more likely to be influenced by local residents who pay the cost of maintaining neighbourhood greenspace (Xiao et al., 2017b). Also, compared with greenspace quantity, environmental inequity for greenspace quality is more associated with neighbourhood SES inequity. This may be because compared with greenspace quantity, maintaining greenspace quality may be more expensive since greenspace quality is associated with surrounding environments (Van Dillen et al., 2012).

4.3. Strengths and limitations

This study has several strengths. First, it is the first attempt to combine street view data with a machine learning approach for assessing greenspace quality in a large scale. Second, we not only compared greenspace quality with quantity, but also compared street view-based greenspace with remote sensing-based greenspace. This allows an examination of greenspace measurement from different perspectives. Third, we also focused on greenspace exposure disparities in terms of urban greenspace quantity and quality in relation to the neighbourhood socioeconomic status. This helps us in further understanding greenspace exposure disparities in the Chinese context.

Our study also had some limitations. First, the boundary of neighbourhood is defined using administrative boundaries and the findings may be affected by the Modifiable Areal Unit Problem (MAUP) (Fotheringham and Wong, 1991). Second, our training sample size was based on 2000 images and it may still be too small. Third, we only focused on greenspace in urban areas in Guangzhou which may not be generalizable to other parts of China, including rural areas. Fourth, aggregated green view indexes were derived in 2016 while neighbourhood socioeconomic conditions variables were derived in 2010, so socioeconomic changes within local areas is not captured. Fifth, since images were collected on a specific date, they may not be sufficiently representative to reflect dynamic changes of greenspace through time. This may be a concern as Guangzhou has a subtropical climate, so most of its vegetation stays green all year round. Still, the static and time specific street view images fail to capture seasonal fluctuation greeneryand may lead to potential bias. Sixth, we only include images from horizontal angle (vertical angle=0), which may not be enough to reflect actual visible greenery. Seventh, local attributes of greenspace quality are usually captured using data collected within urban parks, whereas information collected outside of park boundaries provides indicators that are more closely related to neighbourhood street quality rather than greenspace per se. Therefore, no single measure of greenspace quality fully captures the breadth of greenspace quality; further research could usefully assess the appropriateness of various quality measures in different contexts. Eighth, our approach may not be able to identify all kinds of greenness, since greenness is heterogeneous and there are substantial differences within various kinds of greenness. Ninth, the variation in image quality (i.e. saturation, tint and clarity) may cause potential bias. Last, there are some gated communities in our study area where street view images may not be collected.

4.4. Policy implication and issues for further research

Since greenspace quality may be more relevant to residents' health, policy makers should pay attention to the attributes for greenspace quality such as the accessibility, variation and safety of greenspace instead of only considering greenspace quantity. For example, it may be beneficial to enhance the diversity of street vegetation sin order to promote variation and naturalness. Second, greenspace exposure disparities should also attract more attention, since people living in lower SES areas are less likely to be exposed to greenspace and can only enjoy lower quality greenspaces, so as public facilities, more government investment in greenspace facilities to enhance the quality of these resources should be considered, particularly in ow SES neighbourhoods.. Last, relevant evaluation indicator of greenspace quality should be incorporated into urban planning policies and regulations in China.

Future research should pay attention to the following aspects: 1) There is value in attempting to enlarge the training sample size; online scoring system can be used for collecting people's perceived greenspace quality from across national contexts, so that researchers can further identify whether cultural background and other characters may have impact on assessment of greenspace quality. 2) Previous epidemiological studies mainly focus on the effect of greenspace quantity and future research should also consider greenspace quality using our new methods. 3) Since greenspace quantity and quality may reflect different aspects of greenspace, future research should examine whether the pathways through which greenspace quantity and quality have influence on health vary.

5. Conclusion

This study is the first to develop a new method to assess greenspace quality based on street view data and machine learning method. Further, it focuses on greenspace exposure disparities regarding how urban greenspace quantity and quality are linked to the neighbourhood socioeconomic status in the Chinese context. Results show that the distribution of aggregated NDVI, street view greenness quantity and quality index value shows significant differences that NDVI is relatively more abundant in the suburb, but less abundant in inner-city, while street view greenness is both abundant in the suburb and inner-city. Hence, the correlation analysis shows that no evidence can support three aggregated greenness indexes are significantly associated with each other which indicates that they measure different aspects of greenspace. In terms of three measurements of greenery, environmental inequity for street view greenness is more associated with neighbourhood SES inequity than NDVI. Also, street view greenness quality is more sensitive to the change of neighbourhood SES than quantity. To achieve the goal of promoting urban greening and health through urban planning and design in Chinese urban settings, policymakers and planners are advised to pay more attention to greenspace quality and greenspace exposure disparities in urban area.

Reference

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, *27*(2), 93-115.

Apparicio, P., Landry, S., & Lewnard, J. (2017). Disentangling the effects of urban form and socio-demographic context on street tree cover: A multi-level analysis from Montréal. Landscape and Urban Planning, 157, 422-433.

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.

Astell-Burt, T., Mitchell, R., & Hartig, T. (2014). The association between green space and mental health varies across the lifecourse. A longitudinal study. *J Epidemiol Community Health*, *68*(6), 578-583.

Barbosa, O., Tratalos, J. A., Armsworth, P. R., Davies, R. G., Fuller, R. A., Johnson, P., & Gaston, K. J. (2007). Who benefits from access to green space? A case study from Sheffield, UK. Landscape and Urban Planning, 83(2-3), 187-195.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Brindley, P., Cameron, R. W., Ersoy, E., Jorgensen, A., & Maheswaran, R. (2019). Is more always better? Exploring field survey and social media indicators of quality of urban greenspace, in relation to health. *Urban Forestry & Urban Greening*, *39*, 45-54.

Chen, G. (2016). The heterogeneity of housing-tenure choice in urban China: A case study based in Guangzhou. *Urban Studies*, *53*(5), 957-977.

Chen, W. Y., & Hu, F. Z. Y. (2015). Producing nature for public: Land-based urbanization and provision of public green spaces in China. *Applied Geography*, 58, 32-40.

Cliff, A., & Ord, K. (1972). Testing for Spatial Autocorrelation Among Regression Residuals. *Geographical analysis*, *4*(3), 267-284.

Dadvand, P., Rivas, I., Basagaña, X., Alvarez-Pedrerol, M., Su, J., Pascual, M. D. C., . . . Sunyer, J. (2015). The association between greenness and traffic-related air pollution at schools. *Science Of the Total Environment*, *523*, 59-63.

de Vries, S., Van Dillen, S. M., Groenewegen, P. P., & Spreeuwenberg, P. (2013). Streetscape greenery and health: stress, social cohesion and physical activity as mediators. *Social Science & Medicine*, *94*, 26-33.

Dzhambov, A., Hartig, T., Markevych, I., Tilov, B., & Dimitrova, D. (2018a). Urban residential greenspace and mental health in youth: Different approaches to testing multiple pathways yield different conclusions. *Environmental Research, 160*, 47-59.

Dzhambov, A. M., Markevych, I., Hartig, T., Tilov, B., Arabadzhiev, Z., Stoyanov, D., . . . Dimitrova, D. D. (2018b). Multiple pathways link urban green-and bluespace to mental health in young adults. *Environmental Research*, *166*, 223-233.

Feng, X., & Astell-Burt, T. (2017). Residential green space quantity and quality and child well-being: a longitudinal study. *American Journal of Preventive Medicine*, *53*(5), 616-624.

Fotheringham, A. S., & Wong, D. W. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment And Planning A*, *23*(7), 1025-1044.

Francis, J., Wood, L. J., Knuiman, M., & Giles-Corti, B. (2012). Quality or quantity? Exploring the relationship between Public Open Space attributes and mental health in Perth, Western Australia. *Social Science & Medicine, 74*(10), 1570-1577.

Gascon, M., Triguero-Mas, M., Martínez, D., Dadvand, P., Forns, J., Plasència, A., & Nieuwenhuijsen, M. J. (2015). Mental health benefits of long-term exposure to residential green and blue spaces: a systematic review. *International Journal Of Environmental Research And Public Health*, *12*(4), 4354-4379.

Haklay, M., & Weber, P. (2008). Openstreetmap: User-generated street maps. *IEEE Pervasive Computing*, *7*(4), 12-18.

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

He, S., Wu, F., Webster, C., & Liu, Y. (2010). Poverty concentration and determinants in China's urban low‐income neighbourhoods and social groups. International journal of urban and regional research, 34(2), 328-349.

Helbich, M., Yao, Y., Liu, Y., Zhang, J., Liu, P., & Wang, R. (2019). Using deep learning to examine street view green and blue spaces and their associations with geriatric depression in Beijing, China. *Environment International*, *126*, 107-117.

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.

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).

Jim, C., & Shan, X. (2013). Socioeconomic effect on perception of urban green spaces in Guangzhou, China. *Cities*, *31*, 123-131.

Kabisch, N., & Haase, D. (2014). Green justice or just green? Provision of urban green spaces in Berlin, Germany. *Landscape and Urban Planning*, 122, 129-139.

Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology, 15*(3), 169-182.

Labib, S. M., Huck, J. J., & Lindley, S. (2020). Modelling and mapping eye-level greenness visibility exposure using multi-source data at high spatial resolutions. *Science of The Total Environment*, 143050.

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.

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

La Rosa, D., Takatori, C., Shimizu, H., & Privitera, R. (2018). A planning framework to evaluate demands and preferences by different social groups for accessibility to urban greenspaces. *Sustainable cities and society*, 36, 346-362.

Li, X., Santi, P., Courtney, T. K., Verma, S. K., & Ratti, C. (2018). Investigating the association between streetscapes and human walking activities using Google Street View and human trajectory data. *Transactions in GIS*, 22(4), 1029-1044.

Li, X., Zhang, C., Li, W., & Kuzovkina, Y. A. (2016). Environmental inequities in terms of different types of urban greenery in Hartford, Connecticut. *Urban Forestry & Urban Greening, 18*, 163-172.

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

Liu, Y., & Wu, F. (2006). Urban poverty neighbourhoods: Typology and spatial concentration under China’s market transition, a case study of Nanjing. Geoforum, 37(4), 610-626.

Long, J., Shelhamer, E., & Darrell, T. (2015). *Fully convolutional networks for semantic segmentation.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Lu, Y. (2018). Using Google Street View to investigate the association between street greenery and physical activity. *Landscape And Urban Planning*.

Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A. M., . . . Nieuwenhuijsen, M. J. (2017). Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environmental Research*, *158*, 301-317.

Mears, M., Brindley, P., Jorgensen, A., & Maheswaran, R. (2020). Population-level linkages between urban greenspace and health inequality: The case for using multiple indicators of neighbourhood greenspace. *Health & Place*, 102284.

Mears, M., Brindley, P., Maheswaran, R., & Jorgensen, A. (2019). Understanding the socioeconomic equity of publicly accessible greenspace distribution: The example of Sheffield, UK. *Geoforum*, 103, 126-137.

Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*(1/2), 17-23.

Nieuwenhuijsen, M. J., Khreis, H., Triguero-Mas, M., Gascon, M., & Dadvand, P. (2017). Fifty Shades of Green: Pathway to Healthy Urban Living. *Epidemiology, 28*(1).

Perkins, H. A., Heynen, N., & Wilson, J. (2004). Inequitable access to urban reforestation: the impact of urban political economy on housing tenure and urban forests. *Cities*, *21*(4), 291-299.

Plascak, J. J., Rundle, A. G., Babel, R. A., Llanos, A. A., LaBelle, C. M., Stroup, A. M., & Mooney, S. J. (2020). Drop-And-Spin Virtual Neighborhood Auditing: Assessing Built Environment for Linkage to Health Studies. *American Journal of Preventive Medicine*, 58(1), 152-160.

Shen, Y., Sun, F., & Che, Y. (2017). Public green spaces and human wellbeing: Mapping the spatial inequity and mismatching status of public green space in the Central City of Shanghai. *Urban Forestry & Urban Greening*, 27, 59-68.

Sreetheran, M., Van Den Bosch, C. C. K., 2014. A socio-ecological exploration of fear of crime in urban green spaces – A systematic review. Urban For. Urban Green. 13(1), 1-18.

Su, S., Zhang, Q., Pi, J., Wan, C., & Weng, M. (2016). Public health in linkage to land use: Theoretical framework, empirical evidence, and critical implications for reconnecting health promotion to land use policy. *Land Use Policy*, 57, 605-618.

Sun, C., Lin, T., Zhao, Q., Li, X., Ye, H., Zhang, G., ... & Zhao, Y. (2019). Spatial pattern of urban green spaces in a long-term compact urbanization process—A case study in China. *Ecological indicators*, 96, 111-119.

Toikka, A., Willberg, E., Mäkinen, V., Toivonen, T., & Oksanen, J. (2020). The green view dataset for the capital of Finland, Helsinki. *Data in Brief*, 105601.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127-150.

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments 1. *Journal of Environmental Psychology*, *11*(3), 201-230.

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

Wang, R., Helbich, M., Yao, Y., Zhang, J., Liu, P., Yuana, Y., & Liu, Y. (2019a). Urban greenery and mental wellbeing in adults: Cross-sectional mediation analyses on multiple pathways across different greenery measures. *Environmental Research*.

Wang, R., Liu, Y., Lu, Y., Zhang, J., Liu, P., Yao, Y., & Grekousis, G. (2019b). Perceptions of built environment and health outcomes for older Chinese in Beijing: a big data approach with street view images and deep learning technique. *Computers, Environment and Urban Systems*.

Wang, R., Yuan, Y., Liu, Y., Zhang, J., Liu, P., Lu, Y., & Yao, Y. (2019c). Using street view data and machine learning to assess how perception of neighborhood safety influences urban residents' mental health. *Health & Place*.

Wu, F. (2007). The poverty of transition: from industrial district to poor neighbourhood in the city of Nanjing, China. Urban Studies, 44(13), 2673-2694.

Wu, W., Dong, G., Yeran, S. U. N., & Yun, Y. (2020). Contextualized effects of Park access and usage on residential satisfaction: A spatial approach. *Land Use Policy*, 94, 104532.

Xiao, Y., Li, Z., & Webster, C. (2016). Estimating the mediating effect of privately-supplied green space on the relationship between urban public green space and property value: Evidence from Shanghai, China. *Land Use Policy*, *54*, 439-447.

Xiao, Y., Lu, Y., Guo, Y., & Yuan, Y. (2017a). Estimating the willingness to pay for green space services in Shanghai: Implications for social equity in urban China. *Urban Forestry & Urban Greening*, *26*, 95-103.

Xiao, Y., Wang, Z., Li, Z., & Tang, Z. (2017b). An assessment of urban park access in Shanghai–Implications for the social equity in urban China. *Landscape And Urban*

Planning, *157*, 383-393.

Xu, C., Haase, D., Pribadi, D. O., & Pauleit, S. (2018). Spatial variation of green space equity and its relation with urban dynamics: A case study in the region of Munich. *Ecological Indicators*, 93, 512-523.

Yang, L., Liu, J., Lu, Y., Ao, Y., Guo, Y., Huang, W., ... & Wang, R. (2020). Global and local associations between urban greenery and travel propensity of older adults in Hong Kong. *Sustainable Cities and Society*, 102442.

Yang, Y., He, D., Gou, Z., Wang, R., Liu, Y., & Lu, Y. (2019). Association between street greenery and walking behavior in older adults in Hong Kong. *Sustainable Cities and Society*, 51, 101747.

Yao, Y., Liang, Z., Yuan, Z., Liu, P., Bie, Y., Zhang, J., . . . Guan, Q. (2019). A human-machine adversarial scoring framework for urban perception assessment using street-view images. *International Journal of Geographical Information Science*, 1-22.

Ye, Y., Richards, D., Lu, Y., Song, X., Zhuang, Y., Zeng, W., & Zhong, T. (2018). Measuring daily accessed street greenery: A human-scale approach for informing better urban planning practices. *Landscape And Urban Planning*.

Ye, Y., Zeng, W., Shen, Q., Zhang, X., & Lu, Y. (2019). The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images. *Environment and Planning B: Urban Analytics and City Science*, 46(8), 1439-1457.

You, H. (2016). Characterizing the inequalities in urban public green space provision in Shenzhen, China. *Habitat International*, 56, 176-180.

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning*, 180, 148-160.

Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralba, A. (2017). *Scene parsing through ade20k dataset.* Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Zhou, B., Zhao, H., Puig, X., Xiao, T., Fidler, S., Barriuso, A., & Torralba, A. (2019). Semantic understanding of scenes through the ade20k dataset. *International Journal of Computer Vision*, *127*(3), 302-321.

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.

The distribution of greenspace quantity and quality and their association with neighbourhood socioeconomic conditions in Guangzhou, China: A new approach using deep learning method and street view images

Ruoyu Wang*, Zhiqiang Feng, Jamie Pearce, Yao Yao, Li Xiaojiang, Ye Liu

Ruoyu Wang* (Corresponding author)
Institute of Geography, School of GeoSciences, University of Edinburgh, Edinburgh, UK; Email: R.Wang-54@sms.ed.ac.uk

Zhiqiang Feng
Institute of Geography, School of GeoSciences, University of Edinburgh, Edinburgh, UK; Email: zhiqiang.feng@ed.ac.uk

Jamie Pearce
Centre for Research on Environment Society and Health, School of GeoSciences, University of Edinburgh, Edinburgh, UK; Email: jamie.pearce@ed.ac.uk

Yao Yao
School of Geography and Information Engineering, China University of Geosciences, Wuhan, China; Email: yaoy@cug.edu.cn

Li Xiaojiang
MIT Senseable City Lab, Department of Urban Studies and Planning, Massachusetts Institute of Technology, Room 9-250, 77 Massachusetts Avenue, Cambridge, MA 02139, United States; Email: lixiaojiang.gis@gmail.com

Ye Liu
School of Geography and Planning, Sun Yat-sen University, Guangzhou, China; Email: liuye25@mail.sysu.edu.cn

Acknowledgement:

This work was supported by the National Natural Science Foundation of China [41801306] and the Innovative Research and Development Team Introduction Program of Guangdong Province (grant number 2017ZT07X355). We would like to thank the anonymous reviewer for their helpful and positive comments which have improved the quality of the paper.

ABSTRACT

Awareness is mounting that urban greenspace is beneficial for residents' health. While a plethora of studies have focused on greenspace quantity, scant attention has been paid to greenspace quality. Existing methods for assessing greenspace quality is either highly labor-intensive and/or prohibitively time-consuming. This study develops a new machine learning method to assess greenspace quality based on street view images collected from Guangzhou, China. It also examines whether greenspace exposure disparities are linked to the neighbourhood socioeconomic status (SES). The validation process indicated that our scoring system achieved high accuracy for predicting street view-based greenspace quality outside the training data. Results also show that there were marked differences in spatial distribution between aggregated NDVI (Normalized Difference Vegetation Index), street view greenness quantity and quality. Regression models show that neighbourhood SES is not associated with NDVI. Although neighbourhood SES is associated with both street view greenness quantity and quality index value, street view greenness quality is more sensitive to the change of neighbourhood SES. Our work suggests that policymakers and planners are advised to pay more attention to greenspace quality and greenspace exposure disparities in urban area.

Keywords

Greenspace; Socioeconomic conditions; Street view; Machine learning; Environmental disparity; China

1. Introduction

Awareness is mounting that urban greenspace is beneficial for residents' health (Gascon et al., 2015; Hartig et al., 2014; Markevych et al., 2017; Nieuwenhuijseng et al.,, 2017; Wu et al., 2020). Several meta-reviews identify three main potential pathways through which greenness exposure promotes health (Gascon et al., 2015; Markevych et al., 2017). First, greenspace can reduce people's exposure to environmental stressors such as air pollution, noise and heat waves (Dadvand et al., 2015; Dzhambov et al.,, 2018a; Dzhambov et al., 2018b). Second, greenspace can restore people's capacities. Attention restoration theory (ART) (Kaplan, 1995) and stress reduction theory (SRT) (Ulrich et al., 1991) suggest that greenspace can restore people's attention by reducing stress and pressure. Last, greenspace can build people's capacities such as encouraging more physical activity and facilitating social cohesion (Su et al., 2016; Wang et al., 2019a; Yang et al., 2019; Yang et al., 2020).

Previous studies have reported environmental inequities in terms of urban greenery exposure in developed countries (Li et al., 2016; Xu et al., 2018; Zhou and Kim, 2013). Since greenspace can have potential health benefits, environmental disparities for greenspace such as unequal access or exposure to greenspace may result in disproportionate health benefits for different social groups (Jensen et al., 2004; Li et al., 2016). This kind of environmental disparities is also often associated with neighbourhood socioeconomic conditions (Apparicio et al., 2012, 2017; Barbosa et al., 2007; Jensen et al., 2004; Li et al., 2016; Landry and Chakraborty, 2009). Normally, neighbourhoods of high SES (socioeconomic status) often have greater financial resources, cultural and social capital, and political influence to maintain and enhance greenspace (Li et al., 2016; Li et al., 2015a), and potentially have more demand for greenspace quality (Jim and Shan, 2013). As a result, residents in high socioeconomic status neighbourhoods have better access to greenspace both in terms of quantity and quality. For example, Li et al. (2016) found that neighbourhoods in Hartford, Connecticut, USA with higher income have more street greenery than those with lower income. Similarly, in a study of six cities in Illinois, USA. However, Mears et al. (2020) found that although deprived areas in England had better access to greenspace, the greenspace was usually smaller in size, and worse in quality. Due to higher population density, deprived areas were disadvantaged with lower per capita greenspace. Therefore, the association between greenspace exposure and SES is complex and not always consistent. In China, although most of the greenspace is public greenspace and is provided by the government, it still distributes unequally across different neighbourhoods in terms of SES (Sun et al., 2019; You, 2016). First, local government finance usually is associated with neighbourhood socioeconomic conditions (e.g. taxes from property management fees or rents), so neighbourhoods with higher SES are more likely to support its local government to provide sufficient and better greenspace. Also, since greenspace may increase the land value in China, so local governments may follow land-based development process and are keen to provide more greenspace in neighbourhoods with higher SES (Chen and Hu, 2015). Second, neighbourhoods with more greenspace have higher housing price or rent in China, so disadvantaged social groups are less likely to afford the properties there (Xiao et al., 2017a). For example, You (2016) found inequalities in greenspace provision are associated with neighbourhood-level SES in Shenzhen while Shen et al. (2017) pointed out that disparities of greenspace provision exist for neighbourhoods with different levels of SES in Shanghai.

Despite the growing awareness of the importance greenspace quantity for population health, the role of greenspace quality has received less attention (Brindley et al., 2019). Compared with greenspace quantity which is an objective characteristic, greenspace quality reflects more about people's subjective attitudes towards surrounding greenness (Brindley et al., 2019). van Dillen et al. (2012) indicates that quality tends to be a marker for local people's eagerness to use the greenspace and the affordances they gain from this utilization. Furthermore, a high aesthetical value is likely to improve the restorative experience which leads to reduction of stress. Previous epidemiological studies mainly focus on the effect of greenspace availability, access or quantity on health (Gascon et al., 2015; Markevych et al., 2017), so many scholars argued that future research on neighborhood greenspace and health should focus more on its quality rather quantity (Van Dillen et al., 2012). A limited number of studies have compared the health benefit of both quantity and quality of greenspace and found that the quality of greenspace is more relevant to residents' health outcomes (Astell-Burt et al., 2014; Francis et al., 2012; Van Dillen et al., 2012). For example, Francis et al. (2012) found that residents living in neighbourhoods with high quality greenspace had lower odds of psychological distress, but this association was insignificant with greenspace quantity. The reason may be that quality reflects people's perception of the greenspace which directly influences the actual use of greenspace. However, Mears et al. (2020) found that whilst some quantity and quality indicators were not prominently associated with health outcomes, which highlights an urgent need for research including different measures of greenspace exposure. The omission of quality in research is not only due to the vagueness in definition but also due to methodological issues in operationalisation. (Brindley et al., 2019). Existing studies normally used one of two methods to evaluate greenspace quality including questionnaires and SSO (systematic social observation) (de Vries et al., 2013; Feng and Astell-Burt, 2017; Van Dillen et al., 2012). Both methods have obvious limits including being labor-intensive, time-consuming and difficult to apply across a large study area (Lu, 2018).

Given the challenges of collecting information about green space quality at large scale for epidemiological analyses, recently there has been interest in developing new methods for auto-extracting spatial data that provide indicators of quality. Most notably, the recent development of machine learning approaches combined with online mapping data has enabled the automated extraction of sentiments from social media text such as Flickr and Twitter data (Brindley et al., 2019), and ground objects (i.e., trees and grasses) from interactive panoramas such as street view images (Helbich et al., 2019; Labib et al., 2020; Larkin and Hystad, 2019; Li et al., 2018; Lu, 2018; Toikka et al., 2020; Wang et al., 2019a). For example, Brindley et al. (2019) used social media text to extract people's sentiments towards greenspace in order to better capture an indicator of urban greenspace quality. Hence, street view images have already been used for assessing eye-level greenspace quantity (Helbich et al., 2019; Labib et al., 2020; Larkin and Hystad, 2019; Li et al., 2018; Lu, 2018; Toikka et al., 2020; Wang et al., 2019a). For example, Larkin and Hystad (2019) used different exposure measures of visible greenspace and found weak relationships between street view quantity and other greenspace measures. Helbich et al.(2019) used both NDVI (Normalized Difference Vegetation Index) and street view images to assess greenspace quantity and found only greenspace evaluated by street view images is associated with mental health. Street view images have proven to be useful for field observation, since people can evaluate the local environment based on ground objects in street view images (Wang et al., 2019b; Wang et al., 2019c; Yao et al., 2019), an approach known as virtual systematic social observation (Plascak et al., 2020). For example, Ye et al. (2019) used street view data and machine learning methods to assess street quality, while Zhang et al. (2018) applied the similar approach in identifying different urban perceptions. People's perception of greenspace quality is also based on different ground objects (e.g. the absence of trash cans), so besides assessing greenspace quantity, street view images can also be applied for quality assessment. Lu (2019) used Google street view images to assess greenspace quality and used field observation to validate the results. Their findings showed that the results from street view images is highly correlated with the results from field observation and therefore may be a potentially more efficient way for assessing greenspaace quality.

This study addresses some of these research needs and develops a new method to assess greenspace quality based on street view images collected from Guangzhou, China and a machine learning approach. It also focuses on greenspace exposure disparities in terms of urban greenspace quantity and quality that are linked to the neighbourhood socioeconomic status, which enables us to examine whether socioeconomic disadvantaged populations are exposed to poorer quantity and quality of greenspace. This study extends previous research in several respects. First, instead of only focusing on greenspace quantity, it develops a new method to assess greenspace quality based on street view data which is important for epidemiological studies. Second, although previous studies have tried to use street view images to assess greenspace quality, they still evaluate manually based on comparatively few images, which means their methods cannot be applied readily in large scale studies and may lead to bias. Our new approach relies on machine learning approach which can assess millions of images over a large scale. Third, it further compares greenspace exposure disparities in terms of neighbourhood socioeconomic status between quantity and quality.

2. Method

2.1. Study area

Our analysis was conducted in Guangzhou which is located at the Pearl River in mainland China, part of a metropolitan area with a population of more than 13 million people. Our study area was restricted to seven old districts in the Guangzhou city (Liwan, Yuexiu, Haizhu, Tianhe, Baiyun, Panyu and Huangpu District) (Fig. 1) for two reasons. First, new districts (Huadu, Conghua, Zengcheng and Nansha District) are included in Guangzhou only recently due to administrative order, so they are economically and socially separated from seven districts in main urban zone. Second, new districts tend to have substantially lower population/housing density and have fewer built-up areas, and therefore much less street view data were collected there. The focus of this study is on neighbourhood-level (primary administrative unit) and there are 1677 residential neighbourhoods (*juweihui*) in our study area (average neighborhood size= 1 km^2; average population= 5660 persons). A neighbourhood (juwei) usually consists of several gated communities (*xiaoqu*) and non-gated communities (*xiaoqu*), but a community is often too small and residents' daily activity space is not limited by the boundary of community. , so neighborhoods (juwei) should be the better analytical unit. Hence, using community with a small area of greenspace may lead to, underestimation of residents' greenspace exposure.

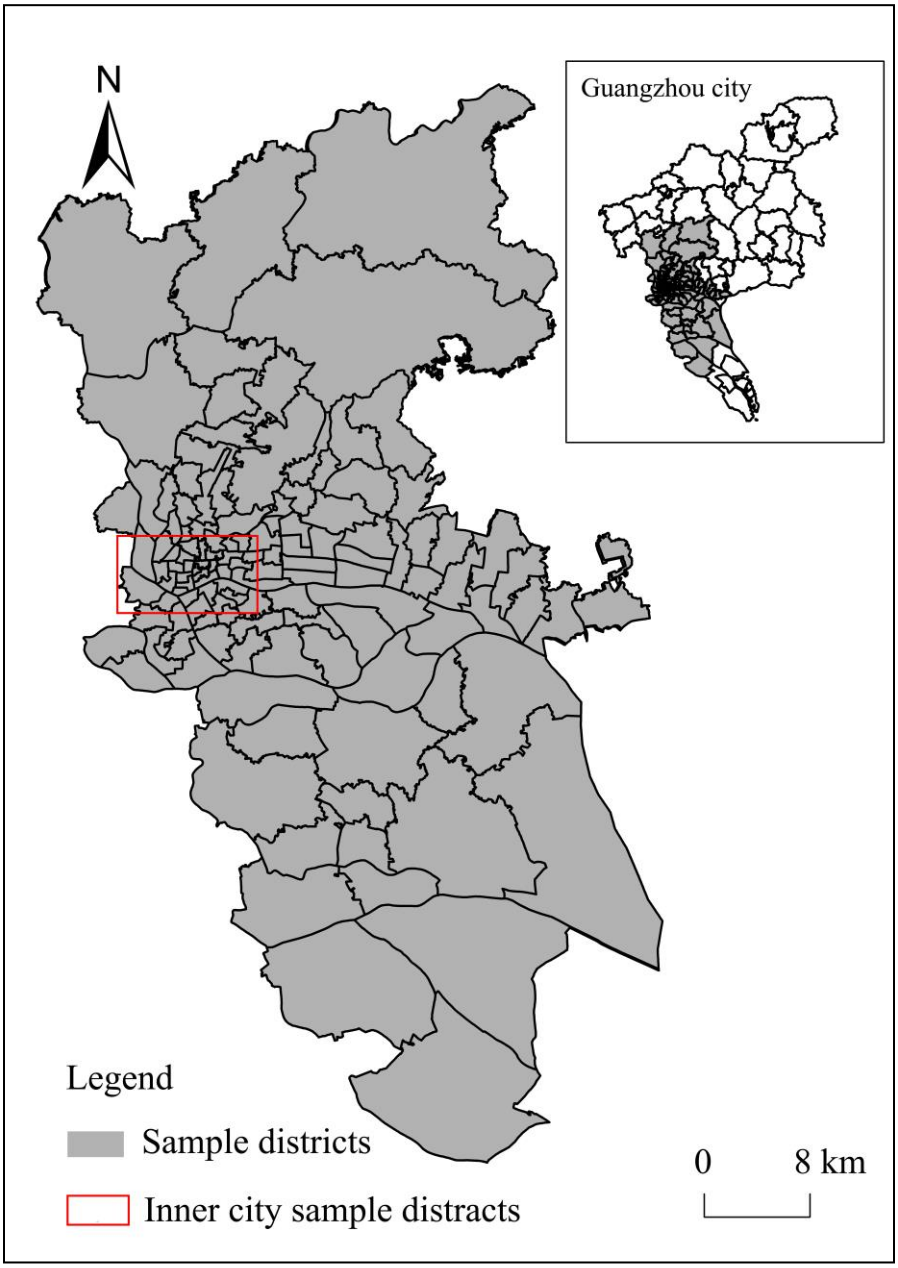


Fig. 1 Locations of the sampled districts in Guangzhou, China

2.2. Data collection

Street view images were collected from Tencent Online Map (downloaded through the Tencent Online Map API), the most comprehensive street view image database in China [https://map.qq.com/] (Helbich et al., 2019). Sampling points were created along each segment of the street network (obtained from OpenStreetMap (Haklay & Weber, 2008)) at 100m intervals following previous studies (Li et al., 2018). Street view images usually have a visual range of more than 50 meters, so 100m intervals can ensure that the street view images cover all the ground objects between any two sampling points. Following previous studies (Helbich et al., 2019; Wang et al., 2019a), we collected four images from four headings (i.e., 0, 90, 180, 270 degree) for each sampling point. In total, we defined 71,286 sampling points which led to 285,144 street view images. A calibration of image brightness was conduct to avoid side effects. In order to obtain a measure of over-head view greenspace, remote-sensing based greenspace was assessed in this study. The data were derived from the Landsat 8 Operational Land Imager and the Thermal Infrared Sensor at a spatial resolution of 30 meters in 2016 (USGS EarthExplorer: https://earthexplorer.usgs.gov/). Last, neighbourhood socioeconomic condition indicators was collected from the sixth census of Guangzhou in 2010, which is part of 2010 China's 10% population sample survey.

2.3. Machine-learning based image segmentation

In order to calculate an eye-level greenspace exposure, following previous studies (Helbich et al., 2019; Wang et al., 2019a), we extracted greenspace objects (e.g., grasses, trees) with a fully convolutional neural network for semantic image segmentation (FCN-8s) (Long et al., 2015) based on the ADE20K dataset of annotated images for training purposes (Zhou et al., 2017; Zhou et al., 2019). The accuracy of the FCN-8s was with 0.814 for the training data and 0.811 for the test data in this study.

2.4. Street view greenspace quantity and quality

Quantity

Following previous studies (Helbich et al., 2019; Wang et al., 2019a), street view greenspace quantity per sampling point was determined as the ratio of the number of greenspace pixels per image summed over the four cardinal directions to the total number of pixels per image summed over the four cardinal directions.

Quality

Fig 2 summarizes the workflow of assessing the street view greenspace quality. First, we constructed our training dataset. Specifically, 2000 images were randomly selected. Then, these images were scored (0 to 10) based on greenspace quality attributes by ten trained investigators who have resided in the research area for at least 3 years. As mentioned in literature review, there are various aspects of greenspace quality, so there are also many different operational items for assessing it. In order to get a robust greenspace quality indicator, we included a wide range of attributes of greenspace quality. The attributes (Cronbach's alpha=0.85) included accessibility (very bad-very good), maintenance (very bad-very good), variation (very monotonous-very varied), naturalness (very unnatural-very natural), colourfulness (very uncolourful-very colourful), clear arrangement (very unsurveyable-very surveyable), shelter (very enclosed-very open), absence of litter (very little trash-very much trash), safety (very unsafe-very safe) and general impression (very negative-very positive) (Lu, 2018; Van Dillen et al., 2012). This provided 10 attributes scores for 2000 images. In the next step, since people evaluate the neighbourhood environment based on ground objects, ground object elements within each street view images can be used to predict residents' perception of the local environment (Wang et al., 2019b; Wang et al., 2019c; Yao et al., 2019). Greenspace quality as one of people's perception of the local environment can also be evaluated through this way. After the image segmentation and the attributes scores of 2000 images, we calculated the proportion of each ground object elements. The random forest model (Breiman, 2001) for automatic rating was trained by fitting the inputted rating scores with the proportion of 151 elements (https://groups.csail.mit.edu/vision/datasets/ADE20K/) in the image segmentations. In this way, 151 ground elements within each image were automatically weighted based on the ten attributes scores. For example, with fewer trash can elements within an image, the absence of litter score would be higher for this image, and the trash can elements are given a weight based on ten attributes scores accordingly. Last, we used this automated scoring system to score all images in study area on these ten attributes.

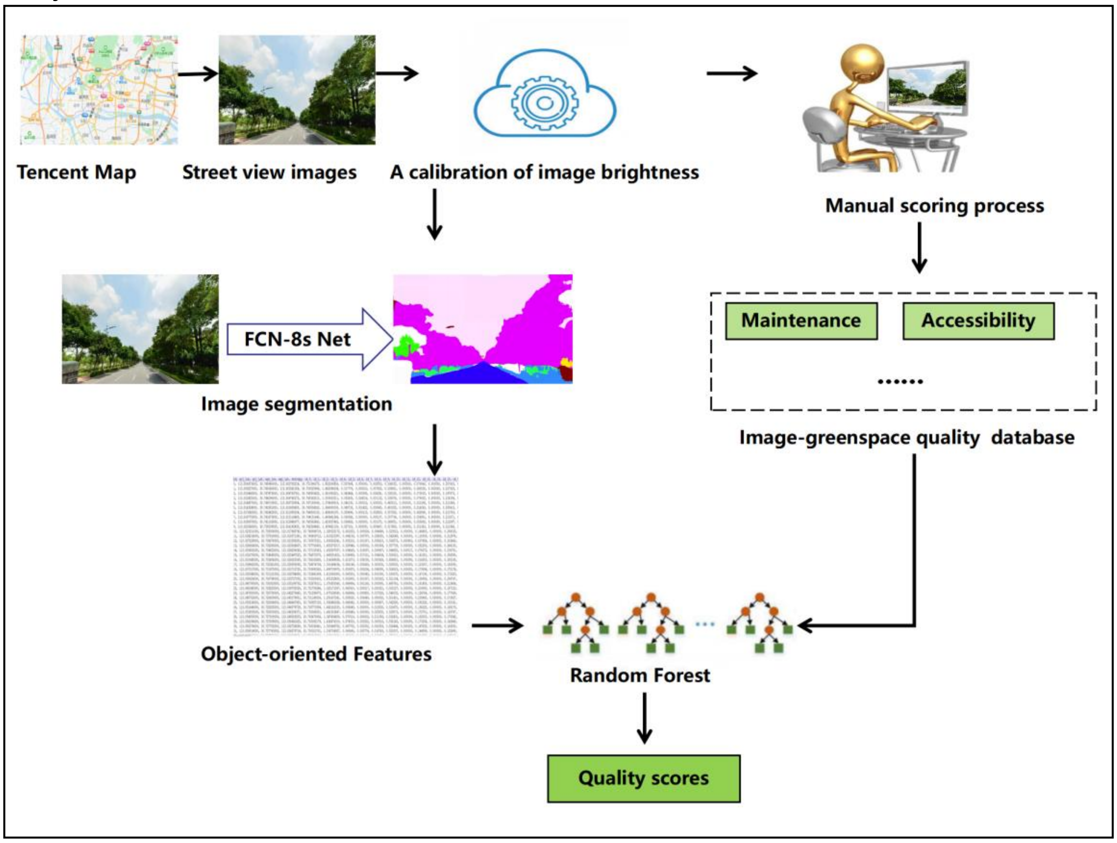


Fig 2 Workflow for assessing greenspace quality.

We took two steps to validate our results. First, one hundred Tencent view images were randomly selected, and attributes of greenspace quality of those images were again assessed manually. The scores from the automated scoring system were highly correlated with attributes of greenness quality of those manually assessed images: accessibility ($r=0.82$, $p<0.05$), maintenance ($r=0.87$, $p<0.05$), variation ($r=0.81$, $p<0.05$), naturalness ($r=0.95$, $p<0.05$), colourfulness ($r=0.92$, $p<0.05$), clear arrangement ($r=0.93$, $p<0.05$), shelter ($r=0.89$, $p<0.05$), absence of litter ($r=0.86$, $p<0.05$), safety ($r=0.98$, $p<0.05$) and general impression ($r=0.91$, $p<0.05$). This process indicates that our scoring system can achieve high accuracy for predicting ten attributes scores outside the training data.

Second, 26 residential neighbourhoods were randomly selected using a multi-stage stratified probability proportionate to population size (PPS) sampling technique and physically visited by three observers and audited with the same audit tool. The results showed reasonable inter-rater reliability (Pearson correlation $r > 0.90$; percentage agreement > 85%). We averaged the score from three observers and calculated the score based on our scoring system for these 26 residential neighbourhood using all the images within the neighbourhood. The correlation between the score from the field audit and scoring system was as follows: accessibility ($r=0.71$, $p<0.05$), maintenance ($r=0.73$, $p<0.05$), variation ($r=0.69$, $p<0.05$), naturalness ($r=0.72$, $p<0.05$), colourfulness ($r=0.68$, $p<0.05$), clear arrangement ($r=0.63$, $p<0.05$), shelter ($r=0.72$, $p<0.05$), absence of litter ($r=0.66$, $p<0.05$), safety ($r=0.88$, $p<0.05$) and general impression ($r=0.81$, $p<0.05$). This process indicates further that the score from our automatic scoring system was correlated with the results from field audit.

The above validation process suggested that our proposed method was suitable for measuring greenspace quality, so we collected scores of attributes of greenspace quality of those images for all images through the proposed automatic scoring system for all sampled neighbourhoods (1677 residential neighbourhoods). Ten attributes for all images achieved excellent internal consistency (Cronbach's alpha=0.88). Following previous studies (Lu, 2018; Van Dillen et al., 2012), the quality of greenspace in each image is the mean value of all 10 attributes. Thus, street view greenspace quality per sampling point was determined as the average greenspace quality score of four image from different cardinal directions. For each neighbourhood, the street view greenspace quality was calculated by the average score of all sampling point within the neighbourhood boundary.

2.5. Remote sensing greenspace quantity

Remote sensing greenspace quantity was assessed by NDVI (Normalized Difference Vegetation Index) (Tucker, 1979). We collected cloud-free images in the greenest season (i.e., June-August) to avoid assessment bias. The value of NDVI was calculated from the following formula: $(NIR - VIS)/(NIR + VIS)$, where NIR stood for reflectance in the near-infrared band and VIS stood for reflectance in the visible region. We aggregated the value of each pixel within the neighbourhood. The value of NDVI varies between -1 and 1and more positive values refer to greener vegetation within the neighbourhood.

2.6. Neighbourhood-level socioeconomic indicators

Following previous studies (Li et al., 2016; Li et al., 2015a; Li et al., 2015b), five socio-economic variables at the neighbourhood-level were selected to represent area-level SES from the 2010 population census data in Guangzhou. These include the proportion of residents with local hukou (registered permanent resident vs registered temporary resident), the proportion of residents with education attainment above high school, unemployment rate and the proportion of residents working in low status occupation, and per capita housing area. Due to the multicollinearity, we used Principal Component Analysis to combine all SES variables into a single indicator. We generated the correlation matrix and then calculated eigenvectors and eigenvalues. After that, we followed Kaiser-Guttman rule and chose one principal component with the largest eigenvalue, which accounted for 86 % of variance explained. The neighbourhood SES index ranged from 0.280 to 9.826. Higher scores mean higher levels of neighbourhood SES. In order to explore the nonlinear relationship between neighbourhood SES and greenness index, we treated neighbourhood SES index as quartile variable.

2.7. Covariates

We also controlled for a series of demographic and built environment variables. First, the proportion of residents aged 0-18, the proportion of residents aged above 65 and the proportion of married residents were controlled. Kabisch and Haase (2014) pointed out that age structure of the neighbourhood may have an influence on greenspace provision. Adolescents and elders have lower mobility than young adults, so they are more likely to benefit more from greenspace within neighbourhood which may be considered in the policy of local greenspace provision (Barbosa et al., 2007). Also, residents aged 0-18 are unemployed and residents aged above 65 are retired, so tend to spend a greater proportion of their day in their local neighbourhood may also influence neighbourhood SES. Second, population density is associated with supply of public facilities and dense neighbourhood may have lower SES since poorer residents are more likely to reside in higher density areas (Liu and Wu, 2006), so we controlled for the average score of this variable for each neighbourhood. Last, the proportion of residents living in houses built before 1979 was included. Old neighbourhoods in China have lower Floor Area Ratio which indicates they have more open space for greenspace (Xiao et al., 2017a). In China, residents living in house built before 1979 usually work in the state sectors which also influence their SES (He et al., 2010; Wu, 2007).

Table 1
Summary statistics for all variables.

Variables (number of neighbourhood=1677)	Mean (Standard Deviation)
Dependent variable	
NDVI	0.105(0.050)
SVG-quantity	0.201(0.089)
SVG-quality	5.560(0.572)
Independent variable	
Neighbourhoods SES quartile (Q1)	0.986(0.210)
Neighbourhoods SES quartile 2 (Q2)	1.416(0.088)
Neighbourhoods SES quartile 3 (Q3)	1.767(0.127)
Neighbourhoods SES quartile 4 (Q4)	2.482(0.577)
Covariates	
Population density (person/km^2)	30704.401(31974.252)
The proportion of residents aged 0-18	0.147(0.046)

The proportion of residents aged above 65	0.082(0.051)
The proportion of residents living in houses built before 1979	0.011(0.014)

2.8. Analysis

2.8.1 Global Moran's I

In order to identify the spatial distribution of greenspace characteristics, we examine the spatial autocorrelation of neighbourhood greenspace quantity and quality . Global Moran's I (Moran, 1950) was used to reflect the overall level of spatial autocorrelation of neighbourhood greenspace quantity and quality.

It was calculated as follow:

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \cdot \sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (1)$$

In the equation, x_i 、 x_j are the level of greenspace quantity or quality in neighbourhood i 、 j ,

w_{ij} is the spatial weight matrix (queen contiguity)， n is the number of neighbourhood in study area. The higher absolute value of Global Moran's I indicates higher level of spatial autocorrelation of neighbourhood greenspace. If Global Moran's I is positive, neighbourhoods with higher level of greenspace cluster with each other and neighbourhoods with lower level of greenspace also cluster with each other. However, if Global Moran's I is negative, neighbourhoods with higher level of greenspace cluster with neighbourhoods with lower level of greenspace while lower level of greenspace cluster with neighbourhoods with higher level of greenspace.

2.8.2 Local Moran's I

Global Moran's I only reflects the overall level of spatial autocorrelation of neighbourhood greenspace, but further we need to identify the spatial autocorrelation level of greenspace for each neighbourhood, so that we can map it and examine the regions where spatial autocorrelation of greenspace is significant for most of the neighbourhoods. We used Local Moran's I (Anselin, 1995) to reflect the spatial relevance of greenspace quantity or quality in each neighbourhood to its neighbors. Local Moran's I reveals the degree of spatial difference and significance between the greenspace quantity or quality of each neighbourhood and its surrounding neighbourhood. It was calculated as follow:

$$I_i = z_i \sum_i W_{ij} z_j \quad (2)$$

In the equation: z_i and z_j are the standardized value of greenspace quantity or quality in neighbourhood i and j; W_{ij} is the spatial weight matrix($\sum_j W_{ij} = 1$). If I_i > 0 and z_i >0, then neighbourhood will be defined as high-high (H-H) region (significant cluster of high values); If I_i < 0 and z_i < 0, then neighbourhood will be defined as low-low (L-L) region (significant cluster of low values); If I_i < 0 but z_i > 0, then neighbourhood will be defined as high-low (H-L) region (significant cluster of outliers in which a high value is surrounded primarily by low values); If I_i > 0 but z_i < 0, then neighbourhood will be defined as low-high (L-H) region (significant cluster of outliers in which a low value is surrounded primarily by high values).

2.8.3 Spatial regression model

In order to link greenspace quantity and quality to neighbourhood socioeconomic conditions spatial regression model was used. It includes spatial lag model (SLM) and spatial error model (SEM) (Cliff and Ord, 1972). If spatial dependence of greenspace exists, OLS (ordinary least squares) models may cause bias. SLM. SLM has nested spatial dependence in dependent variables and the parameter estimation of independent variables while SEM has the parameter estimation of independent variables and error terms, so besides OLS we also adopted SLM and SEM to estimate the relationship between neighbourhood-level socioeconomic and demographic variables on neighbourhood greenspace quantity or quality. It was calculated as follow:

$$y_i = \rho \sum_{j=1}^{n} w_{ij} y_j + \beta x_i + \varepsilon_i \quad (3)$$

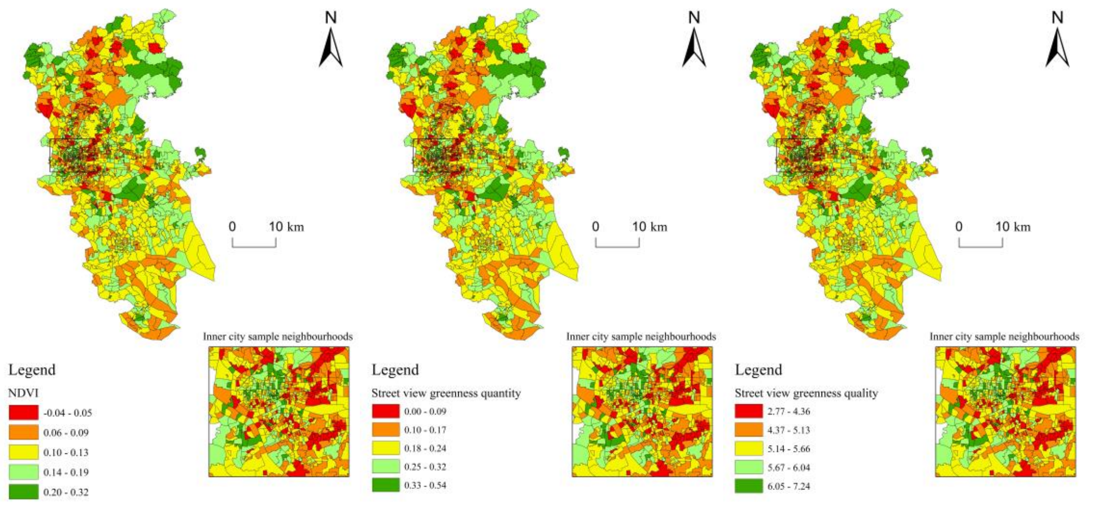
$$y_i = \beta x_i + \delta \sum_{j=1}^{n} w_{ij} \varepsilon_i \quad (4)$$

In the equation: y_i is the level of greenspace quantity or quality in neighbourhood i; y_j is the level of greenspace quantity or quality in neighbourhood j; ρ is the spatial autocorrelation coefficient; w_{ij} is the spatial weight matrix; β is the coefficient of independent variables. x_{it} is the value of independent variables in neighbourhood i; δ is the coefficient of spatial lag explanatory variable, ε_i is the error term. n is the number of neighbourhood in study area.

3. Results

3.1. The distribution of greenspace quantity and quality

Fig. 3 shows the spatial distribution of the aggregated greenspace metrics at the neighbourhood level using NDVI (Fig. 3a), street view greenness quantity (Fig. 3b) and street view greenness quality (Fig. 3c), respectively. Compared with aggregated street view greenness, aggregated NDVI was relatively low in inner-city neighbourhoods. Fig. 4 shows the details of three greenery measurements in study area. In Fig. 4a, the sampled neighbourhoods were low in NDVI, but high in street view greenness. This is likely because neighbourhoods in the inner-city have a large number of street greenery which can be viewed by pedestrians but which are difficult to identify remotely overhead. Also, although street view greenness quantity and quality were both high in inner-city neighbourhoods, there was still a spatial mismatch between them. For example, in Fig. 4b, the sampled neighbourhoods were low in street view greenness quality, but high in street view greenness quantity. In such neighbourhoods, although street greenery was adequate, the green space in the surrounding environment tended to be of low quality. Hence, in Fig. 4c, the sampled neighbourhoods were low in street view greenness quantity, but high in street view greenness quality. Such neighbourhoods are usually in suburb or wealthy downtown area where people can afford villa (larger housing). In such neighbourhoods, street greenery is usually less common but well maintained.



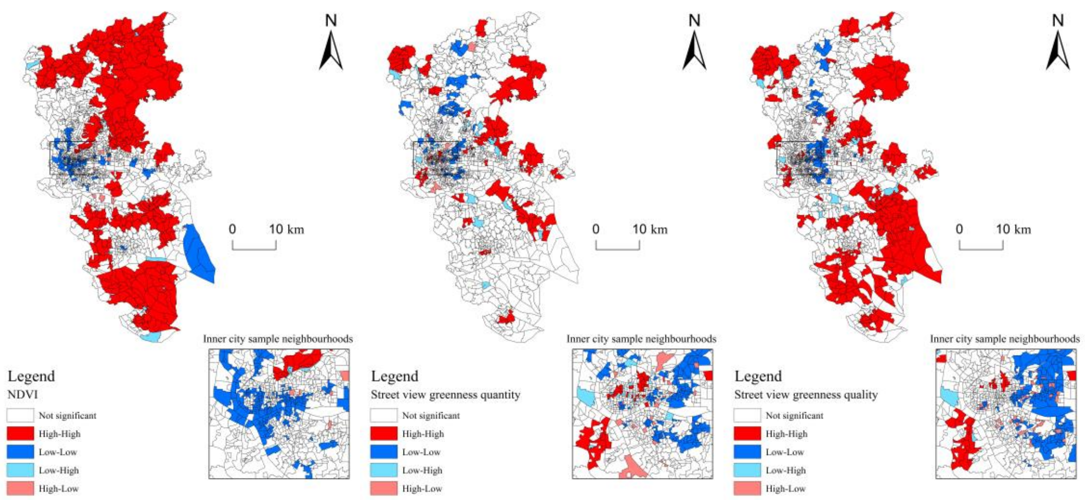
(a) (b) (c)

Fig 3. The distribution of the aggregated green view index values at the neighbourhood level (Natural Breaks): (a)NDVI; (B)Street view greenness quantity; (C)Street view greenness quality

(a) (b) (c)

Fig 4. Comparing three greenery measurements (A) Low level of NDVI and high level of street view greenness; (B) Low level of street view greenness quality and high level of street view greenness quantity; (C) High level of street view greenness quality and low level of street view greenness quantity.

Global Moran's I for distribution of the three aggregated green view index values at the neighbourhood level were all positive and significant at 5% significance level which indicates that distribution of the three aggregated green view index values had positive spatial dependence and space gathering. Figs. 5 displays local Moran's I values in relation to the three types of aggregated green view index values at the neighbourhood level. We only focused on HH and LL clusters, since HL and LH clusters only make up only a small part. Fig. 5a shows that similar to the spatial distribution of the aggregated of NDVI, HH clusters of NDVI were in suburb while LL clusters were in inner-city. Fig. 5b shows that HH, LL clusters of street view greenness quantity all could be observed in inner-city while HH clusters could also be observed in suburb. Fig. 5c shows that similar to street view greenness quantity, HH, LL clusters of street view greenness quality could also be observed in inner-city while HH clusters could be observed in suburb. However, unlike street view greenness quantity, local Moran's I values in relation to the street view greenness quality in inner-city showed spatial characteristic that HH clusters were in the west of inner-city while LL clusters were in the east.



(a) (b) (c)

Fig 5 LISA (Local Indicators of Spatial Association) cluster map of distribution of the aggregated green view index values at the neighbourhood level: (a)NDVI; (B)Street view greenness quantity; (C)Street view greenness quality

3.2. The relationship between three types of aggregated green view index values and neighbourhoods' socioeconomic conditions.

Table 2 shows the Spearman coefficients among three types of aggregated green view index values. The Spearman coefficients of street view greenness quantity versus street view greenness quality and the NDVI showed associations of 0.25 ($p<0.1$) and 0.31 ($p>0.1$). The Spearman coefficients of street view greenness quantity and quality showed associations of 0.78 ($p<0.1$). None of the Spearman coefficients among three types of aggregated green view index values was significant at 5% significance level which suggests they measure different aspects of urban greenness.

Table 2
Results of correlation test for different greenspace measures.

	NDVI	Street view greenness quantity	Street view greenness quality
NDVI	1		
Street view greenness quantity	0.25*	1	
Street view greenness quality	0.31	0.78*	1

*p < 0.10, **p < 0.05, ***p < 0.01.

Table 3, 4 and 5 show the relationship between three types of aggregated green view index values and neighbourhoods' socioeconomic conditions using OLS, SLM and SEM. Compared with OLS, SLM and SEM had higher Adjusted R^2. Also, robust LMLAG values were all significant for three aggregated green view index at 5% significance level while robust LMERR values were not, so we only focused on SLM in Table 3, 4 and 5. SLM in Table 3 showed that neighbourhood SES was not associated with aggregated NDVI values while population density and the proportion of residents aged above 65 were negatively associated with aggregated NDVI values. The significance of Lag Coeff(Rho) indicated that the distribution of aggregated NDVI had positive spatial dependence.

SLM in Table 4 showed that neighbourhood in Q4 SES status had higher aggregated street view greenness quantity values than those in Q1. Also, population density was positively associated with aggregated street view greenness quantity values. The significance of Lag Coeff(Rho) indicated that the distribution of aggregated street view greenness quantity had positive spatial dependence.

SLM in Table 5 showed that neighbourhood in Q2, Q3 and Q4 SES status had higher aggregated street view greenness quality values than those in Q1. Also, population density was positively associated with aggregated street view greenness quality values. The significance of Lag Coeff(Rho) indicated that the distribution of aggregated street view greenness quality had positive spatial dependence.

Table 3

Regression models of NDVI for neighbourhoods in inner-city area, Guangzhou, China.

	OLS		SLM		SEM	
	Coef.(SE)		Coef.(SE)		Coef.(SE)	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Population density		-0.040***(0.002)		-0.017***(0.001)		-0.026***(0.002)
The proportion of residents aged 0-18		0.063*(0.036)		0.024(0.020)		0.039*(0.020)
The proportion of residents aged above 65		-0.076***(0.025)		-0.052***(0.019)		-0.087***(0.021)
The proportion of residents living in house built before 1979		-0.101(0.086)		0.099(0.066)		0.101(0.073)
SES status (referenced: Q1)						
Q2	0.002*(0.001)	0.006*(0.003)	0.002(0.002)	0.003(0.002)	0.002(0.002)	0.002(0.002)
Q3	0.008(0.005)	0.011(0.013)	0.004(0.003)	0.008(0.012)	0.007(0.005)	0.007(0.013)
Q4	0.002(0.003)	0.005(0.003)	0.003(0.003)	0.005(0.003)	0.005(0.003)	0.005(0.003)
Constant	0.099***(0.002)	0.288***(0.009)	0.022***(0.002)	0.119***(0.008)	0.096***(0.003)	0.217***(0.010)
Lag Coeff(Rho)			0.754***(0.016)	0.633***(0.019)		
Lag Coeff(Lambda)					0.757***(0.015)	0.680***(0.019)
R2	0.017	0.363	0.602	0.623	0.602	0.622
AIC	-5297.35	-6011.8	-6508.71	-6700.35	-6508.69	-6660.33
Robust LMLAG			15.327***	75.726***		
Robust LMERR					0.386	11.122***

Coeff. = coefficient; SE = standard error. *p < 0.10, **p < 0.05, ***p < 0.01.

Table 4

Regression models of street view greenness quantity for neighbourhoods in inner-city area, Guangzhou, China.

	OLS		SLM		SEM	
	Coef.(SE)		Coef.(SE)		Coef.(SE)	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Population density		0.017***(0.003)		0.009***(0.003)		0.010**(0.004)
The proportion of residents aged 0-18		0.038(0.057)		0.040(0.051)		0.053(0.052)
The proportion of residents aged above 65		0.046(0.055)		0.034(0.049)		0.032(0.053)
The proportion of residents living in house built before 1979		0.090(0.189)		0.093(0.169)		0.043(0.184)
SES status (referenced: Q1)						
Q2	0.008(0.006)	0.009(0.006)	0.004(0.005)	0.004(0.006)	0.001(0.005)	0.001(0.006)
Q3	0.016(0.016)	0.016(0.011)	0.015(0.015)	0.014(0.016)	0.017(0.017)	0.016(0.017)
Q4	0.027***(0.006)	0.022***(0.008)	0.021***(0.005)	0.018***(0.007)	0.023***(0.006)	0.020***(0.007)
Constant	0.188***(0.004)	0.260***(0.020)	0.093***(0.006)	0.131***(0.019)	0.1905***(0.005)	0.233***(0.022)
Lag Coeff(Rho)			0.484***(0.025)	0.475***(0.026)		
Lag Coeff(Lambda)					0.488***(0.031)	0.479***(0.026)
R2	0.010	0.022	0.217	0.219	0.219	0.221
AIC	-3349.14	-3365.47	-3635.91	-3635.75	-3640.43	-3637.73
Robust LMLAG			3.691**	3.593**		
Robust LMERR					0.044	0.068

Coeff. = coefficient; SE = standard error. *p < 0.10, **p < 0.05, ***p < 0.01.

Table 5

Regression models of street view greenness quality for neighbourhoods in inner-city area, Guangzhou, China.

	OLS Coef.(SE)		SLM Coef.(SE)		SEM Coef.(SE)	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Population density		0.274***(0.020)		0.130***(0.018)		0.176***(0.026)
The proportion of residents aged 0-18		0.592*(0.347)		0.336(0.289)		0.383(0.297)
The proportion of residents aged above 65		0.145(0.331)		0.150(0.276)		-0.037(0.305)
The proportion of residents living in house built before 1979		0.540(1.149)		0.534(0.959)		0.342(1.055)
SES status (referenced: Q1)						
Q2	0.052**(0.024)	0.061**(0.029)	0.026**(0.011)	0.030**(0.012)	0.035**(0.016)	0.031**(0.012)
Q3	0.074**(0.036)	0.056**(0.022)	0.073**(0.031)	0.064***(0.025)	0.101***(0.033)	0.088**(0.037)
Q4	0.226***(0.039)	0.106**(0.047)	0.169***(0.031)	0.116***(0.039)	0.204***(0.035)	0.166***(0.043)
Constant	5.472***(0.027)	6.547***(0.119)	2.111***(0.119)	2.887***(0.177)	5.473***(0.035)	6.227***(0.134)
Lag Coeff(Rho)			0.608***(0.021)	0.558***(0.023)		
Lag Coeff(Lambda)					0.614***(0.021)	0.575***(0.023)
R2	0.019	0.122	0.379	0.388	0.383	0.39
AIC	2861.71	2680.54	2273.17	2225.79	2263.29	2228.21
Robust LMLAG			8.979***	21.663***		
Robust LMERR					0.195	1.065

Coeff. = coefficient; SE = standard error. *p < 0.10, **p < 0.05, ***p < 0.01.

4. Discussion

This study extends previous research on the distribution of different aspects of greenspace and the association between greenspace and neighbourhood socioeconomic conditions in several respects. First, it develops a new method assessing greenspace quality based on street view data and machine learning approach. Compared with traditional measures for assessing greenspace quality, the proposed method is less labor-intensive and more efficient, so it could offer a major step forward in terms of understanding the role of greenspace quality for public health which is important for large scale epidemiological studies. Second, it compares the distribution of greenspace quantity and quality in Guangzhou, China using both street view images (street view greenness) and remote sensing data (NDVI). Results show that there is a difference between the distribution of greenspace quantity and quality. Third, it further investigates the association between greenspace quantity, quality and neighbourhoods' socioeconomic conditions. Compared with NDVI, street view greenness shows greater inequities associated with neighbourhoods' socioeconomic conditions.

4.1. The spatial mismatch between greenspace quantity and quality

We found that NDVI was relatively more abundant in the suburbs, but less abundant in the inner-city. Similar to the spatial distribution of the aggregated of NDVI, LISA cluster map shows that HH clusters of NDVI were in suburb while LL clusters were in inner-city. These findings are consistent with previous studies in the USA (Li et al., 2016; Li et al., 2015a; Li et al., 2015b) and Singapore (Ye et al., 2018). This may be because inner-city areas of Chinese cities tend to be dominated by commercial and residential amenities with few large green facilities. Also, NDVI is more accurate in measuring large green facilities such as park which are more common in the suburbs. As for street view greenness quantity and quality, they are both abundant in the suburb and inner-city. Also, HH, LL, HL and LH clusters of street view greenness all could be observed in inner-city while HH clusters could also be observed in suburb. Ye et al. (2018) and Li et al. (2015, 2016) drew similar conclusions for street view greenness quantity in developed countries. This may be because there is also a demand for greenspace for residents living in inner-city, but since there is little space there, greenspace in inner-city may be more likely to be street plants which can be perceived by pedestrians (Ye et al., 2018). However, street view greenness quality in inner-city shows spatial characteristic that HH clusters were in the west of inner-city while LL clusters were in the east in Guangzhou. A possible explanation is that the west of inner-city in Guangzhou is mainly old town while the east is new town. Residents living in old town in Guangzhou are more likely to own their house while residents living in new town in Guangzhou are more likely to rent (Chen, 2016), and it has been showed by previous studies that owners tend to spend more on maintaining greenspace amenities than renters (Heynen et al. 2006; Perkins et al., 2004).

This study also shows the details of difference among three greenery measurements in inner-city sampled neighbourhoods (Fig. 4). First, inner-city sampled neighbourhoods which are low in

NDVI, but high in street view greenness usually distribute in inner-city which is consistent with the finding in Singapore (Ye et al., 2018). This is due to the high density of street plants and low density of large green facilities in inner-city area. Second, inner-city sampled neighbourhoods which are low in street view greenness quality, but high in street view greenness quantity also usually distribute in inner-city. In such neighbourhoods, street greenery is not well maintained and perceived to be low quality. Third, inner-city sampled neighbourhoods which are low in street view greenness quantity, but high in street view greenness quality usually distribute in both inner-city and suburb. A possible explanation is that these two areas both have villa districts where street plants are not flourish but maintained well. Thus, the spatial mismatch between eye-level and over-head view greenspace as well as between greenspace quantity and quality, indicates that it is important to precisely identify *which* aspect of greenspace should be measured.

4.2. The association between aggregated green view index values and neighbourhood socioeconomic conditions

Consistent with previous studies (Helbich et al., 2019; Wang et al., 2019a), there was no evidence to suggest that NDVI was associated with street view greenness. Nor was there evidence that street view greenness quantity was associated with street view greenness quality. These results indicated that three types of aggregated greenspace index values measure different aspects of greenspace. NDVI measures top-down greenspace, so it does not reflect the cover of street plants (Helbich et al., 2019; Ye et al., 2018). Also, it is a global measurement and covers areas further away from roads, which are less likely to be green. Another explanation is that the resolution of the NDVI data is too coarse to reflect actual top-down greenspace. Therefore, it is important to acknowledge that this study cannot provide definitive evidence that SES disparities are evident, but rather emphasise that NDVI and street view measures identify different components of greenspace. While street view greenness quantity measures human-scale greenspace, it mainly measures eye-level street plants (Helbich et al., 2019; Ye et al., 2018). Compared with greenspace quantity, greenspace quality is more subjective which reflects people’s evaluation of greenspace (Brindley et al., 2019). Although few studies have examined the association between greenspace quantity and greenspace quality, a limited number of studies have found that they may not be significantly correlated with each other, so they may measure different aspects of greenness (Lu, 2018; Van Dillen et al., 2012). Therefore, the above three measures capture slightly different aspects of greenspace and the 'best' measure would depend on the question being posed.

Regression results show that the proportion of residents aged above 65 and the proportion of married residents are negatively associated with the aggregated NDVI index values. However, none of the demographic variables are associated with aggregated street view greenness index values. For built environment variables, only population density is associated with three aggregated greenness index values. Population density is negatively associated with aggregated NDVI index values while it is positively associated with aggregated street view greenness index values. This may be because NDVI is more abundant in the suburbs than the inner-city where population density is higher, while street view greenness is more abundant in the inner-city, but less abundant in suburbs where population density is lower. Most importantly, there was no evidence to suggest SES status has association with the aggregated NDVI index values. However, neighbourhood with higher SES had signifificantly street view greenness quantity values (i.e., 4^{th} quartile) and quality values (i.e., 2^{nd}, 3^{rd} and 4^{th} quartile). These findings indicate that neighbourhoods with higher SES are more likely to have more street view greenspace which is consistent with previous studies (Li et al., 2016; Li et al., 2015a; Li et al., 2015b). This may be because people living in neighbourhood with higher SES are more likely to be able to afford the costs associated with maintaining greenspace (Heynen et al., 2006; Perkins et al., 2004).

In terms of three measurements of greenery, street view greenness quality is more sensitive to the change of SES status (2^{nd}, 3^{rd} and 4^{th} quartile) than quantity (only 4^{th} quartile). On the contrary, NDVI is not associated with SES status. This further indicates the environmental inequity for street view greenness especially street view greenness quality can be explained by neighbourhood SES inequity. A possible explanation is that NDVI can measure large green facilities such as park and their distribution can be planed more equally through urban planning process in China (Xiao et al., 2016; Xiao et al., 2017a). However, street view greenness captures street plants which are more likely to be influenced by local residents who pay the cost of maintaining neighbourhood greenspace (Xiao et al., 2017b). Also, compared with greenspace quantity, environmental inequity for greenspace quality is more associated with neighbourhood SES inequity. This may be because compared with greenspace quantity, maintaining greenspace quality may be more expensive since greenspace quality is associated with surrounding environments (Van Dillen et al., 2012).

4.3. Strengths and limitations

This study has several strengths. First, it is the first attempt to combine street view data with a machine learning approach for assessing greenspace quality in a large scale. Second, we not only compared greenspace quality with quantity, but also compared street view-based greenspace with remote sensing-based greenspace. This allows an examination of greenspace measurement from different perspectives. Third, we also focused on greenspace exposure disparities in terms of urban greenspace quantity and quality in relation to the neighbourhood socioeconomic status. This helps us in further understanding greenspace exposure disparities in the Chinese context.

Our study also had some limitations. First, the boundary of neighbourhood is defined using administrative boundaries and the findings may be affected by the Modifiable Areal Unit Problem (MAUP) (Fotheringham and Wong, 1991). Second, our training sample size was based on 2000 images and it may still be too small. Third, we only focused on greenspace in urban areas in Guangzhou which may not be generalizable to other parts of China, including rural areas. Fourth, aggregated green view indexes were derived in 2016 while neighbourhood socioeconomic conditions variables were derived in 2010, so socioeconomic changes within local areas is not captured. Fifth, since images were collected on a specific date, they may not be sufficiently representative to reflect dynamic changes of greenspace through time. This may be a concern as Guangzhou has a subtropical climate, so most of its vegetation stays green all year round. Still, the static and time specific street view images fail to capture seasonal fluctuation greeneryand may lead to potential bias. Sixth, we only include images from horizontal angle (vertical angle=0), which may not be enough to reflect actual visible greenery. Seventh, local attributes of greenspace quality are usually captured using data collected within urban parks, whereas information collected outside of park boundaries provides indicators that are more closely related to neighbourhood street quality rather than greenspace per se. Therefore, no single measure of greenspace quality fully captures the breadth of greenspace quality; further research could usefully assess the appropriateness of various quality measures in different contexts. Eighth, our approach may not be able to identify all kinds of greenness, since greenness is heterogeneous and there are substantial differences within various kinds of greenness. Ninth, the variation in image quality (i.e. saturation, tint and clarity) may cause potential bias. Last, there are some gated communities in our study area where street view images may not be collected.

4.4. Policy implication and issues for further research

Since greenspace quality may be more relevant to residents' health, policy makers should pay attention to the attributes for greenspace quality such as the accessibility, variation and safety of greenspace instead of only considering greenspace quantity. For example, it may be beneficial to enhance the diversity of street vegetation sin order to promote variation and naturalness. Second, greenspace exposure disparities should also attract more attention, since people living in lower SES areas are less likely to be exposed to greenspace and can only enjoy lower quality greenspaces, so as public facilities, more government investment in greenspace facilities to enhance the quality of these resources should be considered, particularly in ow SES neighbourhoods.. Last, relevant evaluation indicator of greenspace quality should be incorporated into urban planning policies and regulations in China.

Future research should pay attention to the following aspects: 1) There is value in attempting to enlarge the training sample size; online scoring system can be used for collecting people's perceived greenspace quality from across national contexts, so that researchers can further identify whether cultural background and other characters may have impact on assessment of greenspace quality. 2) Previous epidemiological studies mainly focus on the effect of greenspace quantity and future research should also consider greenspace quality using our new methods. 3) Since greenspace quantity and quality may reflect different aspects of greenspace, future research should examine whether the pathways through which greenspace quantity and quality have influence on health vary.

5. Conclusion

This study is the first to develop a new method to assess greenspace quality based on street view data and machine learning method. Further, it focuses on greenspace exposure disparities regarding how urban greenspace quantity and quality are linked to the neighbourhood socioeconomic status in the Chinese context. Results show that the distribution of aggregated NDVI, street view greenness quantity and quality index value shows significant differences that NDVI is relatively more abundant in the suburb, but less abundant in inner-city, while street view greenness is both abundant in the suburb and inner-city. Hence, the correlation analysis shows that no evidence can support three aggregated greenness indexes are significantly associated with each other which indicates that they measure different aspects of greenspace. In terms of three measurements of greenery, environmental inequity for street view greenness is more associated with neighbourhood SES inequity than NDVI. Also, street view greenness quality is more sensitive to the change of neighbourhood SES than quantity. To achieve the goal of promoting urban greening and health through urban planning and design in Chinese urban settings, policymakers and planners are advised to pay more attention to greenspace quality and greenspace exposure disparities in urban area.

Reference

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, *27*(2), 93-115.

Apparicio, P., Landry, S., & Lewnard, J. (2017). Disentangling the effects of urban form and socio-demographic context on street tree cover: A multi-level analysis from Montréal. Landscape and Urban Planning, 157, 422-433.

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.

Astell-Burt, T., Mitchell, R., & Hartig, T. (2014). The association between green space and mental health varies across the lifecourse. A longitudinal study. *J Epidemiol Community Health*, *68*(6), 578-583.

Barbosa, O., Tratalos, J. A., Armsworth, P. R., Davies, R. G., Fuller, R. A., Johnson, P., & Gaston, K. J. (2007). Who benefits from access to green space? A case study from Sheffield, UK. Landscape and Urban Planning, 83(2-3), 187-195.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Brindley, P., Cameron, R. W., Ersoy, E., Jorgensen, A., & Maheswaran, R. (2019). Is more always better? Exploring field survey and social media indicators of quality of urban greenspace, in relation to health. *Urban Forestry & Urban Greening*, *39*, 45-54.

Chen, G. (2016). The heterogeneity of housing-tenure choice in urban China: A case study based in Guangzhou. *Urban Studies*, *53*(5), 957-977.

Chen, W. Y., & Hu, F. Z. Y. (2015). Producing nature for public: Land-based urbanization and provision of public green spaces in China. *Applied Geography*, 58, 32-40.

Cliff, A., & Ord, K. (1972). Testing for Spatial Autocorrelation Among Regression Residuals. *Geographical analysis*, *4*(3), 267-284.

Dadvand, P., Rivas, I., Basagaña, X., Alvarez-Pedrerol, M., Su, J., Pascual, M. D. C., . . . Sunyer, J. (2015). The association between greenness and traffic-related air pollution at schools. *Science Of the Total Environment*, *523*, 59-63.

de Vries, S., Van Dillen, S. M., Groenewegen, P. P., & Spreeuwenberg, P. (2013). Streetscape greenery and health: stress, social cohesion and physical activity as mediators. *Social Science & Medicine*, *94*, 26-33.

Dzhambov, A., Hartig, T., Markevych, I., Tilov, B., & Dimitrova, D. (2018a). Urban residential greenspace and mental health in youth: Different approaches to testing multiple pathways yield different conclusions. *Environmental Research, 160*, 47-59.

Dzhambov, A. M., Markevych, I., Hartig, T., Tilov, B., Arabadzhiev, Z., Stoyanov, D., . . . Dimitrova, D. D. (2018b). Multiple pathways link urban green-and bluespace to mental health in young adults. *Environmental Research, 166*, 223-233.

Feng, X., & Astell-Burt, T. (2017). Residential green space quantity and quality and child well-being: a longitudinal study. *American Journal of Preventive Medicine, 53*(5), 616-624.

Fotheringham, A. S., & Wong, D. W. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment And Planning A, 23*(7), 1025-1044.

Francis, J., Wood, L. J., Knuiman, M., & Giles-Corti, B. (2012). Quality or quantity? Exploring the relationship between Public Open Space attributes and mental health in Perth, Western Australia. *Social Science & Medicine, 74*(10), 1570-1577.

Gascon, M., Triguero-Mas, M., Martínez, D., Dadvand, P., Forns, J., Plasència, A., & Nieuwenhuijsen, M. J. (2015). Mental health benefits of long-term exposure to residential green and blue spaces: a systematic review. *International Journal Of Environmental Research And Public Health, 12*(4), 4354-4379.

Haklay, M., & Weber, P. (2008). Openstreetmap: User-generated street maps. *IEEE Pervasive Computing, 7*(4), 12-18.

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

He, S., Wu, F., Webster, C., & Liu, Y. (2010). Poverty concentration and determinants in China's urban low‐income neighbourhoods and social groups. International journal of urban and regional research, 34(2), 328-349.

Helbich, M., Yao, Y., Liu, Y., Zhang, J., Liu, P., & Wang, R. (2019). Using deep learning to examine street view green and blue spaces and their associations with geriatric depression in Beijing, China. *Environment International, 126*, 107-117.

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.

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).

Jim, C., & Shan, X. (2013). Socioeconomic effect on perception of urban green spaces in Guangzhou, China. *Cities, 31*, 123-131.

Kabisch, N., & Haase, D. (2014). Green justice or just green? Provision of urban green spaces in Berlin, Germany. *Landscape and Urban Planning*, 122, 129-139.

Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology, 15*(3), 169-182.

Labib, S. M., Huck, J. J., & Lindley, S. (2020). Modelling and mapping eye-level greenness visibility exposure using multi-source data at high spatial resolutions. *Science of The Total Environment*, 143050.

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.

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

La Rosa, D., Takatori, C., Shimizu, H., & Privitera, R. (2018). A planning framework to evaluate demands and preferences by different social groups for accessibility to urban greenspaces. *Sustainable cities and society*, 36, 346-362.

Li, X., Santi, P., Courtney, T. K., Verma, S. K., & Ratti, C. (2018). Investigating the association between streetscapes and human walking activities using Google Street View and human trajectory data. *Transactions in GIS*, 22(4), 1029-1044.

Li, X., Zhang, C., Li, W., & Kuzovkina, Y. A. (2016). Environmental inequities in terms of different types of urban greenery in Hartford, Connecticut. *Urban Forestry & Urban Greening, 18*, 163-172.

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

Liu, Y., & Wu, F. (2006). Urban poverty neighbourhoods: Typology and spatial concentration under China’s market transition, a case study of Nanjing. Geoforum, 37(4), 610-626.

Long, J., Shelhamer, E., & Darrell, T. (2015). *Fully convolutional networks for semantic segmentation.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Lu, Y. (2018). Using Google Street View to investigate the association between street greenery and physical activity. *Landscape And Urban Planning*.

Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A. M., . . . Nieuwenhuijsen, M. J. (2017). Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environmental Research*, *158*, 301-317.

Mears, M., Brindley, P., Jorgensen, A., & Maheswaran, R. (2020). Population-level linkages between urban greenspace and health inequality: The case for using multiple indicators of neighbourhood greenspace. *Health & Place*, 102284.

Mears, M., Brindley, P., Maheswaran, R., & Jorgensen, A. (2019). Understanding the socioeconomic equity of publicly accessible greenspace distribution: The example of Sheffield, UK. *Geoforum*, 103, 126-137.

Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika, 37*(1/2), 17-23.

Nieuwenhuijsen, M. J., Khreis, H., Triguero-Mas, M., Gascon, M., & Dadvand, P. (2017). Fifty Shades of Green: Pathway to Healthy Urban Living. *Epidemiology, 28*(1).

Perkins, H. A., Heynen, N., & Wilson, J. (2004). Inequitable access to urban reforestation: the impact of urban political economy on housing tenure and urban forests. *Cities*, *21*(4), 291-299.

Plascak, J. J., Rundle, A. G., Babel, R. A., Llanos, A. A., LaBelle, C. M., Stroup, A. M., & Mooney, S. J. (2020). Drop-And-Spin Virtual Neighborhood Auditing: Assessing Built Environment for Linkage to Health Studies. *American Journal of Preventive Medicine*, 58(1), 152-160.

Shen, Y., Sun, F., & Che, Y. (2017). Public green spaces and human wellbeing: Mapping the spatial inequity and mismatching status of public green space in the Central City of Shanghai. *Urban Forestry & Urban Greening*, 27, 59-68.

Sreetheran, M., Van Den Bosch, C. C. K., 2014. A socio-ecological exploration of fear of crime in urban green spaces – A systematic review. Urban For. Urban Green. 13(1), 1-18.

Su, S., Zhang, Q., Pi, J., Wan, C., & Weng, M. (2016). Public health in linkage to land use: Theoretical framework, empirical evidence, and critical implications for reconnecting health promotion to land use policy. *Land Use Policy*, 57, 605-618.

Sun, C., Lin, T., Zhao, Q., Li, X., Ye, H., Zhang, G., ... & Zhao, Y. (2019). Spatial pattern of urban green spaces in a long-term compact urbanization process—A case study in China. *Ecological indicators*, 96, 111-119.

Toikka, A., Willberg, E., Mäkinen, V., Toivonen, T., & Oksanen, J. (2020). The green view dataset for the capital of Finland, Helsinki. *Data in Brief*, 105601.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment, 8*(2), 127-150.

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments 1. *Journal of Environmental Psychology, 11*(3), 201-230.

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

Wang, R., Helbich, M., Yao, Y., Zhang, J., Liu, P., Yuana, Y., & Liu, Y. (2019a). Urban greenery and mental wellbeing in adults: Cross-sectional mediation analyses on multiple pathways across different greenery measures. *Environmental Research*.

Wang, R., Liu, Y., Lu, Y., Zhang, J., Liu, P., Yao, Y., & Grekousis, G. (2019b). Perceptions of built environment and health outcomes for older Chinese in Beijing: a big data approach with street view images and deep learning technique. *Computers, Environment and Urban Systems*.

Wang, R., Yuan, Y., Liu, Y., Zhang, J., Liu, P., Lu, Y., & Yao, Y. (2019c). Using street view data and machine learning to assess how perception of neighborhood safety influences urban residents' mental health. *Health & Place*.

Wu, F. (2007). The poverty of transition: from industrial district to poor neighbourhood in the city of Nanjing, China. Urban Studies, 44(13), 2673-2694.

Wu, W., Dong, G., Yeran, S. U. N., & Yun, Y. (2020). Contextualized effects of Park access and usage on residential satisfaction: A spatial approach. *Land Use Policy*, 94, 104532.

Xiao, Y., Li, Z., & Webster, C. (2016). Estimating the mediating effect of privately-supplied green space on the relationship between urban public green space and property value: Evidence from Shanghai, China. *Land Use Policy, 54*, 439-447.

Xiao, Y., Lu, Y., Guo, Y., & Yuan, Y. (2017a). Estimating the willingness to pay for green space services in Shanghai: Implications for social equity in urban China. *Urban Forestry & Urban Greening, 26*, 95-103.

Xiao, Y., Wang, Z., Li, Z., & Tang, Z. (2017b). An assessment of urban park access in Shanghai–Implications for the social equity in urban China. *Landscape And Urban*

Planning, *157*, 383-393.

Xu, C., Haase, D., Pribadi, D. O., & Pauleit, S. (2018). Spatial variation of green space equity and its relation with urban dynamics: A case study in the region of Munich. *Ecological Indicators*, 93, 512-523.

Yang, L., Liu, J., Lu, Y., Ao, Y., Guo, Y., Huang, W., ... & Wang, R. (2020). Global and local associations between urban greenery and travel propensity of older adults in Hong Kong. *Sustainable Cities and Society*, 102442.

Yang, Y., He, D., Gou, Z., Wang, R., Liu, Y., & Lu, Y. (2019). Association between street greenery and walking behavior in older adults in Hong Kong. *Sustainable Cities and Society*, 51, 101747.

Yao, Y., Liang, Z., Yuan, Z., Liu, P., Bie, Y., Zhang, J., . . . Guan, Q. (2019). A human-machine adversarial scoring framework for urban perception assessment using street-view images. *International Journal of Geographical Information Science*, 1-22.

Ye, Y., Richards, D., Lu, Y., Song, X., Zhuang, Y., Zeng, W., & Zhong, T. (2018). Measuring daily accessed street greenery: A human-scale approach for informing better urban planning practices. *Landscape And Urban Planning*.

Ye, Y., Zeng, W., Shen, Q., Zhang, X., & Lu, Y. (2019). The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images. *Environment and Planning B: Urban Analytics and City Science*, 46(8), 1439-1457.

You, H. (2016). Characterizing the inequalities in urban public green space provision in Shenzhen, China. *Habitat International*, 56, 176-180.

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning*, 180, 148-160.

Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralba, A. (2017). *Scene parsing through ade20k dataset.* Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Zhou, B., Zhao, H., Puig, X., Xiao, T., Fidler, S., Barriuso, A., & Torralba, A. (2019). Semantic understanding of scenes through the ade20k dataset. *International Journal of Computer Vision*, *127*(3), 302-321.

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.

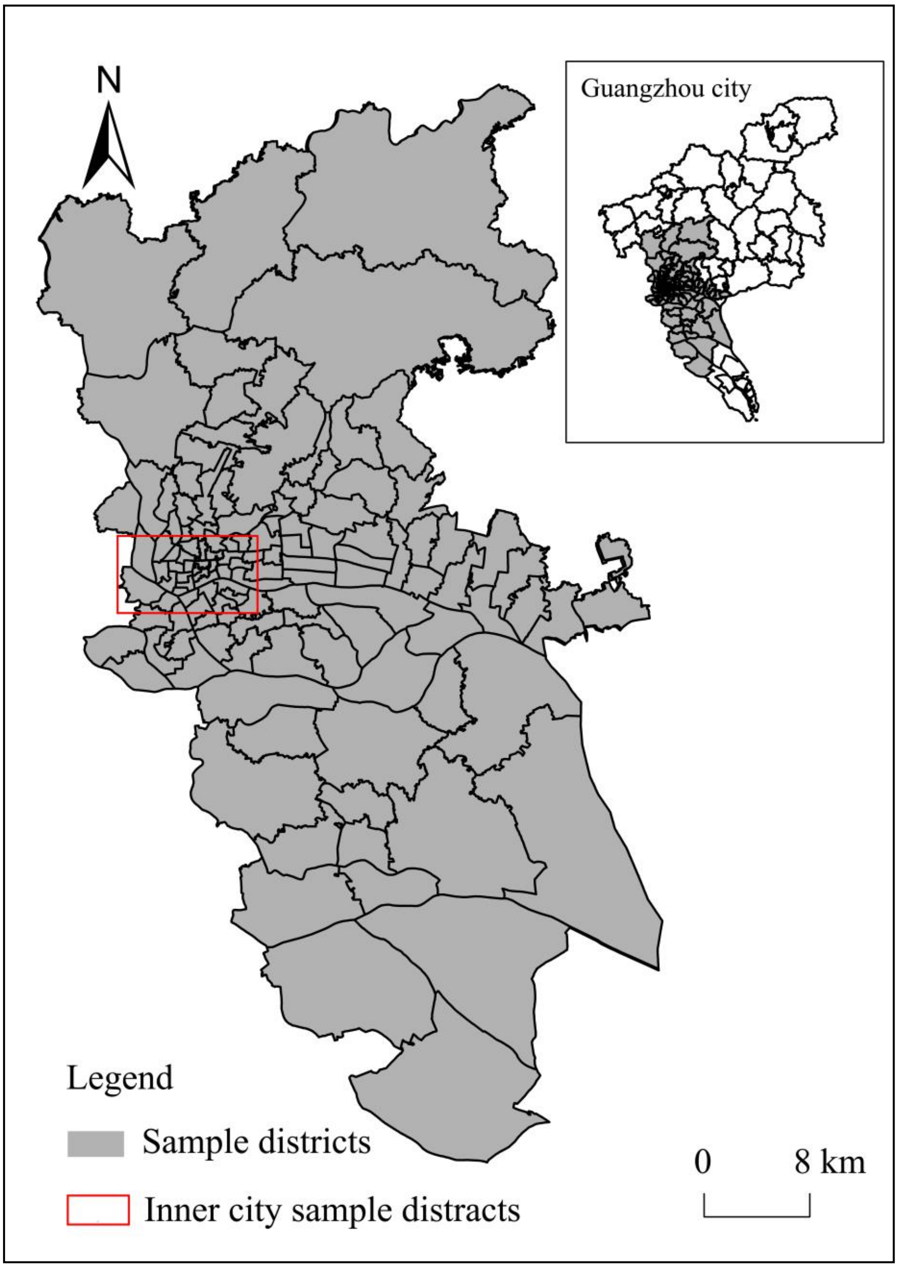


Fig. 1 Locations of the sampled districts in Guangzhou, China

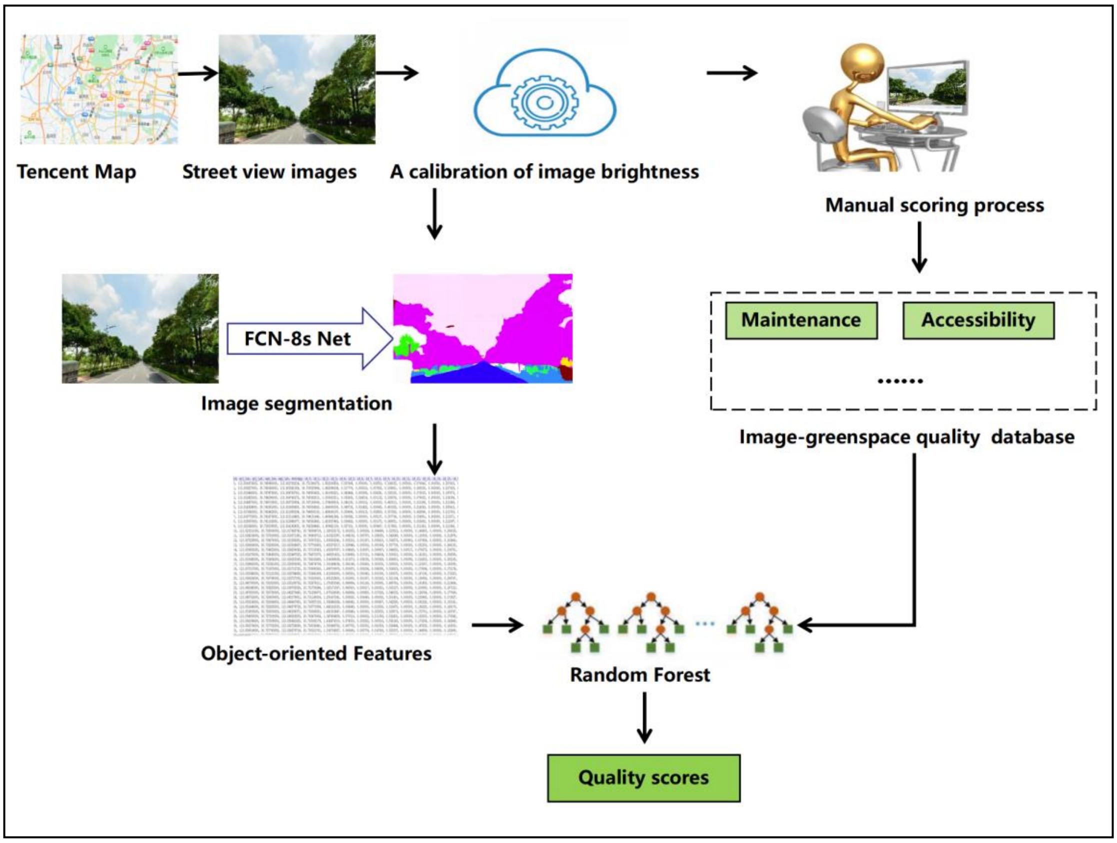
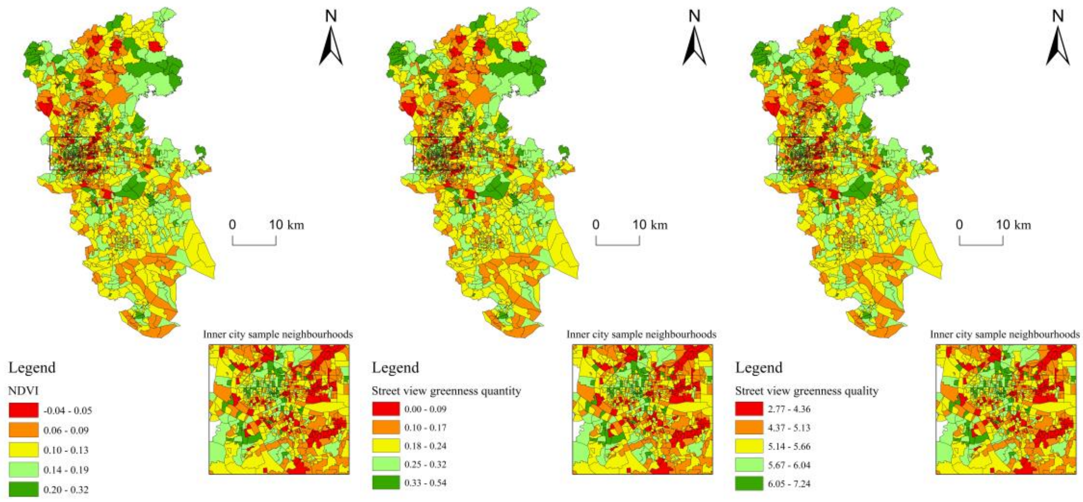


Fig 2 Workflow for assessing greenspace quality.

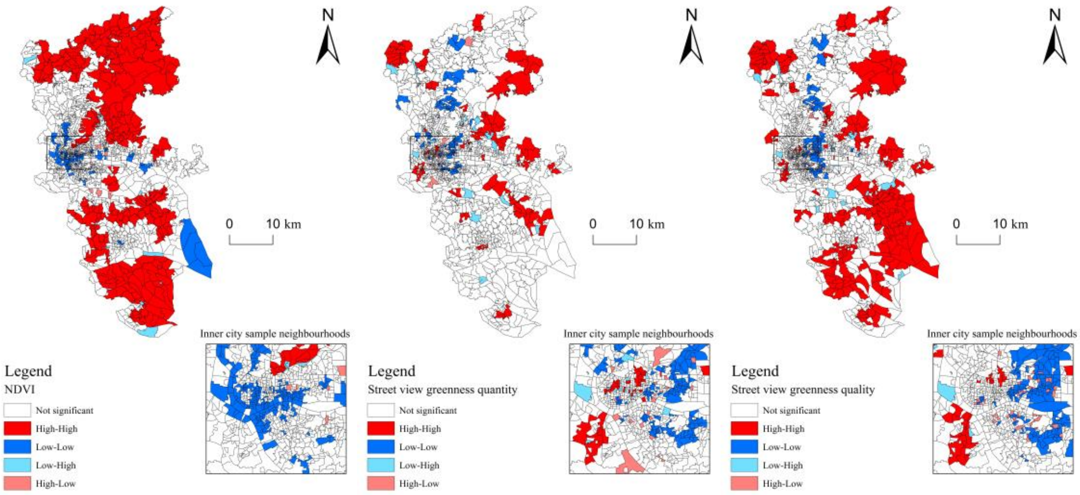


(a) (b) (c)

Fig 3. The distribution of the aggregated green view index values at the neighbourhood level (Natural Breaks): (a)NDVI; (B)Street view greenness quantity; (C)Street view greenness quality

(a) (b) (c)

Fig 4. Comparing three greenery measurements (A) Low level of NDVI and high level of street view greenness; (B) Low level of street view greenness quality and high level of street view greenness quantity; (C) High level of street view greenness quality and low level of street view greenness quantity.



(a) (b) (c)

Fig 5 LISA (Local Indicators of Spatial Association) cluster map of distribution of the aggregated green view index values at the neighbourhood level: (a)NDVI; (B)Street view greenness quantity; (C)Street view greenness quality

Table 1

Summary statistics for all variables.

Variables (number of neighbourhood=1677)	Mean (Standard Deviation)
Dependent variable	
NDVI	0.105(0.050)
SVG-quantity	0.201(0.089)
SVG-quality	5.560(0.572)
Independent variable	
Neighbourhoods SES quartile (Q1)	0.986(0.210)
Neighbourhoods SES quartile 2 (Q2)	1.416(0.088)
Neighbourhoods SES quartile 3 (Q3)	1.767(0.127)
Neighbourhoods SES quartile 4 (Q4)	2.482(0.577)
Covariates	
Population density (person/km^2)	30704.401(31974.252)
The proportion of residents aged 0-18	0.147(0.046)
The proportion of residents aged above 65	0.082(0.051)
The proportion of residents living in houses built before 1979	0.011(0.014)

Table 2

Results of correlation test for different greenspace measures.

	NDVI	Street view greenness quantity	Street view greenness quality
NDVI	1		
Street view greenness quantity	0.25*	1	
Street view greenness quality	0.31	0.78*	1

*p < 0.10, **p < 0.05, ***p < 0.01.

Table 3

Regression models of NDVI for neighbourhoods in inner-city area, Guangzhou, China.

	OLS		SLM		SEM	
	Coef.(SE)		Coef.(SE)		Coef.(SE)	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Population density		-0.040***(0.002)		-0.017***(0.001)		-0.026***(0.002)
The proportion of residents aged 0-18		0.063*(0.036)		0.024(0.020)		0.039*(0.020)
The proportion of residents aged above 65		-0.076***(0.025)		-0.052***(0.019)		-0.087***(0.021)
The proportion of residents living in house built before 1979		-0.101(0.086)		0.099(0.066)		0.101(0.073)
SES status (referenced: Q1)						
Q2	0.002*(0.001)	0.006*(0.003)	0.002(0.002)	0.003(0.002)	0.002(0.002)	0.002(0.002)
Q3	0.008(0.005)	0.011(0.013)	0.004(0.003)	0.008(0.012)	0.007(0.005)	0.007(0.013)
Q4	0.002(0.003)	0.005(0.003)	0.003(0.003)	0.005(0.003)	0.005(0.003)	0.005(0.003)
Constant	0.099***(0.002)	0.288***(0.009)	0.022***(0.002)	0.119***(0.008)	0.096***(0.003)	0.217***(0.010)
Lag Coeff(Rho)			0.754***(0.016)	0.633***(0.019)		
Lag Coeff(Lambda)					0.757***(0.015)	0.680***(0.019)
R2	0.017	0.363	0.602	0.623	0.602	0.622
AIC	-5297.35	-6011.8	-6508.71	-6700.35	-6508.69	-6660.33
Robust LMLAG			15.327***	75.726***		
Robust LMERR					0.386	11.122***

Coeff. = coefficient; SE = standard error. *p < 0.10, **p < 0.05, ***p < 0.01.

Table 4

Regression models of street view greenness quantity for neighbourhoods in inner-city area, Guangzhou, China.

	OLS Coef.(SE)		SLM Coef.(SE)		SEM Coef.(SE)	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Population density		0.017***(0.003)		0.009***(0.003)		0.010**(0.004)
The proportion of residents aged 0-18		0.038(0.057)		0.040(0.051)		0.053(0.052)
The proportion of residents aged above 65		0.046(0.055)		0.034(0.049)		0.032(0.053)
The proportion of residents living in house built before 1979		0.090(0.189)		0.093(0.169)		0.043(0.184)
SES status (referenced: Q1)						
Q2	0.008(0.006)	0.009(0.006)	0.004(0.005)	0.004(0.006)	0.001(0.005)	0.001(0.006)
Q3	0.016(0.016)	0.016(0.011)	0.015(0.015)	0.014(0.016)	0.017(0.017)	0.016(0.017)
Q4	0.027***(0.006)	0.022***(0.008)	0.021***(0.005)	0.018***(0.007)	0.023***(0.006)	0.020***(0.007)
Constant	0.188***(0.004)	0.260***(0.020)	0.093***(0.006)	0.131***(0.019)	0.1905***(0.005)	0.233***(0.022)
Lag Coeff(Rho)			0.484***(0.025)	0.475***(0.026)		
Lag Coeff(Lambda)					0.488***(0.031)	0.479***(0.026)
R2	0.010	0.022	0.217	0.219	0.219	0.221
AIC	-3349.14	-3365.47	-3635.91	-3635.75	-3640.43	-3637.73
Robust LMLAG			3.691**	3.593**		
Robust LMERR					0.044	0.068

Coeff. = coefficient; SE = standard error. *p < 0.10, **p < 0.05, ***p < 0.01.

Table 5

Regression models of street view greenness quality for neighbourhoods in inner-city area, Guangzhou, China.

	OLS Coef.(SE)		SLM Coef.(SE)		SEM Coef.(SE)	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Population density		0.274***(0.020)		0.130***(0.018)		0.176***(0.026)
The proportion of residents aged 0-18		0.592*(0.347)		0.336(0.289)		0.383(0.297)
The proportion of residents aged above 65		0.145(0.331)		0.150(0.276)		-0.037(0.305)
The proportion of residents living in house built before 1979		0.540(1.149)		0.534(0.959)		0.342(1.055)
SES status (referenced: Q1)						
Q2	0.052**(0.024)	0.061**(0.029)	0.026**(0.011)	0.030**(0.012)	0.035**(0.016)	0.031**(0.012)
Q3	0.074**(0.036)	0.056**(0.022)	0.073**(0.031)	0.064***(0.025)	0.101***(0.033)	0.088**(0.037)
Q4	0.226***(0.039)	0.106**(0.047)	0.169***(0.031)	0.116***(0.039)	0.204***(0.035)	0.166***(0.043)
Constant	5.472***(0.027)	6.547***(0.119)	2.111***(0.119)	2.887***(0.177)	5.473***(0.035)	6.227***(0.134)
Lag Coeff(Rho)			0.608***(0.021)	0.558***(0.023)		
Lag Coeff(Lambda)					0.614***(0.021)	0.575***(0.023)
R2	0.019	0.122	0.379	0.388	0.383	0.39
AIC	2861.71	2680.54	2273.17	2225.79	2263.29	2228.21
Robust LMLAG			8.979***	21.663***		
Robust LMERR					0.195	1.065

Coeff. = coefficient; SE = standard error. *p < 0.10, **p < 0.05, ***p < 0.01.

Declaration of interests

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

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: