Social Science & Medicine 379 (2025) 118190

Contents lists available at ScienceDirect

Social Science & Medicine

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

How mobility-based exposure measures may mitigate the underestimation of the association between green space exposures and health

Yang Liu [a,b], Mei-Po Kwan [a,b,*], Liuyi Song [b], Changda Yu [b], Yuhan Cui [a,b]

[a] Department of Geography and Resource Management, the Chinese University of Hong Kong, Hong Kong Special Administrative Region of China
[b] Institute of Space and Earth Information Science, the Chinese University of Hong Kong, Hong Kong Special Administrative Region of China

ARTICLE INFO

Handling Editor: Professor Jamie Pearce

Keywords:
Green space exposure
Mobility
Neighborhood effect averaging
NEAP
Association estimation
Mitigating factor

ABSTRACT

Recent urban green space research highlighted that mobility-based measures of green space exposure may significantly mitigate a particular type of exposure measurement error (contextual errors) of residence-based measures. In this study, we examined an important manifestation of the contextual errors of residence-based measures: neighborhood effect averaging. We analytically illustrated that the contextual errors of residence-based measures may lead to a considerable underestimation of the associations between green space exposures and human health, and the reduction of such underestimation can be quantified through a mitigating factor. We employed data from a cross-sectional survey to assess the usefulness of our analytics. Based on participants' 7-day GPS trajectories, we derived residence-based and mobility-based measures of participants' exposures to green space using a spatiotemporally weighted approach. Logistic regression was employed to estimate the associations between green space exposures and participants' overall health. We derived consistent and significant mitigating factors based on our analytics from the magnitudes of the estimated associations or the variances of green space exposure distributions. Our results indicate that mobility-based measures reduced about 20.9 % – 52.3 % of the underestimation of the associations between green space exposure and health, which reflected the considerable influence of exposure measurement errors. Our study sheds light on how contextual errors may obfuscate the association between green space exposures and human health, which may also be true for other mobility-dependent environmental factors. This has crucial implications for a broad range of environmental and public health studies that need accurate estimation of health impacts.

1. Introduction

To examine how green space may affect human health, past research used various green space exposure measures to assess their association. For instance, Dzhambov et al. (2020) observed that most studies used spatial measures of greenspace (e.g., the normalized difference vegetation index (NDVI), land cover and land use metrics, and distance to green space). Y. Liu et al. (2023c) identified over 70 approaches for green space exposure measures. However, there is no universally accepted measure of individual-level green space exposure due to the different conceptualizations, definitions, and operationalizations of green space in past studies. As observed in past environmental health and health geography studies, there may be different measurement errors and biases when different exposure measurement approaches are used (Gryparis et al., 2009; Hatch and Thomas, 1993; Rhomberg et al., 2011). These errors and biases may stem from the spatial misalignment in the locations of the exposure data and those of the health data (Gryparis et al., 2009) and the spatial and temporal misalignment between exposure levels derived from monitoring station data and individuals' locations (Hatch and Thomas, 1993; Richmond-Bryant and Long, 2020). Since these errors originate from the spatial and temporal contexts of measurement (e.g., different delineations of areal units), they are generally called contextual errors.

In green space research, additional contextual errors may stem from multiple sources and lead to multiple methodological issues. The first is the uncertain geographic context problem (UGCoP), which highlights the misalignment between the true geographic areas of exposure (people may have access to green space around the locations they visit during

* Corresponding author. Department of Geography and Resource Management, the Chinese University of Hong Kong, Hong Kong Special Administrative Region of China.

E-mail addresses: loren-yangliu@link.cuhk.edu.hk (Y. Liu), mpkwan@cuhk.edu.hk (M.-P. Kwan), songliuyi@link.cuhk.edu.hk (L. Song), yuchangda@link.cuhk.edu.hk (C. Yu), YuhanCUI@link.cuhk.edu.hk (Y. Cui).

https://doi.org/10.1016/j.socscimed.2025.118190
Received 5 January 2025; Received in revised form 30 April 2025; Accepted 9 May 2025
Available online 11 May 2025

the day) and the areal units (e.g., residential neighborhoods) used to measure individual exposures (i.e., improper delineations of the geographic areas used to assess exposure) (Kwan, 2012; Perchoux et al., 2014; Spielman and Yoo, 2009; Spielman et al., 2013). Correspondingly, multiple studies highlighted that it may be more proper to employ activity spaces rather than purely neighborhoods for accurate environmental exposure measures (Perchoux et al., 2013). The prompt development of global positioning system (GPS) tracking techniques facilitated the delineation of activity spaces (Chaix, 2018; Chaix et al., 2013). The second methodological issue is algorithmic and representational uncertainties, which suggest that different models and representations of green space may be associated with independent health pathways (Kwan, 2016; Yu et al., 2024). Meanwhile, spatial, temporal, and value-range non-stationarities further complicate the associations between green space exposures and human health outcomes (Bertram and Rehdanz, 2015; Y. Liu et al., 2023d; Liu et al., 2024c). Recently, a new manifestation of contextual errors was proposed, which is referred to as the neighborhood effect averaging problem (NEAP) (Kwan, 2018). The NEAP highlights the tendency that individual-level mobility-based exposures to environmental factors (e.g., green space) would tend towards the mean level of the participants or population of a study area when compared to their residence-based exposures. It occurs because most people move around in their daily lives, and as a result, their mobility-based exposures (i.e., approximations of the true exposures) tend to deviate from their residence-based exposures. For example, when people living in areas with high green space levels travel to areas with lower green space levels for work or leisure, their actual exposure to green space would be lower than estimated based on their residential neighborhoods. Conversely, when people living in areas with lower green space levels travel to areas with higher green space levels, their actual exposure would be higher than estimated. This means that mobility-based exposure measures can better approximate the true exposures to mobility-dependent environmental factors of most people, and measuring people's exposure using mobility-based measures would mitigate contextual errors and lead to a more concentrated distribution around the average exposure level of the population or participants in a study area than residence-based measures. The occurrence of the NEAP has been observed by over a dozen studies for a range of mobility-dependent environmental factors, including air pollution (Dewulf et al., 2016; Kim and Kwan, 2021a, 2021b), infectious diseases (Huang and Kwan, 2022), traffic congestion (Kim and Kwan, 2019), and green space (Gyanwali et al., 2024; Wang et al., 2024; Xu et al., 2023) in different study areas (e.g., Los Angeles, Chicago, Virginia, West Virginia, and Kentucky in the U.S., and Beijing, Xining, and Hong Kong in China).

Previous methodological studies have highlighted the importance of mitigating contextual errors to reduce the risks of Type II errors and wrong conclusions when using statistical frameworks to explore environment-health associations (Liu et al., 2023d; Liu et al., 2024c). This argument emphasizes that the insignificant associations between residence-based exposure measures and human health outcomes may be found to be significant when using mobility-based exposure measures in the same health outcome modeling (Liu and Kwan, 2024; Liu et al., 2023d). Further, as indicated by the manifestation of the NEAP, there may be another totally different improvement in human health outcome modeling. Considering the effect of the NEAP on the variance of the exposure distribution, the more accurate and concentrated exposure distribution using mobility-based measures than using residence-based measures (i.e., smaller variance of green space exposure distribution using mobility-based measures than using residence-based measures) may lead to higher estimated effect sizes. Since mobility-based measures conceptually refer to more accurate total exposure to mobility-dependent environmental factors in people's daily lives, the higher estimated effect size using mobility-based measures means the mitigation of the underestimation using the residence-based measures. More importantly, these arguments on effect sizes also highlight that mobility-based measures may still mitigate the biases and lead to more accurate health association estimations than residence-based measures even though both mobility-based and residence-based measures can yield significant health associations in the same health outcome modeling in some cases.

However, several research gaps still hinder the studies of the NEAP when examining the health impacts of green space. First, the impacts of mobility-based measures on mitigating the underestimation of health effect sizes are not intuitively and straightforwardly understandable, which undermines the discussion of the NEAP and the importance of mobility-based exposure measures. Correspondingly, a rigorous analytical articulation of the mechanism using a statistical framework is still needed. Second, since empirical studies that confirm the mitigation magnitude of the underestimation are quite limited to date, corresponding rigorous experiments are also needed to provide strong evidence for confirming it.

To fill these research gaps and to advance our knowledge on how the NEAP may impact the association estimation between green space exposures and human health outcomes (i.e., the possible health impact of green space), we conduct an empirical study using data collected from a cross-sectional survey in Hong Kong with 940 participants. Our study objectives are to 1) examine whether the NEAP exists in green space exposure assessment using our Hong Kong data set, 2) analytically and empirically articulate how mobility-based measures may mitigate the underestimation of the association between green space exposures and overall health outcomes due to contextual errors and the NEAP, and 3) conduct a sensitivity analysis on this mitigating effect to confirm whether the distribution concentration of measured exposures is transmitted to the estimation of the association between green space exposures and overall health outcomes.

2. Analytical derivation of how mobility-based measures may mitigate the underestimation of the association between green space exposures and health outcomes

In this section, we analytically illustrate how mobility-based exposure measures may reduce the contextual errors of residence-based exposure measures (via deriving and defining a mitigating factor λ) and how reducing these errors may influence the magnitude of the estimated association between an environmental factor and the pertinent health outcome.

To begin, we focus on estimating the magnitude of the association between an environmental factor and the possible health impact through linear ordinary least squares (OLS) regression because it is the baseline technique used in health studies and is straightforward and concise for articulating our arguments (Dismuke and Lindrooth, 2006; Judkins and Porter, 2016). Linear OLS regression generally assumes normal distributions ($N(\mu, \sigma^2)$) of the health outcomes and explanatory variables parameterized with the expectation μ and variance σ^2. Given the dependent health outcome $Y \sim N\left(\mu_y, \sigma_y^2\right)$ and an influencing environmental factor $X \sim N(\mu_x, \sigma_x^2)$, we generally assume the linear association as

$$Y = I + X\beta + \varepsilon \tag{1}$$

where β is the association magnitude between X and Y, I is the constant intercept, and ε is the random error that cannot be explained by the model. Correspondingly, the magnitude of the estimated association $\widehat{\beta}$ and the constant intercept $\widehat{I}$ can be derived using OLS as

$$\widehat{\beta} = \left[(X - \overline{X})^T (X - \overline{X})\right]^{-1} (X - \overline{X})^T (Y - \overline{Y}) \tag{2}$$

$$\widehat{I} = \overline{Y} - \overline{X}\widehat{\beta} \tag{3}$$

where $\overline{X}$ and $\overline{Y}$ are the estimated μ_x and μ_y from observations, respectively. To be more concise in our analytics (Fig. 1A), we remove the

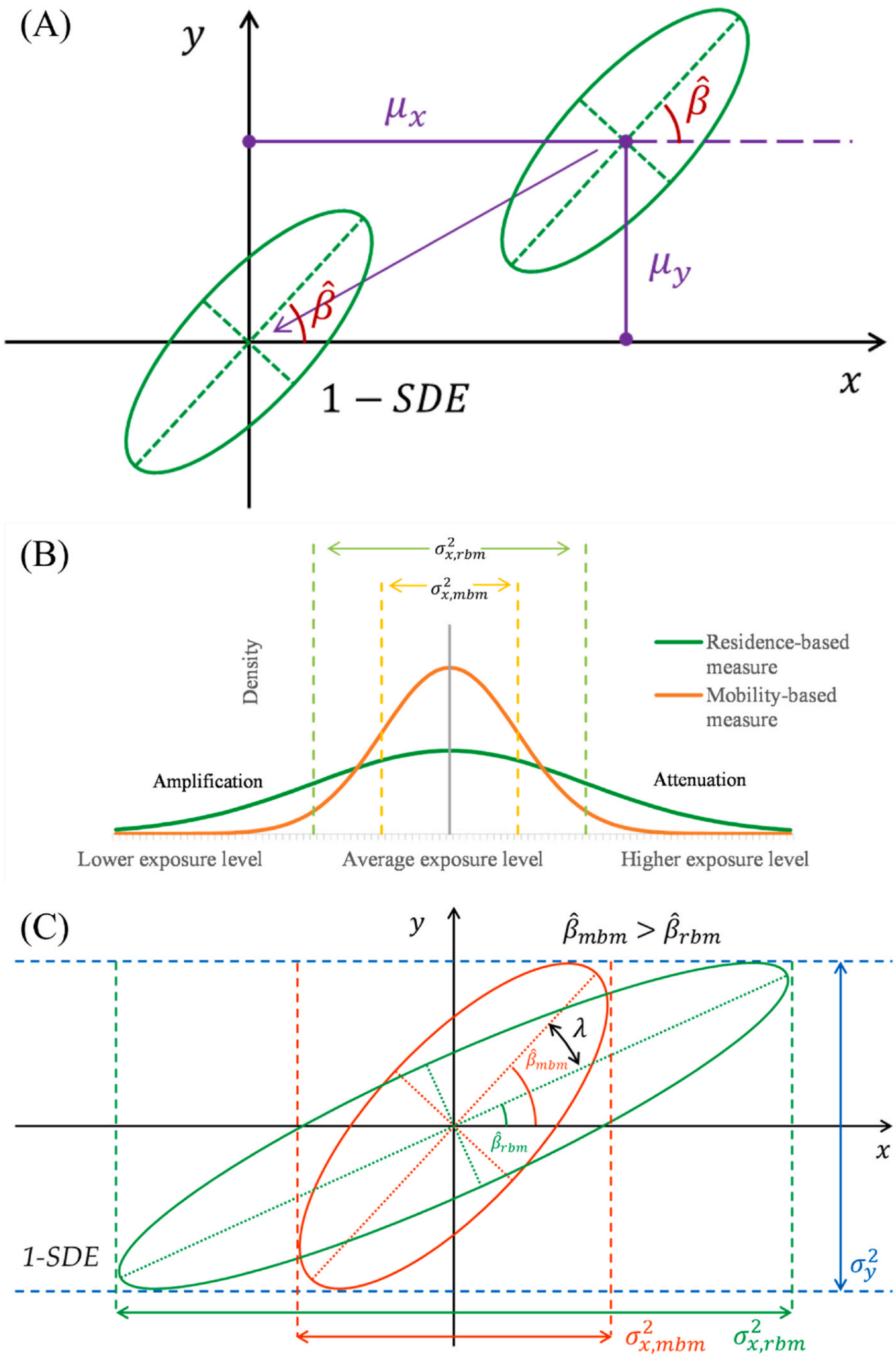

Fig. 1. Graphic illustrations of the analytics in this work: (A) the removal of central tendency; (B) the manifestation of the NEAP using variance σ^2, and (C) the manifestation of the mitigating factor λ. 1-SDE: the one-standard-deviation ellipse of the co-distribution of $[X; Y]$ in an $\mathbb{R}^2$ Cartesian space. The slope of the longer axis of the 1-SDE corresponds to the estimated association magnitude, and the eccentricity of the 1-SDE corresponds to the confidence level of the estimated association magnitude.

central tendency in $[X; Y]$ as

$$X' = X - \mu_x \quad (4)$$

$$Y' = Y - \mu_y \quad (5)$$

such that, $X' \sim N(0, \sigma_x^2)$ and $Y' \sim N\left(0, \sigma_y^2\right)$. In this case, $\widehat{I} \equiv 0$ and Equations (1) and (2) can be rewritten as

$$Y' = X'\beta + \varepsilon \quad (6)$$

$$\widehat{\beta} = \left(X'^T X'\right)^{-1} X'^T Y' \quad (7)$$

As a result, we can remove the constant intercept and purely discuss the magnitude of the estimated association $\widehat{\beta}$ in a similar symmetric form.

Then, let us incorporate the impact of mobility-based measure on the distribution of X. Given the residence-based measure of an environmental factor $X_{rbm} \sim N\left(0, \sigma^2_{x,rbm}\right)$ and the mobility-based measure of the same environmental factor $X_{mbm} \sim N\left(0, \sigma^2_{x,mbm}\right)$, it is apparent to assume

$\sigma^2_{x,rbm} > \sigma^2_{x,mbm}$ according to the definition of the NEAP (Fig. 1B). Here we define a mitigating factor $\lambda \in (1, +\infty)$ as

$$\frac{\sigma^2_{x,rbm}}{\sigma^2_{x,mbm}} = \frac{Var(X_{rbm})}{Var(X_{mbm})} = \frac{X^T_{rbm}X_{rbm}}{X^T_{mbm}X_{mbm}} = \lambda^2 \tag{8}$$

such that

$$\sigma^2_{x,mbm} = \frac{\sigma^2_{x,rbm}}{\lambda^2} \tag{9}$$

and

$$X^T_{rbm}X_{rbm} = \lambda^2 X^T_{mbm}X_{mbm} = X^T_{mbm}\lambda^T\lambda X_{mbm} = (\lambda X_{mbm})^T(\lambda X_{mbm}) \tag{10}$$

i.e., X_{rbm} follows the equivalent normal distribution of λX_{mbm} ($X_{rbm} \sim \lambda X_{mbm}$). The factor λ quantitatively illustrates how mobility-based measures reduce the contextual errors in residence-based measures and how the distribution of the mobility-based measures is closer to the true distribution of exposure measures than the residence-based measures, whereby we can call λ the mitigating factor. Next, we illustrate how λ can be transmitted from the exposure measures to the association estimation through a mitigated distribution.

Now let us turn back to the estimation of association magnitude $\widehat{\beta}$

$$\begin{aligned} \widehat{\beta}_{rbm} &= \left(X^T_{rbm}X_{rbm}\right)^{-1}X^T_{rbm}Y \\ &= \left(X^T_{mbm}\lambda^T\lambda X_{mbm}\right)^{-1}(\lambda X_{mbm})^T Y \\ &= \frac{\lambda}{\lambda^2}\left(X^T_{mbm}X_{mbm}\right)^{-1}X^T_{mbm}Y = \frac{1}{\lambda}\widehat{\beta}_{mbm} \end{aligned} \tag{11}$$

i.e., $\widehat{\beta}_{rbm} = \frac{1}{\lambda}\widehat{\beta}_{mbm}$. Note $\widehat{\beta}_{mbm} > \widehat{\beta}_{rbm}$ since $\lambda > 1$. Thus, as presented above, we analytically illustrated how the mitigating factor λ reduces contextual errors and the variance of exposure levels, which in turn influences the magnitude of the estimated association between green space exposures and health outcomes.

A more straightforward graphic illustration of Equation (11) is shown in Fig. 1C. As observed in past studies, the NEAP can manifest as either upward averaging (amplification) or downward averaging (attenuation), leading to a more concentrated distribution of exposure levels and a smaller variance when using mobility-based measures (when compared to using residence-based measure (i.e., $\sigma^2_{x,rbm} > \sigma^2_{x,mbm}$)). Consequentially, we can analytically illustrate that the magnitude of the association estimated with mobility-based measures would be larger than that obtained with residence-based measures (Equation (11)). Since residence-based exposure measures contain considerable contextual errors that are reduced by mobility-based measures, *the higher estimated association obtained with mobility-based measures indicates how much residence-based measures underestimate the association.* However, no empirical study has evaluated how much residence-based measures may underestimate such association and how much mobility-based measures may reduce contextual errors.

In the following sections, we employed a cross-sectional dataset with 940 participants collected in Hong Kong to empirically examine these issues. Participants' residence-based and mobility-based exposures to green space were derived using GPS trajectories, fine-grained remote sensing data, and spatiotemporally weighted approaches. Logistic regression was used to derive the mitigating factor λ and assess its influence on the association between participants' green space exposures and self-reported health. Finally, a bootstrapping approach and multiple t-tests were used to analyze the sensitivity and significance of the derived λ. In the following subsections, we describe our data and methods in detail.

3. Methods

3.1. Study area

Our cross-sectional survey was conducted in Hong Kong, which is one of the world's most urbanized and populated megacities. Hong Kong had a population of about 7.5 million by the end of 2023 and all residents are fully urbanized (Department, 2023). Its green space coverage is about 75 % of its land area, while the downtown areas have green space coverage ranging from 14.45 % to 44.24 % (Tian et al., 2011). The vastly uneven distribution of green space in Hong Kong indicates a high likelihood of considerable differences between people's mobility-based and residence-based exposures to green space in their everyday lives. Thus, the study area is excellent for examining the extent to which residence-based measures may underestimate the association between green space exposure and health and mobility-based measures may reduce contextual errors.

There are two types of residential communities in Hong Kong: old towns developed before the 1950s, and new towns developed later (Y. Liu et al., 2023b). Compared to old towns, new towns in Hong Kong were developed with better plans to meet the needs of the booming population, such as much taller buildings, wider roads, and more public open spaces (Y. Liu et al., 2023b). We chose three old towns and three new towns for our survey (Fig. 2). The selected new towns are Tin Shui Wei (TSW), Sha Tin (ST), and Kwai Tsing (KTS). The selected old towns are Sham Shui Po (SSP), Central and Western (CW), and Kwun Tong (KTO). KTO was a typical old town and was recently renovated. However, the renovation did not significantly alter the green space settings in KTO, and we still consider KTO as an old town in our survey. The geographic settings of each community are summarized in Table 1.

3.2. Data collection

In this study, green space is defined as any land patch covered by vegetation (Y. Liu et al., 2023c). It was delineated using PlanetScope multispectral remote sensing images at 3 m spatial resolution (Team, 2017). The normalized difference vegetation index (NDVI) and a thresholding approach were employed to identify green space and non-green space after radiometric calibration and atmospheric correction (Y. Liu et al., 2023d). The accuracy and validity of the delineated green space have been carefully evaluated in several studies (Kan et al., 2023; D. Liu et al., 2023a; Y. Liu et al., 2023d).

The remote sensing images used in this study were captured in January 2021. Although there is some mismatch between the timing of these images and our survey data (which were collected from March 2021 to April 2023), it would not affect the results much because Hong Kong has a humid subtropical climate and the green space in Hong Kong is ever-green and remains largely unchanged throughout the year. In addition, Hong Kong is a highly urbanized city without rapid or massive land cover changes over time. We therefore did not consider seasonal changes and landcover changes in our study (Liu et al., 2024b).

In the survey, participants in each selected community were recruited through stratified sampling, which aimed to generate representative samples of the respective communities based on their socio-demographic compositions (age, gender, monthly household income, education level, and marital status). The survey protocol was approved by the *Survey and Behavioral Research Ethics (SBRE) Committee* of the authors' university. Written informed consent was obtained from each participant before data were collected from them. As Table 2 indicates, our sample has more female participants than male participants due to the slightly lower response rate of male residents. Overall, the socio-demographic profiles of the respective subsamples of each selected community closely represent the respective populations. Our participants cover diverse socio-demographic groups along multiple socio-demographic axes (Table 2). About 74 % of participants are employed and need to commute to work. However, other participants may also

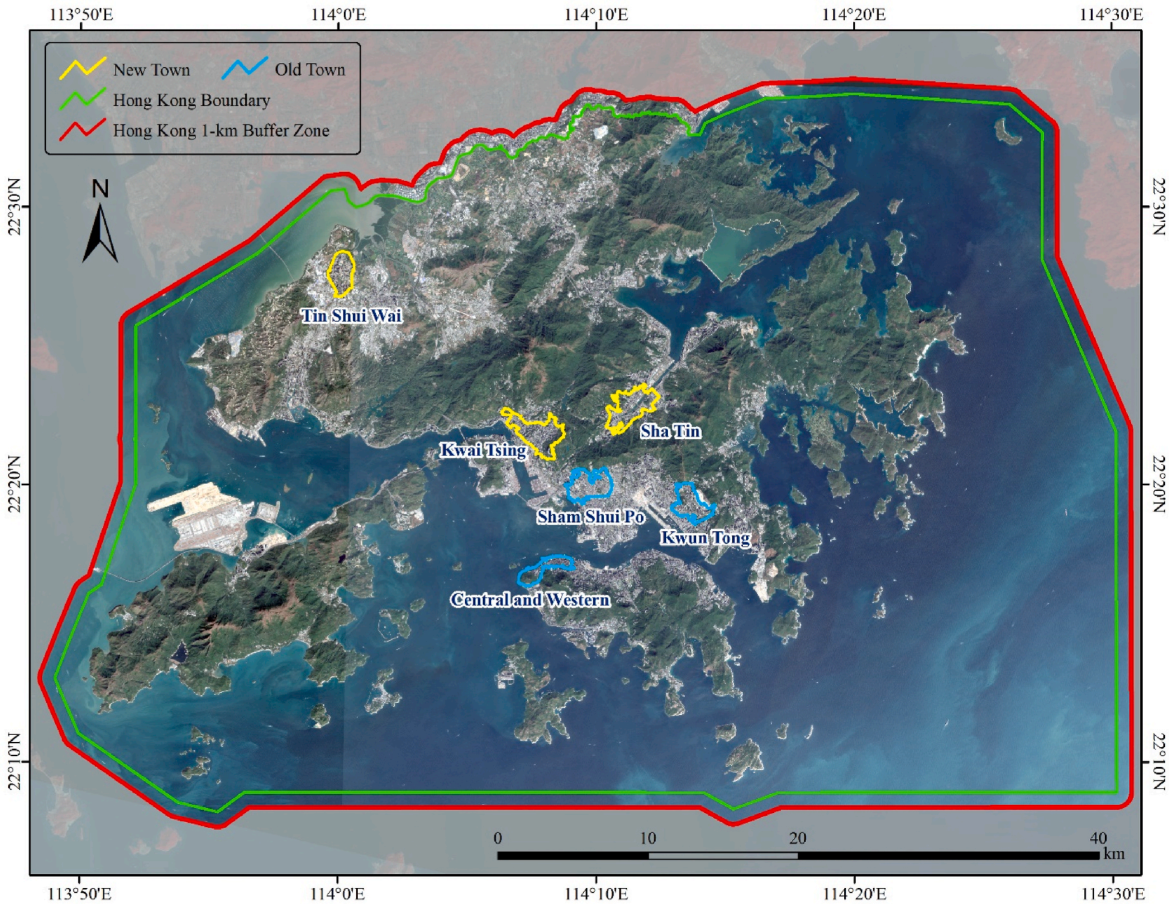

Fig. 2. The representative communities in our surveys.

Table 1
The geographic settings and sample sizes of the representative communities in our surveys.

Town type	Name	Area	Population	Sample size
Old towns	Sham Shui Po (SSP)	5.35 km^2	~300 k	104
	Central and Western (CW)	2.83 km^2	~175 k	89
	Kwun Tong (KTO, renovated)	3.93 km^2	~446 k	182
New towns	Tin Shui Wei (TSW)	4.32 km^2	~300 k	104
	Sha Tin (ST)	5.64 km^2	~315 k	289
	Kwai Tsing (KTS)	5.84 km^2	~441 k	172

have considerable mobility due to activities such as shopping and leisure activities. About 82 % of participants have an activity space size larger than 1 km^2 during weekday daytime, and 63 % during weekend daytime.

As part of a larger project, the survey data were collected using an integrated individual environmental exposure assessment system (IEEAS) developed by Wang et al. (2021) and implemented in subsequent studies (Kou et al., 2020; Ma et al., 2021). The system collected real-time exposure data from each participant using GPS, mobile sensors, activity-travel diaries, health questionnaires, and geographic ecological momentary assessment (GEMA). During the surveys days, each participant carried a GPS-equipped smartphone, which recorded

Table 2
The socio-demographic profiles and self-reported overall health statuses in either the old towns or the new towns in our surveys.

		New towns (TSW, ST, and KTS)		Old and renovated towns (SSP, CW, and KTO)	
Gender	Male	187	33.1 %	126	33.6 %
	Female	378	66.9 %	249	66.4 %
Age	18–24	84	14.9 %	49	13.1 %
	25–44	321	56.8 %	225	60.0 %
	45–64	160	28.3 %	101	26.9 %
Monthly household income [a]	Low	149	26.4 %	96	25.6 %
	Middle	232	41.0 %	137	36.5 %
	High	184	32.6 %	142	37.9 %
Education level [b]	Low	140	24.8 %	78	20.8 %
	Middle	308	54.5 %	222	59.2 %
	High	117	20.7 %	75	20.0 %
Marital status [c]	Single	311	55.0 %	220	58.7 %
	Married	208	36.8 %	128	34.1 %
	Others	46	8.2 %	27	7.2 %
Overall health status	Good	468	82.8 %	325	86.7 %
	Bad	97	17.2 %	50	13.3 %
Total		565	100.0 %	375	100.0 %

[a] Monthly household income: the low-income group has an income of less than 20,000 Hong Kong dollars (HKD), the middle-income group has an income of 20,000–39,999 HKD, and the high-income group has an income of 40,000 HKD or above.

[b] Education level: the low group graduated from middle school or lower, the middle group is with a bachelor's degree or certification, and the high group is with a master's degree or higher.

[c] Other marital statuses include those divorced and widowed.

the location of each of their visited places at a 1-s resolution using three independent APPs, including AirCasting, GPS Logger, and Six Foot (Wang et al., 2025). AirCasting is an APP for monitoring air pollution that also collects GPS data. The time participants spent at different locations (e.g., home and other out-of-home locations) was derived from the GPS data. The GPS data collected were aligned spatially and temporally by cross-checking the GPS, activity-travel diary, and sensor data (only the GPS location sequences of the sensor data were employed in this study). Participants' GPS trajectories were recorded consecutively for 7 days. These trajectories originally had a temporal resolution of 1 s. They were then filtered using a spatiotemporal clustering approach (Jiang et al., 2017; Li et al., 2008; Martin et al., 2023) and resampled to 1-min temporal resolution to reduce measurement uncertainties (Wang et al., 2025). After excluding surveys with missing or incomplete data (e.g., incomplete GPS trajectories with completeness lower than 98 %), we finally have valid data from 940 participants (Table 2).

As part of a larger project, the health questionnaires and activity-travel diaries collected comprehensive data on participants' socio-demographic attributes, physical and mental health statuses, physical activity, sleep quality, neighborhood environments, and so on. In this study, participants' self-reported overall health status is used as an indicator of their current health condition. It was recorded on a 6-point scale in response to the question: "Overall, your health condition is?" where the answers were "excellent" (1) to "terrible" (6). Due to the few responses of bad health statuses, the scores were re-coded into two categories: good health (excellent, very good, and good health) and bad health (bad, very bad, and terrible).

3.3. Green space exposure measures

In this study, we derive a participant's green space exposure as the green space area coverage ratio in a buffer zone around either the person's home location (i.e., residence-based measures) or visited locations (i.e., mobility-based measures). The residence-based measure of green space exposure E_{rbm}^{GS} is defined as

$$E_{rbm}^{GS} = \frac{A(buf \cap GS)}{A(buf)} \tag{12}$$

where buf is a buffer zone around a participant's home location, GS is the green space inventory, and function A returns the size of the area, either the green space within the buffer zone or the buffer zone itself. The mobility-based measure of green space exposure E_{mbm}^{GS} can be defined in a similar format through a spatiotemporally weighted approach:

$$E_{mbm}^{GS} = \sum \frac{A(buf_i \cap GS)}{A(buf_i)} \cdot \frac{t_{i+1} - t_i}{D} \tag{13}$$

where buf_i is the buffer zone around the i-th visited location along a participant's activity-travel trajectory, t_i is the timestamp of the i-th visited location and t_{i+1} is the next timestamp, and D is the duration of the entire survey period (i.e., 7 days) for a participant. As indicated by Equation (13), the green space exposure at different places is disproportionally weighted through visit durations, and the green space exposure at a long-staying place (e.g., home or office) is weighted more importantly than a quickly passing-by place (e.g., a park on the way home).

Both E_{rbm}^{GS} and E_{mbm}^{GS} are unitless scalars ranging from 0 to 1. A larger value of E_{rbm}^{GS} means more exposure to green space at a participant's home location, and a larger value of E_{mbm}^{GS} means more exposure to green space along a participant's activity-travel trajectories. The buffer radii for both measures are set to 300 m. Note that a recent paper has put forward an analytical framework for identifying the best buffer size for examining the health impacts of green space for a specific study area (Liu et al., 2024a). The "optimal" buffer size was identified by minimizing contextual errors and evaluating the statistical significance of the association between green space and human health. Buffer zones from 1.5 m to 2500 m were tested and the results indicate that the most causally relevant green space exposures are derived using buffer zones from 100 to 500 m and 300 m is the best. Such configuration was also used in several previous studies (Y. Liu et al., 2023d; Liu et al., 2024c; Zheng et al., 2024). ArcGIS Pro and R were employed for deriving the exposure measures.

3.4. Statistical analyses

We conducted three-phase statistical analyses to examine how mobility-based exposure measures may reduce the underestimation of the association between green space and health by residence-based exposure measures using IBM SPSS Statistics. In the first phase, we used frequency plots and multiple F-tests to confirm the occurrence of neighborhood effect averaging in the selected communities. In the second phase, we employed logistic regression based on maximum-likelihood estimation to estimate the association between green space exposures and participants' overall health. Logistic regression is more suitable for handling our binary health outcome (Pohlman and Leitner, 2003). Further, it does not assume the normal distribution of the exposure measures and then loosens the requirements for the explanatory variables. In this study, the model is designed as

$$\log\left(\frac{P(good\ health)}{1 - P(good\ health)}\right) = I + B_m^{GS} E_m^{GS} + \sum B_j Cov_j + \sum B_k Com_k + \varepsilon \tag{14}$$

where $P(good\ health)$ is the probability of a participant being in good health, I is the intercept, E_m^{GS} is the green space exposure and B_m^{GS} is the corresponding association between green space exposure and participants' overall health. m stands for either residence-based or mobility-based measures. ε is the random error that cannot be explained by our model. To control for the potential effects of participants' socio-demographic attributes and the geographic contexts of different communities (i.e., new towns versus old towns), we also employed a group of covariates and dummy variables: Cov_j stands for a group of covariates to control for the influence of participants' socio-demographic attributes, including gender, age, household income level, education level, and marital status; Com_k stands for a group of dummy variables to control for the influence of geographic contexts (see more details in Table 2). The coefficients of these controlled covariates and dummy variables are B_j and B_k, respectively. A preliminary analysis using multilevel binary logistic regression indicates that the variability of either the intercept or the coefficient of green space exposure is not significant between different communities in either the old towns or the new towns. Thus, it is appropriate to employ a single-level but more concise binary logistic model.

In the last phase, we tested the sensitivity of the derived mitigating factor λ. According to Equations (8) and (11), we can derive λ from either the estimated coefficients of green space exposures or from the variances of green space exposure distributions as below

$$\lambda_b = \frac{B_{mbm}^{GS}}{B_{rbm}^{GS}} \tag{15}$$

$$\lambda_v = \sqrt{\frac{Var(E_{rbm}^{GS})}{Var(E_{mbm}^{GS})}} \tag{16}$$

where λ_b is the mitigating factor derived from the estimated coefficients of green space exposures, and λ_v is the mitigating factor derived from the variances of green space exposure measures. Function Var returns the variance of a group of observations. Equations (8)–(11) indicate that the mitigation (i.e., reduction) of the underestimation (indicated by λ_b)

originates from the concentration in the variance of the exposure measures (indicated by λ_v). Correspondingly, λ_b and λ_v should be statistically consistent with each other in practice. Note that there is currently no available method for evaluating the uncertainties in the derived λ and thus no analytical approach is available for testing the consistency between λ_b and λ_v. Instead, we employed a bootstrapping approach to simulate the distributions of the derived λ and evaluated the standard errors of the expectation of λ (Beasley and Rodgers, 2012; Haukoos and Lewis, 2005). We randomly chose two-thirds of the participants from our total sample, repeated the logistic regression, and derived λ_b and λ_v from the subset. We repeated these steps 500 times and derived 500 pairs of λ_b and λ_v, respectively. Finally, multiple t-tests were employed to test the differences between 1) λ_b and λ_v derived from the subsets; 2) λ_b or λ_v derived from the subsets and the null hypothesis, i.e., $\lambda = 1$; and 3) λ_b or λ_v derived from the subsets and the corresponding λ derived from the full set.

4. Results

4.1. *Confirmation of the occurrence of the NEAP*

We derived the residence-based and mobility-based measures of green space exposures for the 940 participants. To confirm the occurrence of the NEAP, we plotted the frequency of green space exposure levels in the new towns and the old towns. For the new towns, we observed apparent NEAP (Fig. 3A). For the participants in the new towns, clear upward averaging can be observed using mobility-based measures at the lower end of the plot and clear downward averaging can be observed at the higher end. In contrast, the frequency of green space exposure levels near the average level of the entire sample is higher using mobility-based measures than residence-based measures. These are typical manifestations of the NEAP. We thus conclude that for the new towns, using mobility-based exposure measures mitigates the contextual errors of residence-based exposure measures (Kwan, 2012).

However, for the old towns, the differences in the exposure-level frequencies between mobility-based and residence-based measures have a slightly different pattern (Fig. 3B). We can observe a clear upward averaging in the old towns but the downward averaging in green space exposure is milder and less obvious in the higher end (from 0.33 to 1.00). This means that people who live in old towns with high green space exposures may have low daily mobility or travel to areas with similarly high green space levels in their daily lives. However, the overall tendency of the two frequency distributions still indicates the occurrence of the NEAP: the frequency curve of mobility-based exposures shifted toward the mean value when compared to the frequency curve of residence-based exposures. This means that mobility-based measures can still mitigate some contextual errors of residence-based measures in the old towns (Kwan, 2012).

Our argument on how the NEAP may impact the association estimation originates from the concentration of the variance using mobility-based measures of green space exposure rather than that using residence-based measures. Thus, we conducted F-tests on the variance between the two distributions of green space exposure to statistically confirm the occurrence of the NEAP (Table 3). The F score for either the new towns or the old towns shows a value larger than 1 and is significant at the 0.05 level (i.e., the variance of mobility-based measures is significantly smaller than the variance of residence-based measures). These results indicate that the distribution of green space exposure measures is more dispersed when using residence-based measures than when using mobility-based measures, which means the occurrence of the NEAP. To differentiate the specific manifestations between the old and new towns, we name the case of the new towns the symmetric NEAP case and the case in the old towns the asymmetric NEAP case. We also tested the difference in mean values between the residence-based and mobility-based measures of green space exposure in either the new towns or the old towns (Table 3). We observed an insignificant difference in the new towns and a significant difference in the old towns. However, articulated by Equations (4)–(7), the differences in the mean values of green space exposure only affect the intercepts in the regression models and do not affect the estimation of health associations. These differences thus will not affect our discussions of underestimated health associations.

4.2. *Estimated association magnitudes and the mitigating factor λ*

As indicated by the analytics presented in Section 2, mobility-based measures of green space exposure may reduce the contextual errors of residence-based measures and lead to smaller variance. They may *mitigate the underestimation of the association between green space and health outcomes*, and we should observe a much larger association magnitude using mobility-based measures than residence-based measures. In this study, we employed logistic regression and 940 participants' self-reported overall health statuses as an example. As the model results in Table 4 indicate, green space exposure and health have positive associations using either residence-based or mobility-based measures in the new towns and negative associations using residence-based measures in the old towns. These results are consistent with previous studies and the contradictory associations can be well explained by the different geographic contexts and different primary pathways in either the old towns or the new towns (Y. Liu et al., 2023d; Liu et al., 2024c). For example, residents may have closer contact with green space in the compact old towns, and the risk of vector-borne diseases is higher since green space may serve as the habitat of these vectors. In contrast, green space is better planned in the new towns and farther from residents, whereby the risk of vector-borne diseases may dramatically decrease. Instead, green space may prompt residents' mental health and be associated with better overall health. Our estimated associations between green space exposure and health for the new towns are statistically

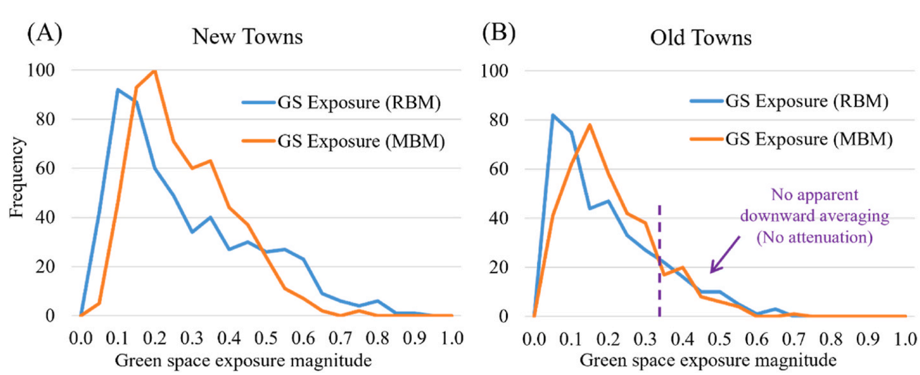

Fig. 3. The frequency plot of green space exposures in (A). the new towns and (B). the old towns. RBM: residence-based measures, MBM: mobility-based measures. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 3
F-tests of the variance concentration for the confirmation of the NEAP and t-tests of the differences in mean values. RBM: residence-based measures, MBM: mobility-based measures.

Test of the difference in variances [a]	Variance of exposure using RBM	Variance of exposure using MBM	RBM set degree of freedom	MBM set degree of freedom		F score	p-value
New towns	$3.46*10^{-2}$	$1.69*10^{-2}$	564	564		2.052	<0.0001
Old towns	$1.91*10^{-2}$	$1.35*10^{-2}$	374	374		1.406	= 0.0005
Test of the difference in mean values [b]	**Mean value of exposure using RBM**	**Mean value of exposure using MBM**	**d value**	**Standard error**	**Degree of freedom**	**t value**	**p-value**
New towns	0.2555	0.2502	0.005284	0.004682	564	1.129	0.260
Old towns	0.1682	0.1803	−0.012081	0.004721	374	−2.559	0.011

[a] F-test of difference in variance.
[b] Paired-sample *t*-test of difference in mean values.

Table 4
The estimated associations between green space exposures and participants' self-reported overall health status and corresponding derived mitigating factors λ. RBM: residence-based measures, MBM: mobility-based measures. CI: confidence interval. S.E.: standard error. The Variance Inflation Factor (VIF) values of the variables in the four models range from 1.065 to 2.481, indicating that multicollinearity may not be problematic in our modeling.

	New Towns (symmetric NEAP case)		Old Towns (asymmetric NEAP case)	
N	565		375	
E_m^{GS} approach (m)	RBM	MBM	RBM	MBM
Association B_m^{GS}	1.303	1.681	−1.754	0.944
Odd ratio	3.681	5.374	0.173	2.571
S.E.of B_m^{GS}	0.719	1.004	1.192	1.509
CI of B_m^{GS} (α = 0.10)	(0.120, 2.486)	(0.030, 3.332)	(-3.715, 0.207)	(-1.538, 3.426)
CI of the odd ratio	(1.127, 12.013)	(1.030, 27.994)	(0.024, 1.230)	(0.215, 30.753)
p-value	0.070	0.094	0.141	0.531
λ_b [a]	1.290		−0.538	
Variance of E_m^{GS}	3.465×10^{-2}	1.689×10^{-2}	1.905×10^{-2}	1.355×10^{-2}
λ_v [a]	1.432		1.186	

[a] Please see Equations (15) and (16) for the derivation of λ_b and λ_v.

significant at the 0.10 level, while our estimated associations for the old towns are not statistically significant. Since this work aims to discuss how mobility-based measures may mitigate the underestimation of the associations rather than confirming the health impact of green space at a very high confidence level, the estimated associations between green space exposures and overall health outcomes, especially in the new towns, would adequately enable our more important discussion of the mitigating factor λ.

As indicated by Equations (8) and (11), we can derive the mitigating factor λ through either estimated association magnitudes (i.e., λ_b) or the variances of green space exposure measures (i.e., λ_v), and they theoretically should both be larger than 1 by definition. In our cases (Table 4), the λ_b for the new towns is 1.290 while the λ_v for the new towns is 1.432, which agrees well with our analytics. In contrast, the λ_b for the old towns is −0.538 while the λ_v for the old towns is 1.186. Since we did not observe statistically significant associations in the old towns using either residence-based or mobility-based measures, the derivation of λ_b in the old towns indeed becomes a case of 0/0, whose evaluation is mathematically undefined. The uncertainties in the estimated associations for the old towns lead to an unpredictable λ_b even though we can observe a λ_v that complies with our analytics. However, we do not know the uncertainties in λ and thus cannot claim that the derived λ_b and λ_v are consistent with each other. Thus, in the next subsection, we provide a more in-depth discussion of the mitigating factor λ using bootstrapping.

4.3. Sensitivity analysis of the mitigating factor λ

We employed a bootstrapping approach to repeatedly derive λ 500 times using randomly selected subsets of our sample. The likelihood distributions (in logarithmic values) of derived λ_b and λ_v are shown in Fig. 4. The distributions of derived λ_b and λ_v are both approximately unimodal distributions, which enables us to systematically test the expectations of the mitigating factor λ. Moreover, we found that the λ_v derived from green space exposure variances is more precise than the λ_b derived from estimated association magnitudes. The precision of the estimation is indicated by the width of the value range. The ranges of the derived λ_v are within (1.36, 1.51) for the new towns, while the ranges of the derived λ_b are within (0.00, 3.00). Similarly, the ranges of the derived λ_v are within (1.10, 1.28) for the old towns, while the ranges of the derived λ_b are most within (−3.00, 0.60). These manifestations agree well with our analytics. Since λ_b can be influenced by several other factors in the modeling, including the association itself, contextual errors, and other covariates, it is reasonable to expect larger uncertainties (wider value ranges) and observe lower precision in the derivation of λ_b than the derivation of λ_v.

We further employed multiple t-tests for an in-depth discussion of the consistency among the derived λ from different approaches (Fig. 5). We found that both the derived λ_b and λ_v for the new towns are not statistically different from each other at the 0.05 level using either the full sample set or subsets (Fig. 5A green linkage). Moreover, both λ_b and λ_v in the new towns are statistically different from 1 (the null hypothesis of no change in associations) at the 0.05 level. These results indicate that the derived λ_b and λ_v in the new towns are consistent with each other and indicate mitigation of the underestimation at a very high confidence level, which is a solid confirmation of our analytics. On the other hand, we observed that the associations between green space exposures and participants' overall health outcomes in the old towns are not significantly different from 0 in Table 4. Consequentially, mitigating the underestimation of 0 still yields 0. In this case, it is not possible to derive a reliable λ_b in the undefined form of 0/0, indicated by the significant disparity between λ_b and λ_v in Fig. 5B.

5. Discussions

5.1. Interpretation of this study

In this study, we argued that the NEAP may lead to underestimations of the associations between people's green space exposures and health outcomes, and mobility-based measures of green space exposure may mitigate the underestimations by reducing contextual errors. We analytically illustrated the mechanism through a mitigating factor λ and empirically assessed our analytics using data from 940 participants from Hong Kong.

We confirmed the occurrence of the NEAP in green space exposure in both the new and old towns of Hong Kong, successfully predicted the

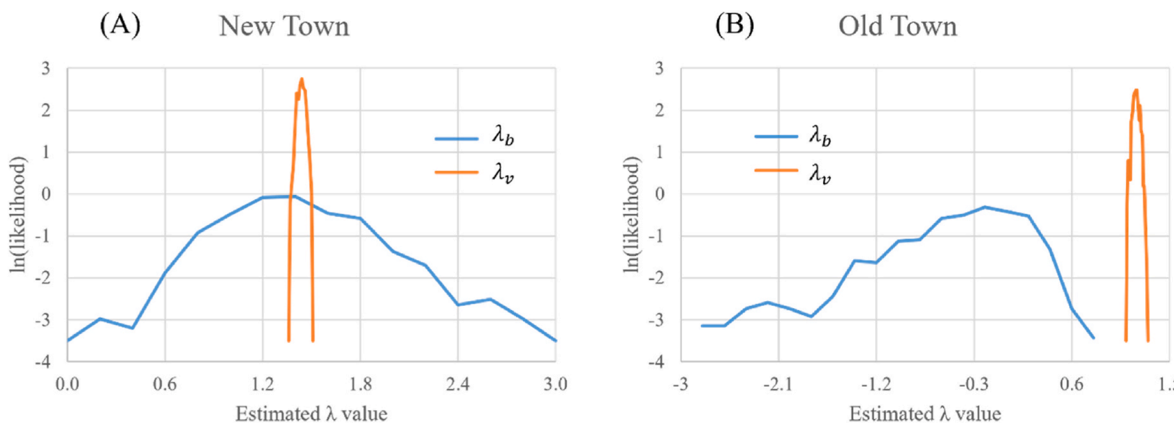

Fig. 4. The likelihood distributions (in logarithmic values for better illustration) of derived λ using bootstrapping: (A). λ_b and λ_v in the new towns, and (B). λ_b and λ_v in the old towns.

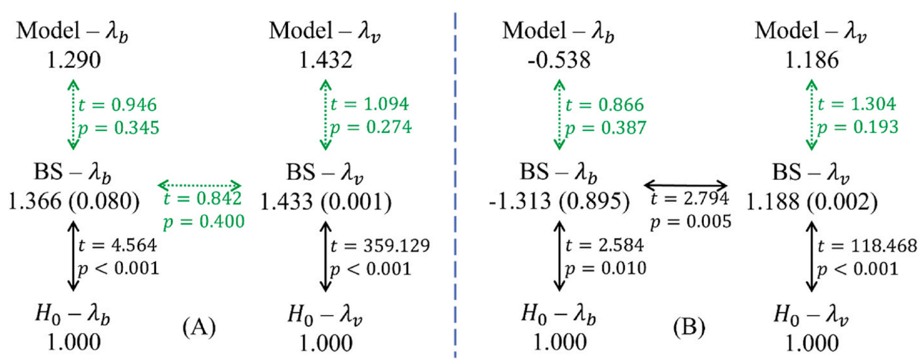

Fig. 5. The t-tests of consistency between λ_b and λ_v (A). in the new towns and (B). in the old town. H_0: the null hypothesis of no change in the estimated associations, i.e., $\lambda = 1$; BS: average value (expectation) of λ using bootstrapping, two-thirds randomly selected samples, 500 repeats, the standard error of expectation is enclosed by parenthesis; Model: the estimated λ using the full sample set, no available theory enables the uncertainty evaluation in this case. Two-tailed t-tests, differences are highlighted in green if not statistically different from 0 at the 0.05 level. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

underestimations in the associations between green space and health outcomes for the new towns, and derived the magnitude of the underestimations through the mitigating factor λ. In our new town model, we observed significant associations between green space exposure and participants' overall health outcomes, using either residence-based or mobility-based measures. The S.E. of the estimated associations are similar. These manifestations indicate that the estimated effect sizes using either residence-based or mobility-based exposure measures are of equivalent *precision* to avoid Type II errors in our new town models. However, the significant mitigating factor (λ) observed in the new town models further indicates that the mobility-based measures lead to a more *accurate* estimated effect size by mitigating the contextual errors in the residence-based measures.

Note that the change in the estimated associations (mitigation magnitude of underestimation) can be associated with λ_b as

$$C(B) = \frac{B^{GS}_{mbm} - B^{GS}_{rbm}}{B^{GS}_{rbm}} = \frac{B^{GS}_{mbm}}{B^{GS}_{rbm}} - 1 = \lambda_b - 1 \tag{17}$$

In the new towns, the derived λ_b is about 1.366, with a 95 % confidence interval of (1.209, 1.523). These results indicate that about 20.9 % – 52.3 % of the underestimation can be mitigated with very high confidence by using the mobility-based measures of green space exposure rather than using the residence-based measures. The mitigated underestimation resulted from mitigating the contextual errors, especially both the upward and downward averaging manifested by the NEAP.

5.2. Implication of this study

Multiple methodological studies have highlighted the Type II errors in health outcome modeling using residence-based measures (Liu and Kwan, 2024; Y. Liu et al., 2023d). Previous observations indicate that there may be insignificant associations between residence-based measures and health outcomes but significant associations with mobility-based measures in the same health outcome modeling. This study shows a totally new and different manifestation. Although residence-based and mobility-based measures can both have significant associations with health outcomes in some cases, the modeling using mobility-based measures may not only avoid wrong conclusions (Type II errors) but also yield more accurate estimated associations by mitigating the contextual errors in the residence-based measures. As our results indicate, about 20.9 % – 52.3 % of the underestimated associations between green space exposures and health outcomes (due to contextual errors) can be mitigated by using mobility-based measures of green space exposure compared to when using residence-based measures. This is a considerable amount of exposure measurement error in estimating green space's possible health impacts. It highlights that critical attention is needed when measuring green space exposure and estimating its health impacts if we want to obtain accurate estimations of green space's health impacts, which become increasingly practical and important for various public health concerns.

Based on our analytics, it is also apparent that mobility-based exposure measures can help reduce the underestimation of the health impacts of an environmental factor due to contextual errors and the

NEAP. In this study, we employed green space exposure as an example to evaluate the analytics we put forward. Our analytics and the mitigating factor will also be useful for mitigating contextual errors and the underestimation of the health impacts of other mobility-dependent environmental factors, such as air pollution (Dewulf et al., 2016; Kim and Kwan, 2021a, 2021b), infectious diseases (Huang and Kwan, 2022), traffic congestion (Kim and Kwan, 2019), and urban noise. Our analytics indeed provide significant suggestions and implications for a broad scope of epidemiological, environmental, and public health studies that share similar interests in the health impacts of various environmental factors.

However, we need to highlight that our analytics and the mitigation of underestimated effect sizes can only be valid when the NEAP exists in people's exposure to environmental factors. Since we observed an apparent NEAP in our study area and significant associations in the new towns, we can validate our analytics with very strong evidence. In other contexts where the NEAP does not occur or for other environmental factors that are not mobility-dependent, it is not proper to directly apply our conclusions.

5.3. Limitations of this study

Although we successfully illustrated our arguments using cross-sectional data from Hong Kong, this study still faces some limitations. First, we observed significant positive associations between green space and health in our study, but we are unclear on how the adverse health impacts of other environmental impact factors (e.g., air pollutants) may have influenced our results. Second, our sample may be inadequate for uncovering the disparities between different socio-demographic groups when mitigating the underestimated associations between green space and health. Future data collections with larger sample sizes, more precise health outcome measures, and other environmental exposure measures may further deepen our understanding of these methodological issues. Third, we employed bootstrapping to estimate the distribution of the mitigating factor (λ) and the repeat times may slightly affect our estimation of λ's expectation. We conducted a sensitivity test with 5000 repeats. The expectation of λ_b is 1.399 with a 95 % CI of (1.297, 1.500). However, it does not change any of our discussions and conclusions. Finally, our study employed participants' self-reported overall health, and we did not highlight concrete pathways since it is beyond the scope of this paper. Y. Liu et al. (2024c) have analytically articulated the disparities between the paradigms of residence-based and mobility-based exposure measures. Mobility-based measures show advantages in most cases of mobility-dependent environmental exposures, such as green space exposure examined in this study. However, residence-based exposures may be more relevant than mobility-based exposures regarding some specific health concerns. For example, exposure to indoor vegetation, neuroendocrine disturbance from room light at night, and air pollution from cooking may primarily occur at home for housekeepers. They are, by definition, not suitable to be examined by using mobility-based exposure measures. We thus clarify that theorized pathways are necessary for generating useful substantive insights in public health studies.

6. Conclusions

In this study, we demonstrated that mobility-based measures of green space exposure may significantly mitigate the underestimation of the associations between green space exposures and health due to contextual errors and the NEAP. The reduction in such underestimation can be quantified through a mitigating factor. We found that participants in the new towns of Hong Kong experienced the NEAP in their green space exposures. We derived the consistent mitigating factors that are statistically different from 1 at a very high confidence level. The mitigated underestimation ranges from 20.9 % to 52.3 %, which is a striking change in the estimated associations between green space exposure and human health. While green space exposure is employed as an example in this study, our analytics and the mitigating factor can also be applied to study other mobility-dependent environmental factors. Thus, our study has crucial implications for a broad range of environmental and public health studies that need accurate estimation of the possible health impacts of a range of environmental factors.

CRediT authorship contribution statement

Yang Liu: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization. **Mei-Po Kwan:** Writing – review & editing, Writing – original draft, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Liuyi Song:** Writing – review & editing. **Changda Yu:** Writing – review & editing. **Yuhan Cui:** Writing – review & editing.

Consent to participate

Informed (written) consent was obtained from all subjects involved in the study before data collection began.

Consent to publish

Not applicable to this study.

Ethics approval

The study was approved by the *Survey and Behavioral Research Ethics (SBRE) Committee* of the Chinese University of Hong Kong (Reference No. SBRE-19-123 approved on January 8, 2020 and Reference No. SBRE (R)-21-005 approved on November 1, 2021).

Funding

This research was supported by grants from the Hong Kong Research Grants Council (General Research Fund Grant no. 14605920, 14606922, 14603724; Collaborative Research Fund Grant no. C4023-20 GF; Research Matching Grants RMG 8601219, 8601242, 3110151), a grant from the Research Committee on Research Sustainability of Major Research Grants Council Funding Schemes (3133235) of the Chinese University of Hong Kong (CUHK), and a grant from the Vice-Chancellor's One-off Discretionary Fund (Smart and Sustainable Cities: City of Commons) (4930787) of CUHK. The funders had no role in the study design, data collection, data analysis, decision to publish, or preparation of the manuscript.

Conflicts of interest

The authors declare they have no conflicts of interest related to this work to disclose.

Data availability

The data that have been used are confidential.

References

Beasley, W.H., Rodgers, J.L., 2012. Bootstrapping and Monte Carlo methods. In: APA Handbook of Research Methods in Psychology, Vol 2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological. American Psychological Association, pp. 407–425. https://doi.org/10.1037/13620-022.

Bertram, C., Rehdanz, K., 2015. The role of urban green space for human well-being. Ecol. Econ. 120, 139–152.

Chaix, B., 2018. Mobile sensing in environmental health and neighborhood research. Annu. Rev. Publ. Health 39, 367–384. https://doi.org/10.1146/annurev-publhealth-040617-013731. Volume 39, 2018.

Chaix, B., Méline, J., Duncan, S., Jardinier, L., Perchoux, C., Vallée, J., Merrien, C., Karusisi, N., Lewin, A., Brondeel, R., Kestens, Y., 2013. Neighborhood environments, mobility, and health: towards a new generation of studies in environmental health research. Rev. Epidemiol. Sante Publique 61, S139–S145. https://doi.org/10.1016/j.respe.2013.05.017.

Department, C.a.S., 2023. Mid-year Population for 2023 [15 Aug 2023]. The Government of the Hong Kong Special Administrative Region. Retrieved December 29 from. https://www.censtatd.gov.hk/en/press_release_detail.html?id=5265.

Dewulf, B., Neutens, T., Lefebvre, W., Seynaeve, G., Vanpoucke, C., Beckx, C., Van de Weghe, N., 2016. Dynamic assessment of exposure to air pollution using mobile phone data. Int. J. Health Geogr. 15, 1–14.

Dismuke, C., Lindrooth, R., 2006. Ordinary least squares. Methods Des. Outcomes Res. 93 (1), 93–104.

Dzhambov, A.M., Browning, M.H., Markevych, I., Hartig, T., Lercher, P., 2020. Analytical approaches to testing pathways linking greenspace to health: a scoping review of the empirical literature. Environ. Res. 186, 109613.

Gryparis, A., Paciorek, C.J., Zeka, A., Schwartz, J., Coull, B.A., 2009. Measurement error caused by spatial misalignment in environmental epidemiology. Biostatistics 10 (2), 258–274.

Gyanwali, S., Karki, S., Jang, K.M., Crawford, T., Zhang, M., Kim, J., 2024. Implications for spatial non-stationarity and the neighborhood effect averaging problem (NEAP) in green inequality research: evidence from three states in the USA. J. Geogr. Syst. 1–24.

Hatch, M., Thomas, D., 1993. Measurement issues in environmental epidemiology. Environ. Health Perspect. 101 (Suppl. 4), 49–57.

Haukoos, J.S., Lewis, R.J., 2005. Advanced statistics: bootstrapping confidence intervals for statistics with "Difficult" distributions. Acad. Emerg. Med. 12 (4), 360–365. https://doi.org/10.1197/j.aem.2004.11.018.

Huang, J., Kwan, M.-P., 2022. Uncertainties in the assessment of COVID-19 risk: a study of people's exposure to high-risk environments using individual-level activity data. Ann. Assoc. Am. Geogr. 112 (4), 968–987.

Jiang, S., Ferreira, J., Gonzalez, M.C., 2017. Activity-based human mobility patterns inferred from mobile phone data: a case study of Singapore. IEEE Trans. Big Data 3 (2), 208–219.

Judkins, D.R., Porter, K.E., 2016. Robustness of ordinary least squares in randomized clinical trials. Stat. Med. 35 (11), 1763–1773.

Kan, Z., Kwan, M.-P., Cai, J., Liu, Y., Liu, D., 2023. Nonstationary relationships among individuals' concurrent exposures to noise, air pollution and greenspace: a mobility-based study using GPS and mobile sensing data. Health Place 83, 103115.

Kim, J., Kwan, M.-P., 2019. Beyond commuting: ignoring individuals' activity-travel patterns may lead to inaccurate assessments of their exposure to traffic congestion. Int. J. Environ. Res. Publ. Health 16 (1), 89.

Kim, J., Kwan, M.-P., 2021a. Assessment of sociodemographic disparities in environmental exposure might be erroneous due to neighborhood effect averaging: implications for environmental inequality research. Environ. Res. 195, 110519.

Kim, J., Kwan, M.-P., 2021b. How neighborhood effect averaging might affect assessment of individual exposures to air pollution: a study of ozone exposures in Los Angeles. Ann. Assoc. Am. Geogr. 111 (1), 121–140.

Kou, L., Tao, Y., Kwan, M.-P., Chai, Y., 2020. Understanding the relationships among individual-based momentary measured noise, perceived noise, and psychological stress: a geographic ecological momentary assessment (GEMA) approach. Health Place 64, 102285.

Kwan, M.-P., 2012. The uncertain geographic context problem. Ann. Assoc. Am. Geogr. 102 (5), 958–968.

Kwan, M.-P., 2016. Algorithmic geographies: big data, algorithmic uncertainty, and the production of geographic knowledge. Ann. Assoc. Am. Geogr. 106 (2), 274–282.

Kwan, M.-P., 2018. The neighborhood effect averaging problem (NEAP): an elusive confounder of the neighborhood effect. Int. J. Environ. Res. Publ. Health 15 (9), 1841.

Li, Q., Zheng, Y., Xie, X., Chen, Y., Liu, W., Ma, W.-Y., 2008. Mining user similarity based on location history. In: Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.

Liu, D., Kwan, M.-P., Kan, Z., Liu, Y., 2023a. Examining individual-level tri-exposure to greenspace and air/noise pollution using individual-level GPS-based real-time sensing data. Soc. Sci. Med. 329, 116040.

Liu, Y., Kwan, M.-P., 2024. Mobility-oriented measurements of people's exposure to outdoor artificial light at night (ALAN) and the uncertain geographic context problem (UGCoP). PLoS One 19 (4), e0298869. https://doi.org/10.1371/journal.pone.0298869.

Liu, Y., Kwan, M.-P., Kan, Z., 2023b. Inconsistent association between perceived air quality and self-reported respiratory symptoms: a pilot study and implications for environmental health studies. Int. J. Environ. Res. Publ. Health 20 (2), 1491.

Liu, Y., Kwan, M.-P., Wang, J., 2024a. Analytically articulating the effect of buffer size on urban green space exposure measures. Int. J. Geogr. Inf. Sci. 1–22.

Liu, Y., Kwan, M.-P., Wang, J., Cai, J., 2024b. Confounding associations between green space and outdoor artificial light at night: systematic investigations and implications for urban health. Environ. Sci. Ecotechnol., 100436

Liu, Y., Kwan, M.-P., Wong, M.S., Yu, C., 2023c. Current methods for evaluating People's exposure to green space: a scoping review. Soc. Sci. Med., 116303

Liu, Y., Kwan, M.-P., Yu, C., 2023d. The uncertain geographic context problem (UGCoP) in measuring people's exposure to green space using the integrated 3S approach. Urban For. Urban Green. 85, 127972.

Liu, Y., Kwan, M.-P., Yu, C., 2024c. How mobility and temporal contexts may affect environmental exposure measurements: using outdoor artificial light at night (ALAN) and urban green space as examples. Ann. Assoc. Am. Geogr. 1–20. https://doi.org/10.1080/24694452.2024.2319067.

Ma, J., Tao, Y., Kwan, M.-P., Chai, Y., 2021. Assessing mobility-based real-time air pollution exposure in space and time using smart sensors and GPS trajectories in Beijing. In: Smart Spaces and Places. Routledge, pp. 102–116.

Martin, H., Hong, Y., Wiedemann, N., Bucher, D., Raubal, M., 2023. Trackintel: an open-source Python library for human mobility analysis. Comput. Environ. Urban Syst. 101, 101938.

Perchoux, C., Chaix, B., Cummins, S., Kestens, Y., 2013. Conceptualization and measurement of environmental exposure in epidemiology: accounting for activity space related to daily mobility. Health Place 21, 86–93. https://doi.org/10.1016/j.healthplace.2013.01.005.

Perchoux, C., Kestens, Y., Thomas, F., Hulst, A.V., Thierry, B., Chaix, B., 2014. Assessing patterns of spatial behavior in health studies: their socio-demographic determinants and associations with transportation modes (the RECORD Cohort Study). Soc. Sci. Med. 119, 64–73. https://doi.org/10.1016/j.socscimed.2014.07.026.

Pohlman, J.T., Leitner, D.W., 2003. A Comparison of Ordinary Least Squares and Logistic Regression.

Rhomberg, L.R., Chandalia, J.K., Long, C.M., Goodman, J.E., 2011. Measurement error in environmental epidemiology and the shape of exposure-response curves. Crit. Rev. Toxicol. 41 (8), 651–671.

Richmond-Bryant, J., Long, T.C., 2020. Influence of exposure measurement errors on results from epidemiologic studies of different designs. J. Expo. Sci. Environ. Epidemiol. 30 (3), 420–429.

Spielman, S.E., Yoo, E.-h., 2009. The spatial dimensions of neighborhood effects. Soc. Sci. Med. 68 (6), 1098–1105. https://doi.org/10.1016/j.socscimed.2008.12.048.

Spielman, S.E., Yoo, E.-H., Linkletter, C., 2013. Neighborhood contexts, health, and behavior: understanding the role of scale and residential sorting. Environ. Plann. Plann. Des. 40 (3), 489–506. https://doi.org/10.1068/b38007.

Team, P., 2017. Planet Application Program Interface: in Space for Life on Earth. San Francisco, CA, 2017, 40.

Tian, Y., Jim, C.Y., Tao, Y., Shi, T., 2011. Landscape ecological assessment of green space fragmentation in Hong Kong. Urban For. Urban Green. 10 (2), 79–86.

Wang, J., Kou, L., Kwan, M.-P., Shakespeare, R.M., Lee, K., Park, Y.M., 2021. An integrated individual environmental exposure assessment system for real-time mobile sensing in environmental health studies. Sensors 21 (12), 4039.

Wang, J., Kwan, M.-P., Xiu, G., Peng, X., Liu, Y., 2024. Investigating the neighborhood effect averaging problem (NEAP) in greenspace exposure: a study in Beijing. Landsc. Urban Plann. 243, 104970.

Wang, J., Liu, Y., Kwan, M.-P., 2025. Cross-validation between GPS-derived trajectories and activity-travel diaries for transport geography studies. J. Transport Geogr. 126, 104239. https://doi.org/10.1016/j.jtrangeo.2025.104239.

Xu, T., Wang, S., Liu, Q., Kim, J., Zhang, J., Ren, Y., Ta, N., Wang, X., Wu, J., 2023. Vegetation color exposure differences at the community and individual levels: an explanatory framework based on the neighborhood effect averaging problem. Urban For. Urban Green. 86, 128001.

Yu, C., Kwan, M.-P., Liu, Y., 2024. Assessing momentary stress responses to dynamic real-time greenspace exposure: unveiling algorithmic uncertainty and the temporality of exposure context. Soc. Sci. Med. 363, 117411.

Zheng, L., Kwan, M.-P., Liu, Y., Liu, D., Huang, J., Kan, Z., 2024. How mobility pattern shapes the association between static green space and dynamic green space exposure. Environ. Res. 258, 119499.