Social Science & Medicine 399 (2026) 119254

Contents lists available at ScienceDirect

Social Science & Medicine

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

Profiling spatiotemporal contexts associated with compound exposure disadvantages in noise and greenery: A study using mobile sensors and explainable machine learning

Linsen Wang [a], Mei-Po Kwan [a,b,*]

[a] Institute of Space and Earth Information Science, Fok Ying Tung Remote Sensing Science Building, The Chinese University of Hong Kong, Hong Kong Special Administrative Region of China

[b] Department of Geography and Resource Management, Wong Foo Yuan Building, The Chinese University of Hong Kong, Hong Kong Special Administrative Region of China

ARTICLE INFO

Keywords:
Greenery exposure
Noise exposure
Compound disadvantages
Spatiotemporal contexts
Mobile sensors
Explainable machine learning

ABSTRACT

Although exposures to noise or greenery in isolation have been widely studied, limited research has explored the nonlinear associations between spatiotemporal contexts and compound disadvantages from a daily mobility perspective. Drawing on data from mobile sensors, GPS, satellite imagery, points of interest, and activity diaries, this study measures real-time noise and greenery exposures along relevant spatiotemporal contexts. Methodologically, daily mobility is delineated through a spatiotemporal threshold approach, modeling them as stay and move event sequences that underpin the development of a compound disadvantage index. A novel framework, integrating GPBoost with residual-based bootstraps and SHAP, subsequently accommodates time-sequence dependencies and uncertainties to investigate nonlinear relationships between spatiotemporal contexts and compound disadvantages. The main findings indicate that (i) the neighborhood effect averaging problem is apparent in noise and greenery measurements; (ii) socioeconomic status coupled with spatiotemporal contexts can infer individual compound disadvantages; (iii) employment status constitutes the foremost socioeconomic predictor of compound disadvantages; and (iv) the bootstrap-SHAP methodology pinpoints spatiotemporal contexts with statistically significant effects on compound disadvantages, concurrently disclosing thresholds and interactions with socioeconomic statuses.

1. Introduction

Globally, rapid urbanization has transformed landscapes, resulting in marked disturbance and pollution in the audio-visual environment of people's daily lives (Chu et al., 2025; Hartig et al., 2003; Jeon and Jo, 2020). People perceive the environment through audio-visual sensation and the interaction between the two, which influences their well-being (Hasegawa and Lau, 2022). In the urban acoustic environment, noise pollution is a primary concern. A substantial proportion of the global population is exposed to noise pollution (World Health Organization, 2018; Ministry of Ecology and Environment of China, 2020). There is a large body of evidence linking noise to various physical and mental disorders, including hearing loss, cardiovascular disease, insomnia, stress, depression, and cognitive deficit (Basner et al., 2014; Münzel et al., 2014; Ma et al., 2020). In the urban visual environment, the scarcity of beneficial greenery is a major issue. Urban residents have limited exposure to natural environments. Previous studies have revealed that greenery exposure influences human well-being through multiple pathways, including reducing harmful exposure (e.g., noise pollution), restoring mental resources, and promoting physical activities and social cohesion (De Vries et al., 2003, 2013; Hartig et al., 2014; James et al., 2016). The pursuit of a tranquil and naturalistic environment has become an essential goal (Chu et al., 2025).

While sole exposure to noise or greenery has received extensive attention, compound exposure has not been holistically investigated. In this study, compound exposure refers to the simultaneous experience of noise and greenery at the individual level throughout daily activities. At the geographical level, greenery and noise often exhibit spatially

* Corresponding author. Institute of Space and Earth Information Science, Fok Ying Tung Remote Sensing Science Building, The Chinese University of Hong Kong, Hong Kong Special Administrative Region of China.

E-mail addresses: wanglinsen@link.cuhk.edu.hk (L. Wang), mpkwan@cuhk.edu.hk (M.-P. Kwan).

https://doi.org/10.1016/j.socscimed.2026.119254
Received 30 November 2025; Received in revised form 25 March 2026; Accepted 31 March 2026
Available online 1 April 2026
0277-9536/

exclusive competition with a generally negative relationship. However, empirical evidence reveals more nuanced patterns at the individual level. Kan et al. (2023) found only a weak negative association between individual noise and greenery exposures, while Liu et al. (2023) showed that the co-occurrence of high greenery and high noise is not rare, often associated with neighborhood socioeconomic status (SES). These findings suggest that geographical settings may not accurately reflect actual exposure patterns (Liu et al., 2024). Although greenery can attenuate noise through physical mechanisms (Horoshenkov et al., 2013; Ow and Ghosh, 2017) and perceptual restorative effects (Chau et al., 2018; Dzhambov and Dimitrova, 2015), the actual exposure patterns individuals experience may be complex. For instance, Klompmaker et al. (2019) found that combined effects exceeded those of sole exposures, with low greenery potentially magnifying high noise impacts, while Gascon et al. (2018) demonstrated that greenery's influence on well-being varies across geographic spaces. Consequently, a critical research gap exists in holistically measuring and understanding compound exposures, particularly compound disadvantage (defined as concurrent high noise and low greenery exposure), which has important implications for understanding environmental exposure inequalities and their health consequences. Addressing this gap requires approaches that can capture the co-occurrence and interaction of noise and greenery.

Further, existing studies mostly rely on static exposure measurements, with fixed monitoring stations and geographic units employed to measure noise and greenery exposures. These studies assumed uniform environmental exposure within the same geographic unit, typically using residence-based geographic context as the primary exposure measurement unit (Casey et al., 2017; Collins et al., 2020; Hartig et al., 2014; James et al., 2016; Klompmaker et al., 2019; Mitchell and Popham, 2008). With large spatial and temporal coverage (nationwide or longitudinal), these studies effectively revealed environmental inequities and risk factors, guiding public resource allocation and optimization. However, static exposures ignored people's daily mobility, resulting in exposure measurement bias and a misestimation effect. Kwan (2012) has identified this issue as the uncertain geographic context problem (UGCoP), which indicates that residents living in the same neighborhood could be exposed to different geographic contexts. Compared to residence-based exposure, Kim and Kwan (2021) further found that mobility-based exposure might converge toward the average exposure of the whole population, articulated as the neighborhood effect averaging problem (NEAP) proposed by Kwan (2018). For more precise measurement of noise and greenery exposures, dynamic exposure measures should be implemented to capture the spatiotemporal information of residents' daily mobility. Mobile sensors offer promising solutions to address the challenges posed by the UGCoP and the NEAP. Using GPS trajectories and mobile sensors, recent studies have demonstrated the feasibility of measuring dynamic noise exposure (Cai et al., 2023; Wang et al., 2025) and dynamic greenery exposure (Liu et al., 2024; Yu et al., 2024) along individuals' daily activity paths. However, these studies mostly apply a spatiotemporal weighting or sampling approach to real-time data, leaving the spatiotemporal information not fully exploited. For instance, spatiotemporal weighting approaches often aggregate exposure data to individual-level summaries or specific activity locations (e.g., home or workplace) using time-weighted averages, thereby overlooking the actual spatiotemporal contexts during activity episodes. Similarly, time sampling approaches, such as ecological momentary assessment (EMA), collect exposure data at predetermined time points (e.g., 12:00, 20:00), which fragments the complete temporal sequence and fails to capture continuous exposure dynamics throughout the day. These methodological limitations in handling fine-grained spatiotemporal data, combined with the insufficient application of advanced methods for modeling multi-subject time-series data, have prevented researchers from uncovering the nonlinear relationships between individuals' spatiotemporal mobility contexts and compound exposures. Specifically, the aggregation or fragmentation of spatiotemporal information reduces the analytical resolution needed to detect complex, context-dependent patterns. Moreover, traditional statistical methods (e.g., linear regression, generalized linear models) commonly used in environmental exposure research are limited in their ability to capture such nonlinear interactions and threshold effects, particularly when dealing with high-dimensional spatiotemporal data from multiple individuals with hierarchical structures (i.e., repeated measurements nested within individuals).

Fortunately, the emergence of explainable machine learning methods presents novel opportunities for addressing these challenges. For example, GPBoost (i.e., a tree boosting method that combines Gaussian processes and mixed-effects modeling) offers a powerful approach specifically designed to handle massive repeated measurements across individuals over time (Sigrist, 2022). GPBoost is well-suited for this task because it integrates three complementary components: (1) tree boosting captures complex nonlinear relationships and high-order interactions among high-dimensional spatiotemporal features without requiring pre-specified functional forms; (2) Gaussian processes model the spatial and temporal autocorrelation inherent in individuals' movement trajectories, accounting for the fact that exposures at nearby locations or adjacent time points are not independent; and (3) mixed-effects modeling handles the hierarchical data structure by accounting for both between-individual heterogeneity and within-individual variability across repeated measurements. This integration enables GPBoost to efficiently model the complex dependencies in multi-subject time-series data while maintaining computational tractability for large-scale datasets. Coupled with SHAP (SHapley Additive exPlanations), which quantifies the marginal importance of each feature to model predictions (Lundberg et al., 2020), this approach can reveal the nonlinear relationships and interaction effects of spatiotemporal contextual factors on compound exposure patterns.

In summary, three critical research gaps remain. First, existing studies have primarily focused on sole exposure rather than compound exposure patterns (particularly compound disadvantage). Moreover, the reliance on static residence-based measurements fails to account for individuals' daily mobility, leading to inaccurate exposure assessment. Second, methodological limitations in handling fine-grained spatiotemporal data have prevented full exploitation of the rich variability embedded in mobility trajectories, obscuring how compound exposure patterns actually vary across individuals, locations, and time. Third, the insufficient application of advanced methods for modeling multi-subject time-series data has hindered understanding of the nonlinear relationships between spatiotemporal contexts and compound exposure patterns, as well as the interactive effects between SES and spatiotemporal contexts in shaping these patterns. Given these gaps, this study investigates compound disadvantages from a dynamic mobility perspective through three research questions. RQ1: How do static residence-based and dynamic mobility-based measurements differ in capturing compound disadvantage? (addressing Gap 1). RQ2: What are the spatiotemporal distributions of compound disadvantage across individuals, activity locations, and temporal contexts, and to what extent can socioeconomic status (SES) and spatiotemporal contexts explain these variations? (addressing Gap 2). RQ3: Which specific spatiotemporal contexts show significant nonlinear associations with compound disadvantage and do SES interact with spatiotemporal contexts in shaping compound disadvantage patterns? (addressing Gap 3).

2. Data and methodology

To answer these questions, we employed three integrated strategies (Fig. 1). First, portable sound meters, satellite imagery, and GPS-enabled smartphones measured dynamic noise and greenery exposure along participants' daily trajectories, while activity-travel diaries and points of interest (POIs) documented spatiotemporal contexts. We calculated both residence-based exposure and mobility-based exposure to enable direct comparison (RQ1). Second, we analyzed compound exposure at the stay

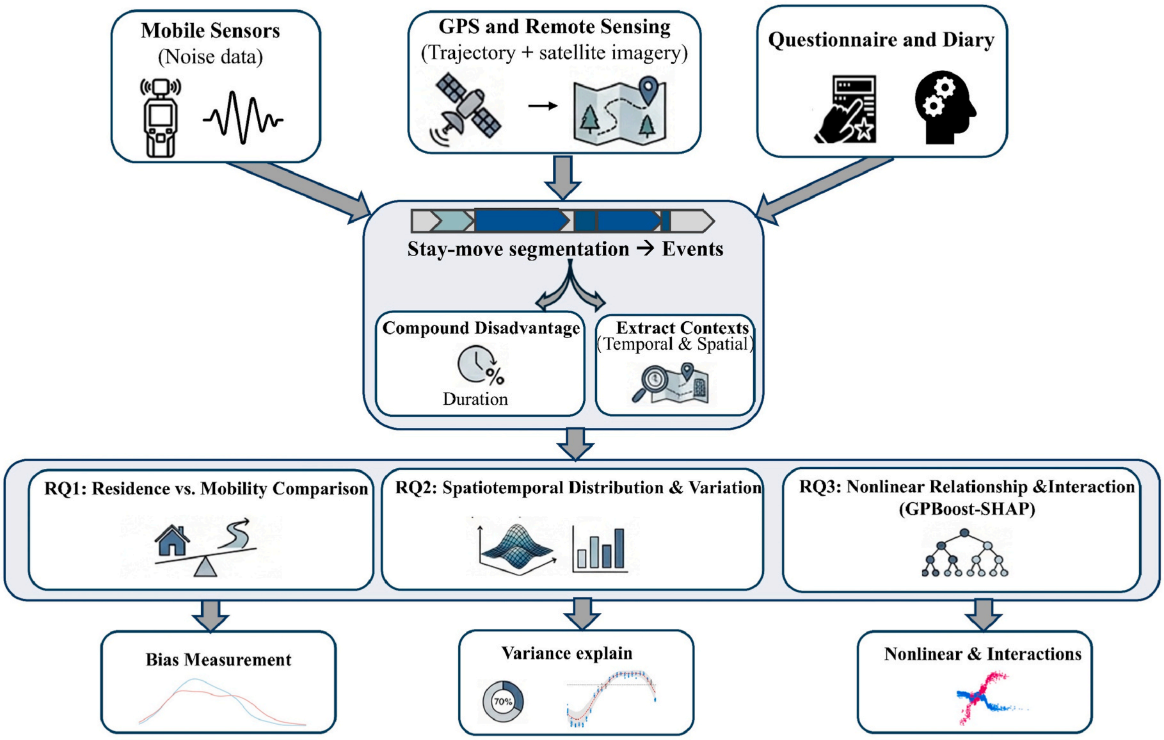

Fig. 1. The framework of the workflow.

and move event level to preserve complete spatiotemporal sequences. For each event, we constructed a compound disadvantage index and linked it with fine-grained contexts. This approach fully exploits spatiotemporal variability to reveal how compound disadvantage is distributed across different individuals, locations, and time periods (RQ2). Third, we employed GPBoost-SHAP, an explainable machine learning framework designed for multi-subject time-series data, to capture complex nonlinear relationships while accounting for spatial-temporal autocorrelation and hierarchical structures. This approach quantifies the nonlinear effects of specific spatiotemporal contexts and identifies the interactive effects between SES and spatiotemporal contexts in shaping compound disadvantage patterns (RQ3).

2.1. Study area and survey

This study was conducted in Hong Kong, which is located in southern China. We collected data from 800 participants between November 2021 and March 2023. Participants were distributed across four quarters with relatively balanced seasonal coverage (Q1: 19.1%, Q2: 25.1%, Q3: 26.6%, Q4: 36.0%). Each participant was surveyed once during a two-day period (one weekday and one weekend day), with no overlap across different time periods. Participants mainly reside in four areas: Kwai Tsing, Sha Tin, Central and Western, and Kwun Tong (Fig. 2). Each area comprises multiple Large Subunit Groups (LSUG), which serve as census tracts in Hong Kong. These four areas were strategically chosen to capture Hong Kong's substantial urban diversity. Hong Kong's urban development history has produced distinct settlement patterns, with areas developed before the 1960s (Central and Western, Kwun Tong) exhibiting different built environment characteristics compared to those developed after the 1980s (Sha Tin, Kwai Tsing). These differences manifest in varying building densities, street configurations, and environmental quality. Geographically, the selected areas span Hong Kong's four principal regions (Hong Kong Island, Kowloon, New Territories East, and New Territories West) ensuring coverage of the territory's major spatial divisions. Functionally and economically, the areas represent a spectrum of urban characteristics: Central and Western combines high-income residential neighborhoods with commercial business districts; Sha Tin integrates residential and retail functions with moderate income levels; Kwun Tong hosts manufacturing and industrial activities; and Kwai Tsing consists predominantly of residential developments, with the latter two areas characterized by comparatively lower income levels. This selection strategy ensures our study sites encompass varied built environment configurations, socioeconomic profiles, and environmental exposure conditions (including noise pollution and greenery availability) representative of Hong Kong's urban landscape. Participant recruitment within each area employed stratified sampling informed by Hong Kong Population Census statistics (https://www.censtatd.gov.hk/en/). The sampling framework matched the demographic composition of each area's population across five key characteristics: age distribution, gender ratio, household income brackets, educational qualification levels, and marital status categories. This methodological approach ensures our participant sample reflects the demographic structure of populations within the four study districts. Ethical approval was obtained from the Survey and Behavioral Research Ethics (SBRE) committee at the authors' institution (Protocol SBRE-19-123, approved 8 January 2020; Protocol SBRE(R)-21-005, approved 1 November 2021). All participants provided written informed consent prior to their participation.

The survey procedure is as follows. First, participants were asked to complete a sociodemographic questionnaire that contained information such as age, gender, household income, education level, employment status, and residential address. Moreover, we defined the neighborhood type using the proportion of public housing within the LSUG that corresponded to each participant's residential address. LSUG with a 100% proportion of public housing were designated as public housing neighborhoods, those with a proportion of public housing greater than 0% but less than 100% as mixed neighborhoods, and those with a 0% proportion of public housing as private housing neighborhoods. Second, each participant was required to carry a portable sound level meter (i.e., SLM-25, which complies with IEC 61672 Type 2 and ANSI S1.4 Type 2 standards) to measure equivalent continuous A-weighted sound pressure levels every 30 s. Sound level calibrators (i.e., CEM SC-05, which

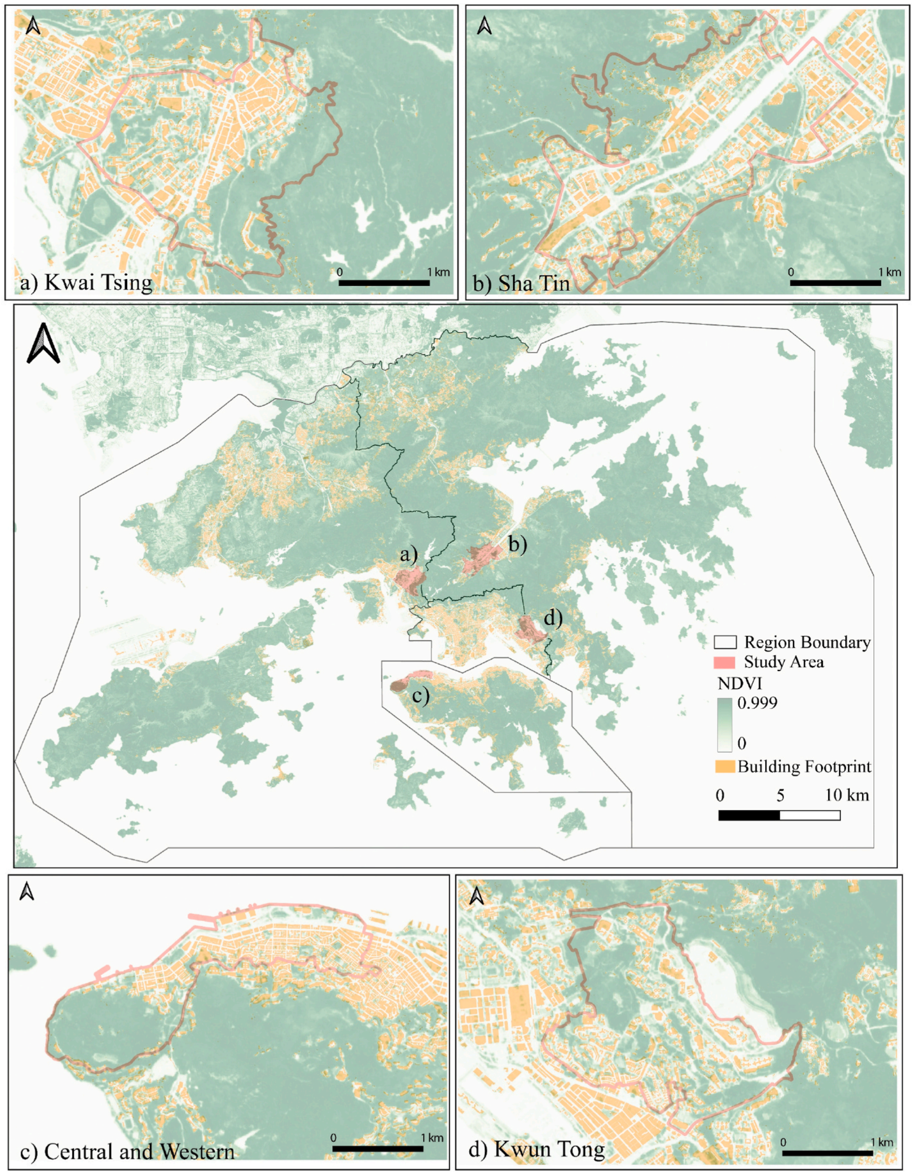

Fig. 2. The study area.

complies with the IEC 60942 standard) were used to calibrate the portable sound meters. Moreover, participants were also requested to carry a mobile phone to record real-time GPS locations. Lastly, each participant filled out a two-day activity diary (one weekday and one weekend day) to document their daily activity types, activity schedules, and the spatial and temporal flexibility of their activities.

We integrated participants' GPS trajectories, sound level meter data, and activity diary entries at a 1-min resolution, yielding 749 valid participants (Table 1) and a total of 1,434,794 observations. 51 participants (6.4%) were excluded based on the following criteria. 1) Missing device records: participants with incomplete GPS trajectories, sound level meter data, or activity diary entries for the required two-day period; 2) Anomalous GPS trajectories: trajectories with coordinates drifting beyond Hong Kong's geographical boundaries or exhibiting instantaneous speeds exceeding 150 km/h (considering Hong Kong's maximum speed limit of 110 km/h; Wang et al., 2017), a threshold designed to account for measurement error while identifying genuine data quality issues; 3) Temporal matching failures: observations that could not be temporally matched across the three data sources (GPS, sound meter, activity diary) at 1-min resolution, preventing reliable integration of spatiotemporal contexts with exposure measurements. These criteria ensure data quality while retaining 93.6% of recruited participants, demonstrating the robustness of our data collection protocol.

2.2. Identification of stay and move events and mobility status

Environmental exposures mainly vary with the changing spatiotemporal contexts of daily life. Stay and move points embody daily activity semantics, reflecting an individual's anchors or activity spaces (Liang et al., 2025). Stay and move events represent a fundamental decomposition of human daily mobility into discrete and meaningful episodes. Stay events capture periods when individuals remain at a location for a substantial duration, typically corresponding to purposeful activities such as working, dining, shopping, or resting at home. Move events represent transitions between stay locations, encompassing various transportation modes (walking, cycling, motorized travel). This event-based conceptualization offers three key advantages for exposure assessment. 1) it aligns exposure measurements with actual activity semantics, enabling investigation of how different activity types relate to environmental exposures; 2) it preserves the temporal sequence and duration of exposures, avoiding information loss inherent in simple spatial or temporal aggregation; and 3) it provides a natural unit of

Table 1
Sociodemographic profiles of valid participants.

Characteristics	Subgroups	Number (%)
Household income (HKD/month)	Less 20k	179 (23.90%)
	20k - 40k	293 (39.12 %)
	Over 40k	277 (36.98 %)
Age (years)	18 - 30	286 (38.18 %)
	31 - 45	286 (38.18 %)
	46 - 66	177 (23.63 %)
Gender	Female	508 (67.82 %)
	Male	241 (32.18 %)
Education level	Below bachelor's degree	249 (33.24 %)
	Bachelor's degree	326 (43.52 %)
	Master's/Doctor's degree	174 (23.23 %)
Employment status	The unemployed/retirees/homemakers	116 (15.49 %)
	Part-time employment/the self-employed	208 (27.77 %)
	Full-time employment	425 (56.74 %)
Neighborhood type	Private housing neighborhood	384 (51.27 %)
	Mixed neighborhood	234 (31.24 %)
	Public housing neighborhood	131 (17.49 %)

analysis that reflects how people structure their daily lives as sequences of activities and transitions (Wang et al., 2025). We thus chose to aggregate real-time exposures to the stay and move events level. Based on a heuristic spatiotemporal threshold method from Wang et al. (2025), we identify stay and move events and mobility status following the workflow in Fig. 3. In the first step, GPS trajectories are preprocessed. Recorded geographic coordinates may drift when GPS satellite signals are weak or lost. Referring to prior GPS denoising approaches, we first calculate instantaneous speeds between consecutive GPS points and remove those exceeding 150 km/h. GPS points outside Hong Kong's geographical boundaries are also excluded. Subsequently, a Kalman filter with a constant velocity model is applied to smooth the trajectories. In the second step, stay and move states are classified. Positions with a roaming distance (average pairwise distance) below 100 m and temporal proximity are grouped as potential stay points. Clusters with stay times exceeding 30 min are then designated as final stay points, with all positions replaced by the cluster centroid. The 30-min threshold follows established conventions in mobility research (Gong et al., 2012; Wang et al., 2025) and is designed to identify meaningful activity episodes (e.g., home, work, shopping, dining) rather than transient pauses (e.g., traffic signals, brief rest stops). While this threshold may exclude some shorter activities, it ensures our compound disadvantage indices primarily reflect exposures during substantive activities that contribute most to cumulative daily exposure. Conversely, GPS trajectories between two stay point clusters are treated as move events, regardless of brief stationary periods (<30 min) that may occur during movement. These brief pauses are incorporated into the overall move event rather than classified as separate stay points. In the third step, referencing methods from Gong et al. (2012) and Martin et al. (2023), mobility statuses are labeled based on the average speed of move events calculated across the entire trajectory segment between consecutive stay points: those below 6 km/h as low-speed movement (corresponding to walking), 6-15 km/h as moderate-speed movement (corresponding to cycling), and above 15 km/h as high-speed movement (corresponding to motorized travel). This classification approach ensures that brief stops during motorized travel do not result in misclassification as low-speed movement; instead, the overall average speed of the move event determines its mobility status. GPS trajectory processing was performed using the Python package MovingPandas (Graser, 2019). The method identified a total of 8427 stay and move events across all participants.

2.3. Measuring greenery exposure, noise exposure, and compound disadvantage indices

Greenery exposure. GPS trajectories and the normalized difference vegetation index (NDVI) derived from Sentinel-2 satellite imagery were used to calculate dynamic greenery exposure. The Sentinel-2 satellite imagery, provided by Google Earth Engine, has a spatial resolution of 10 m. We collected satellite imagery for Hong Kong during the survey period. Given Hong Kong's subtropical climate with limited seasonal variation in evergreen vegetation, we calculated NDVI using the 95th percentile across the entire survey period as background values to minimize cloud cover effects while capturing the stable greenery baseline. Before calculating the NDVI, the *s2cloudless* algorithm was applied to mitigate interferences from clouds and shadows (Skakun et al., 2022). The calculation formula for the NDVI is as follows:

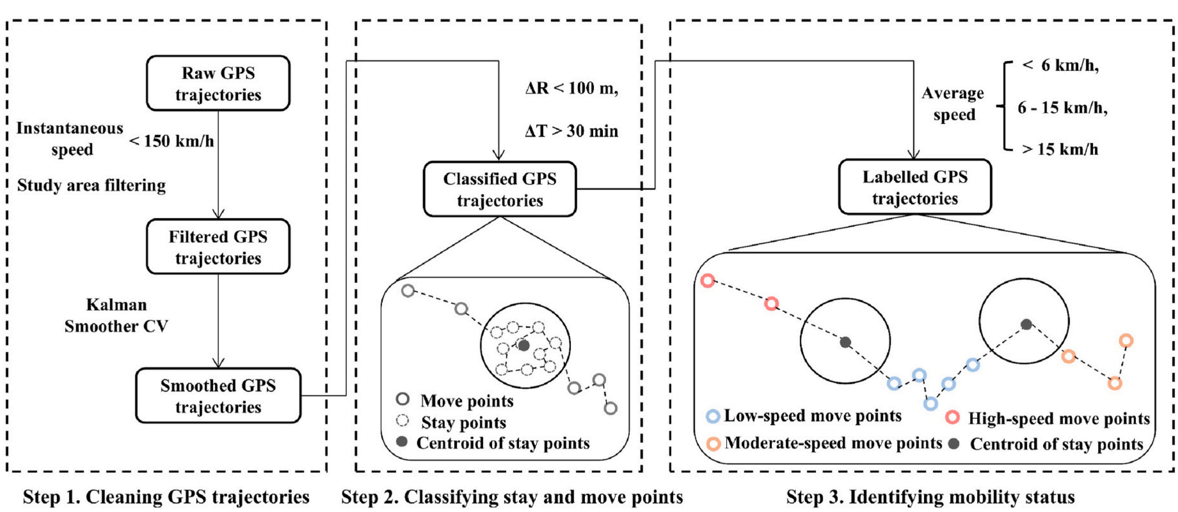

Fig. 3. Workflow of processing the GPS data.

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{1}$$

where NIR denotes the reflectance in the near-infrared band (band 8 in Sentinel-2) and RED denotes the reflectance in the visible red band (band 4 in Sentinel-2). The raw NDVI values range from −1 to 1, with negative values reassigned to 0 to eliminate interference from water bodies (Liu et al., 2025). Following the method of Liu et al. (2025), we selected a 300-m buffer around each GPS point and computed the NDVI to characterize greenery exposure. We also calculated the average exposure for each participant over the two-day survey period to compare residence-based and mobility-based measurements, thereby revealing the NEAP.

Noise exposure. We processed the raw equivalent continuous A-weighted sound pressure levels (LAeq) recorded by portable sound meters. Since LAeq cannot be linearly averaged, we aggregated the corresponding LAeq for each stay and move event using the following formula:

$$LAeq = 10\ \lg\left(\frac{1}{N}\sum\nolimits_{t}^{t_n} 10^{0.1LAeq_i}\right) \tag{2}$$

where N is the number of records within the time window for each stay and move event, $LAeq_i$ is the raw LAeq at 30-s resolution, t is the start time of each stay and move event, and t_n is the end time of each stay and move event. Similarly, we aggregated the corresponding LAeq for each participant during the survey period.

Compound disadvantage index. As defined in Section 1, compound disadvantage refers to the concurrent experience of high noise and low greenery exposure, representing the most detrimental environmental combination. To operationalize this concept, we calculated the proportion of time during which individuals simultaneously experienced high noise (LAeq >60 dBA) and low greenery (NDVI <0.4). Multiple lines of evidence support the selection of these specific thresholds. Prior research has demonstrated that LAeq exceeding 60 dBA and NDVI below 0.4 have significant negative impacts on environmental perception and mental well-being (Wang and Kwan, 2025; Zhang et al., 2020; Song et al., 2024). We further tested alternative threshold combinations. Specifically, more lenient thresholds (LAeq >55 dBA and NDVI <0.5) and more stringent thresholds (LAeq >65 dBA and NDVI <0.3). Five-fold cross-validation revealed that the original combination substantially outperformed both alternatives, providing strong empirical support for this threshold selection (see Supplementary Material A for detailed results). The calculation formula is:

$$CDI_{prop} = \frac{\sum \Delta t_i \cdot I(LAeq_i > 60 \wedge NDVI_i < 0.4)}{T} \tag{3}$$

where Δt_i is the duration of each sub-interval i, I is the indicator function (1 if the condition holds, 0 otherwise), and T is the total duration of the stay and move event.

2.4. Characterizing spatiotemporal contexts of daily mobility

Based on the works of Kou et al. (2021), Sun et al. (2024), Wang and Kwan (2025), and Yu et al. (2026), this study characterized the spatiotemporal contexts of daily activity in terms of four categories: activity-type contexts, activity-space function contexts, urban design contexts, and activity-temporality-state contexts. Table 2 provides a comprehensive mapping of these context categories to related terminology in the literature, along with specific variables and reference frameworks. First, for activity-type contexts, following the activity-space framework (Kwan, 1999; Kou et al., 2021; Sun et al., 2024), we derived the primary activity type for each stay and move event from the activity diaries. These types include work/study, housework, personal affairs, leisure, or travel. We matched stay and move events with activity diaries according to timestamps. However, discrepancies arose due to participants' recall biases or errors in identifying stay and move events (Wang et al., 2025). To mitigate these issues, we identified the primary activity as the mode (most frequent type) among those falling within the start and end times of each stay and move event. Second, based on the framework of land-use diversity (Leslie et al., 2007; Wang and Kwan, 2025), we utilized points of interest (POIs) data to characterize activity-space function contexts. For each stay and move event, we calculated the average proportions of various POI types within a 300-m buffer around the GPS trajectories. The POI data were sourced from Amap, a leading online mapping platform in Hong Kong. Based on the city's spatial function classifications, we made minor adjustments to the original POI categories, resulting in the following: transportation, residence, industry, company, shopping, restaurant, life service, education and culture, entertainment, sport and fitness, recreation and tourism, and healthcare.

Table 2
Spatiotemporal context categories for the established framework.

Context Categories	Related Terms	Variables	Framework
Activity-Type Contexts	Activity purpose	Primary activity type: work/study, housework, personal affairs, leisure, travel	Activity-space (Kwan, 1999; Kou et al., 2021; Sun et al., 2024)
Activity-Space Function Contexts	POI composition	Proportions of 12 POI categories: transportation, residence, industry, company, shopping, restaurant, life service, education & culture, entertainment, sport & fitness, recreation & tourism, healthcare	Land-use diversity (Leslie et al., 2007; Wang and Kwan, 2025)
Urban Design Contexts	Built environment 3Ds	Density (POI density), Diversity (Shannon-Wiener index based on POI composition), Design (tree height)	3Ds framework (Cervero and Kockelman, 1997)
Activity-Temporality-State Contexts	Temporal and mobility characteristics	Day type, Time of day (Time hour), Mobility status, Time inflexibility, Activity location type	Activity-space dynamics (Kwan, 2013; Yu et al., 2026)

Third, referring to the works of Cervero and Kockelman (1997) and Wang et al. (2026), we selected three indicators to represent urban design contexts: density, diversity, and tree height. For each stay and move event, we computed the average POI number within a 300-m buffer surrounding the GPS trajectories, the average Shannon-Wiener diversity index, and the average tree height. The tree height data were obtained from the ETH Global Canopy Height dataset (Lang et al., 2023), which integrated LiDAR data from Global Ecosystem Dynamics Investigation (GEDI) and Sentinel-2 data with a 10-m spatial resolution for the year 2020. Lastly, building on the framework of activity-space dynamics (Kwan, 2013; Yu et al., 2026), we characterized activity-temporality-state contexts using the following variables: day type (weekday or weekend), time of day (hours 0 - 23), mobility status (stay, low-speed movement, moderate-speed movement, or high-speed movement), time inflexibility (flexible, relatively fixed, and fixed), and activity location type (within the home neighborhood or not). Mobility status was determined as described in Section 2.2, while time inflexibility was extracted from the activity diaries. Activity location type was defined using a 500-m buffer around the participant's residence.

2.5. Modeling the nonlinear relationships between spatiotemporal contexts and compound disadvantages using GPBoost-SHAP

Gaussian process boosting (GPBoost). Given that stay and move events were nested within participants (violating independence assumptions), and given the need to capture complex nonlinear relationships between high-dimensional spatiotemporal contexts and compound disadvantage (as discussed in Section 1), this study employed GPBoost. GPBoost is an extension of tree-boosting models that combine tree-boosting and latent Gaussian models (Sigrist, 2022). GPBoost adheres to the framework of mixed effects models, utilizing gradient trees (LightGBM) for fixed effects and Gaussian processes for random effects. This approach provides several advantages, including relaxation of linearity assumptions, the capability to model dependencies in data with grouped structures. The model's formula is as follows:

$$y_i = F(\mathbf{x}_i) + b(\mathbf{z}_i) + \epsilon_i \tag{4}$$

where $F(\cdot)$ is the fixed effect function approximated by an ensemble of boosted decision trees (using LightGBM), with $\mathbf{x}_i$ encompassing variables characterizing SES and spatiotemporal contexts of daily mobility. $b(\cdot)$ is the random effect function modeled by a Gaussian process $b \sim \mathscr{GP}(0, k(\mathbf{z}, \mathbf{z}'))$ with covariance function k, and $\mathbf{z}_i$ denoting the grouping variable based on participant ID. $\epsilon_i \sim \mathscr{N}(0, \sigma^2)$ is independent and identically distributed noise. The compound disadvantage index, as described in Section 2.3, was utilized as the target variable. Both are continuous values, so mean squared error (MSE) was employed as the loss function. GPBoost computations were based on the *gpboost* Python package developed by Sigrist (2022), with hyperparameter tuning performed using the *optuna* Python package developed by Akiba et al. (2019). We have uploaded the codes and logs to the GitHub repository (https://github.com/GEO-AI-for-health/gpboost_compound_exposure_disadvantage).

Shapley additive explanation (SHAP). This study employed SHAP to interpret the fixed effects of the GPBoost model. SHAP is a game theory-based additive explanation method that decomposes model predictions into the marginal importances of individual variables using Shapley values (Shapley, 1953). We computed overall SHAP values, main effects, and pairwise interaction effects using the SHAP interaction framework implemented in Lundberg et al. (2020). The overall effect ϕ_i represents the overall marginal importance of variable i to the model prediction. The main effect $\phi_{i,i}$ isolates the independent importance of variable i after accounting for its interactions with other variables, while pairwise interaction effects $\phi_{i,j}$ ($i \neq j$) quantify the synergistic or antagonistic deviations from additivity between variables i and j. The formulas are as follows:

Overall effect (standard SHAP value):

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(M-|S|-1)!}{M!} \left[f(\mathbf{x}_{S \cup \{i\}}) - f(\mathbf{x}_S) \right] \tag{5}$$

Pairwise interaction effect ($i \neq j$):

$$\phi_{i,j} = \sum_{S \subseteq N \setminus \{i,j\}} \frac{|S|!(M-|S|-2)!}{2(M-1)!} \delta_{i,j}(S) \tag{6}$$

$$\delta_{i,j}(S) = f(\mathbf{x}_{S \cup \{i,j\}}) - f(\mathbf{x}_{S \cup \{i\}}) - f(\mathbf{x}_{S \cup \{j\}}) + f(\mathbf{x}_S) \tag{7}$$

Main effect:

$$\phi_{i,i} = \phi_i - \sum_{j \neq i} \phi_{i,j} \tag{8}$$

where N is the full set of variables with cardinality $M = |N|$, S is a subset of variables excluding the indexed ones, $f(\mathbf{x}_S)$ is the model's prediction using only the variables in S (with other variables masked to the background), and $\delta_{i,j}(S)$ is the second-order difference that isolates the pure interaction between i and j beyond their individual effects. In addition, we used a non-parametric bootstrap method based on residual resampling to estimate uncertainties in SHAP's main and interaction effects, similar to the approach taken by Li, 2022. This procedure was repeated 1000 times with the identical hyperparameters, and the 95% confidence intervals for the main effects were calculated using percentiles from the bootstrap sampling distribution. The codes have also been uploaded to the above GitHub repository.

3. Results

3.1. Measuring the results of noise and greenery exposures

Fig. 4 depicts the spatial distribution of noise and greenery exposures across 1,434,794 observations (1-min resolution), derived from 749 valid participants. The results show that, despite the participants residing in four areas, their trajectories span the entirety of Hong Kong's built-up areas. It indicates that residence-based exposure introduces measurement bias, whereas mobility-based exposure more accurately reflects people's actual noise exposure and greenery exposure. Fig. 4a shows that high noise exposure (80-114 dBA) exhibits network-like distributions, likely corresponding to major roads or transportation corridors dominated by traffic noise. Low-noise areas (26-60 dBA) are distributed across sparser branches or peripheral zones, possibly representing residential suburbs or quieter roads. Participants are more intensively exposed to high-noise environments at urban hubs during daily mobility, with potential hotspots concentrated at road intersections or infrastructure-dense zones. Fig. 4b reveals that high green exposure (0.6-0.8) mainly occurs in extensive patchy areas, possibly corresponding to parks, greenery belts, or suburban zones with abundant trees and open spaces. Lower exposure (0-0.4) dominates in urban cores with sparse vegetation. Compared to Fig. 4a, green exposure exhibits a more patchy and uneven distribution across the network, with peak concentrations in peripheral or dedicated recreational spaces, while urban cores show lower exposure levels. Moreover, noise exposure shows strong spatial heterogeneity: high noise exposure also correlates with high green exposure, and vice versa. It highlights the necessity for fine-grained analysis that considers individuals' daily mobility.

Fig. 5a shows differences in the noise-greenery relationship between residence-based and mobility-based exposure (aggregated to individuals). The findings indicate that the relationship is weak in both residence-based exposure (Spearman r = −0.094, p = 0.012) and mobility-based exposure (Spearman r = −0.067, p = 0.065). Formal statistical testing reveals that these two correlation coefficients do not differ significantly (paired bootstrap test: difference = 0.021, 95% CI = [−0.045, 0.089]). Fig. 5b and c shows the probability density functions of noise exposure and greenery exposure using residence-based and mobility-based measurements. The results clearly show that mobility-based measurements of both noise and greenery exposures show less deviation than residence-based ones, confirming the existence of the neighborhood effect averaging problem (NEAP). Moreover, greenery exposure exhibits downward averaging (i.e., high values of mobility-based exposure shifting toward the mean), while noise exposure shows upward averaging (low values shifting toward the mean). It indicates that compound disadvantages may be amplified by daily mobility, compared to those in home neighborhoods. Fig. 6 further illustrates the compound disadvantage patterns at the event level. Compound disadvantage is significantly higher across all daily events compared to events within home neighborhoods (t = 19.887 and 17.555, respectively, both p < 0.001). These findings confirm that exploring the spatiotemporal contexts of daily mobility and activity is crucial.

3.2. Models fit, random effects, and model explanatory power

To mitigate over-fitting, a 5-fold cross-validation was employed to tune the optimal hyperparameters for GPBoost (see Supplementary Material A for the best validation MSE and R^2 of each fold). We selected the hyperparameters (see the GitHub repository provided in Section 2.3 for details) corresponding to the fold with the highest R^2 to train the

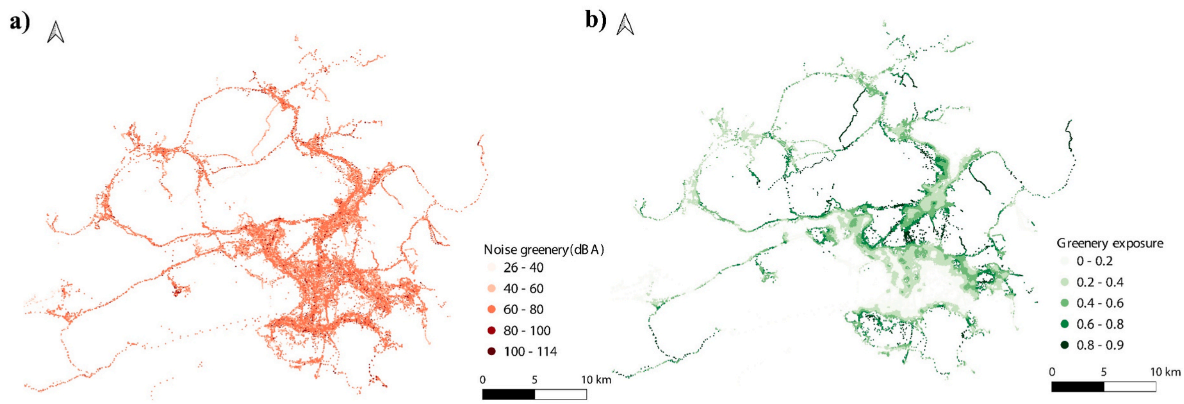

Fig. 4. Spatial visualization of noise (panel a) and greenery exposure (panel b, 1,434,794 observations).

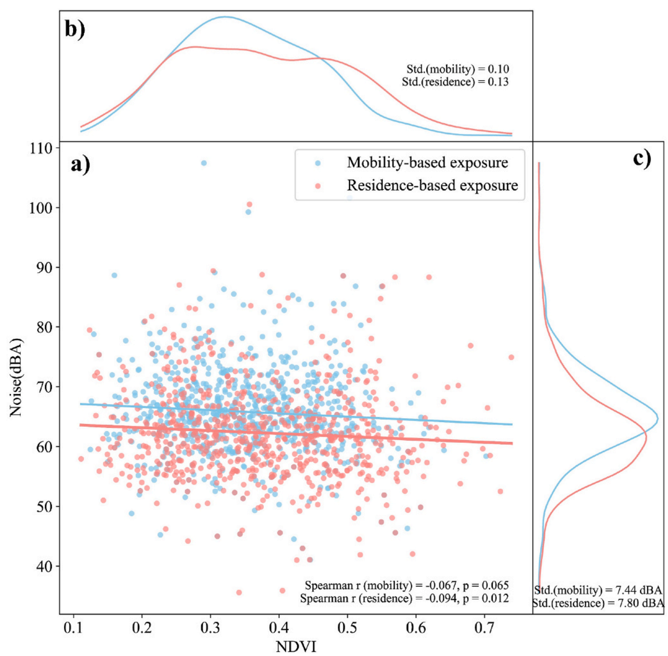

Fig. 5. The noise-greenery relationship (panel a), probability density function of greenery (panel b), and probability density function of noise (panel c), for residence-based exposure and mobility-based exposure (aggregated to individuals).

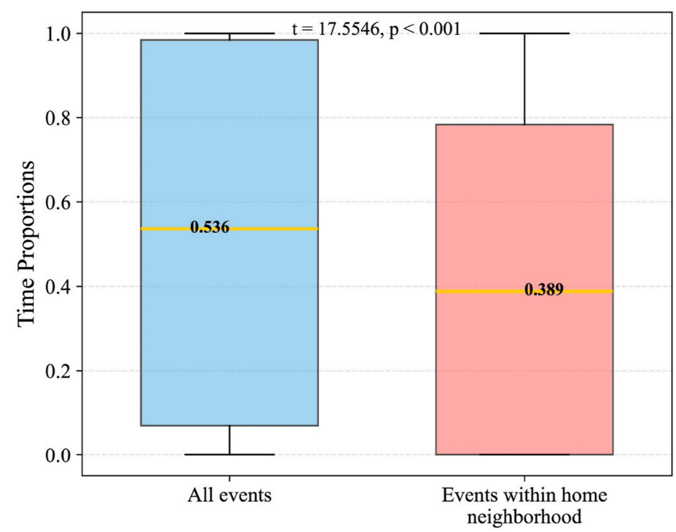

Fig. 6. The statistics and distribution of the compound disadvantage index (stay and move event level).

Table 3
Summary of the GPBoost models (using the entire dataset).

		Time proportions
Variance	Fixed effects	0.1076
	Random effects	0.0108
	Random errors	0.0366
ICC		0.2283
Marginal R^2		0.6943
Conditional R^2		0.7641

GPBoost model using the entire dataset. Table 3 presents the variances of the fixed effects, random effects, and random errors, along with the intraclass correlation coefficient (ICC), marginal R^2, and conditional R^2. The ICC for the models is 0.2283, indicating substantial random effects due to individual clustering, which should be accounted for in the models. The marginal R^2 is 0.6943, and the conditional R^2 is 0.7641. It means that SES and spatiotemporal contexts account for 69.43% of the variance in time proportions of compound exposure disadvantage; when the random effects of individual clustering are included, the full model accounts for 76.41% of the variance in time proportions.

3.3. The overall relationships between socioeconomic status, spatiotemporal contexts and compound disadvantages

Fig. 7 depicts the relative importance and overall relationships of variables with compound disadvantage. The findings show that spatio-temporal contexts exhibit stronger associations than SES variables. First, urban design contexts show the strongest associations, with tree height and POI density ranking in the top two positions and POI diversity ranking in the top ten. Notably, tree height and POI diversity are negatively associated with compound disadvantage, whereas POI density shows a positive association. Second, travel within activity-type contexts and the proportion of residence POIs within activity-space function contexts are both among the top five variables associated with compound disadvantages. Travel shows a positive association, whereas the proportion of residence POIs shows a negative association. Third, within activity-temporality-state contexts, time hour and activity location type (being in the home neighborhood or not) rank within the top five to six variables. Notably, the association of time hour appears potentially nonlinear, and being in the home neighborhood is associated with weaker compound disadvantage, consistent with the findings in Section 3.1. Within SES, employment status is the most important variable. Individuals with full-time employment are associated with less

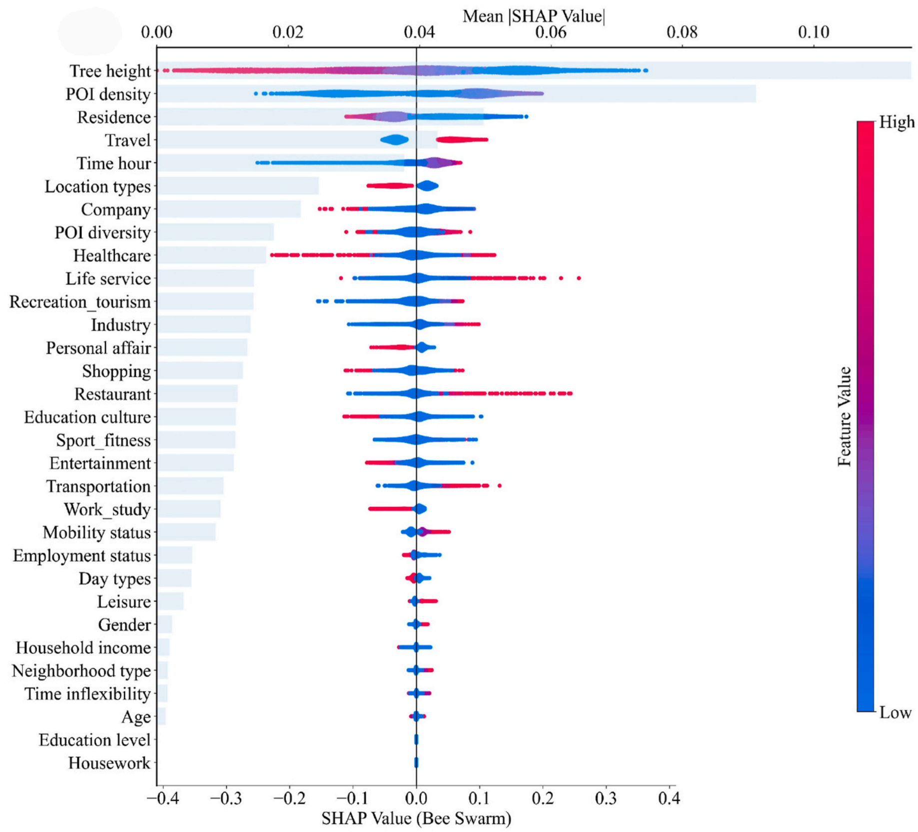

Fig. 7. SHAP summary plots.

compound exposure disadvantage.

To examine the temporal patterns of compound disadvantages along the stay and move events, we included the compound disadvantage indicator from the previous five stay and move events in the set of predictors for GPBoost training and used SHAP to quantify their associations (Fig. 8), noting that the main model does not include lagged CDI variables. The findings show that compound disadvantage at the current event is most strongly associated with the immediately previous event, with association strength decreasing substantially for events two or more steps earlier. This pattern suggests that compound disadvantage exhibits strong short-term predictive relationships within activity sequences, with the most recent event being the strongest predictor of current exposure conditions.

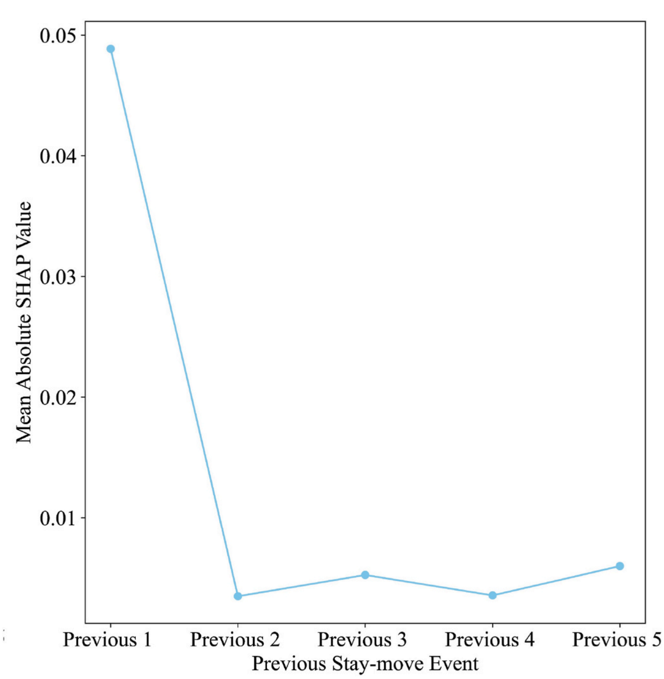

Fig. 8. SHAP-based importance of the compound disadvantage index for the previous five stay and move events.

3.4. The main relationships between socioeconomic status, spatiotemporal contexts and compound disadvantages

This section addresses RQ3 by examining which specific spatiotemporal contexts show significant nonlinear associations with compound disadvantage. To focus on the independent relationships of each variable, we used SHAP main effects rather than overall SHAP values for the partial dependence analysis. As detailed in Section 2.5 (Equation (9)), main effects isolate each variable's association after removing its pairwise interactions with other variables. We report all variables showing statistically significant associations at the 95% confidence level based on bootstrap confidence intervals (1000 iterations). Variables without significant main effects are presented in Supplementary Material B. Fig. 9a shows the partial dependence between SES and compound exposure disadvantage. Note that given the substantial computational cost, we randomly sampled 3000 observations for display. The findings indicate that employment status is the only SES variable that is significantly associated with compound exposure disadvantage at the 95% confidence level. Higher levels of employment participation are associated with a lower compound disadvantage. Fig. 9b presents the partial dependence between activity-type contexts and compound exposure disadvantage. When participants are working or studying, as well as dealing with personal affairs, their levels of compound exposure disadvantage are associated with significantly lower levels. By contrast,

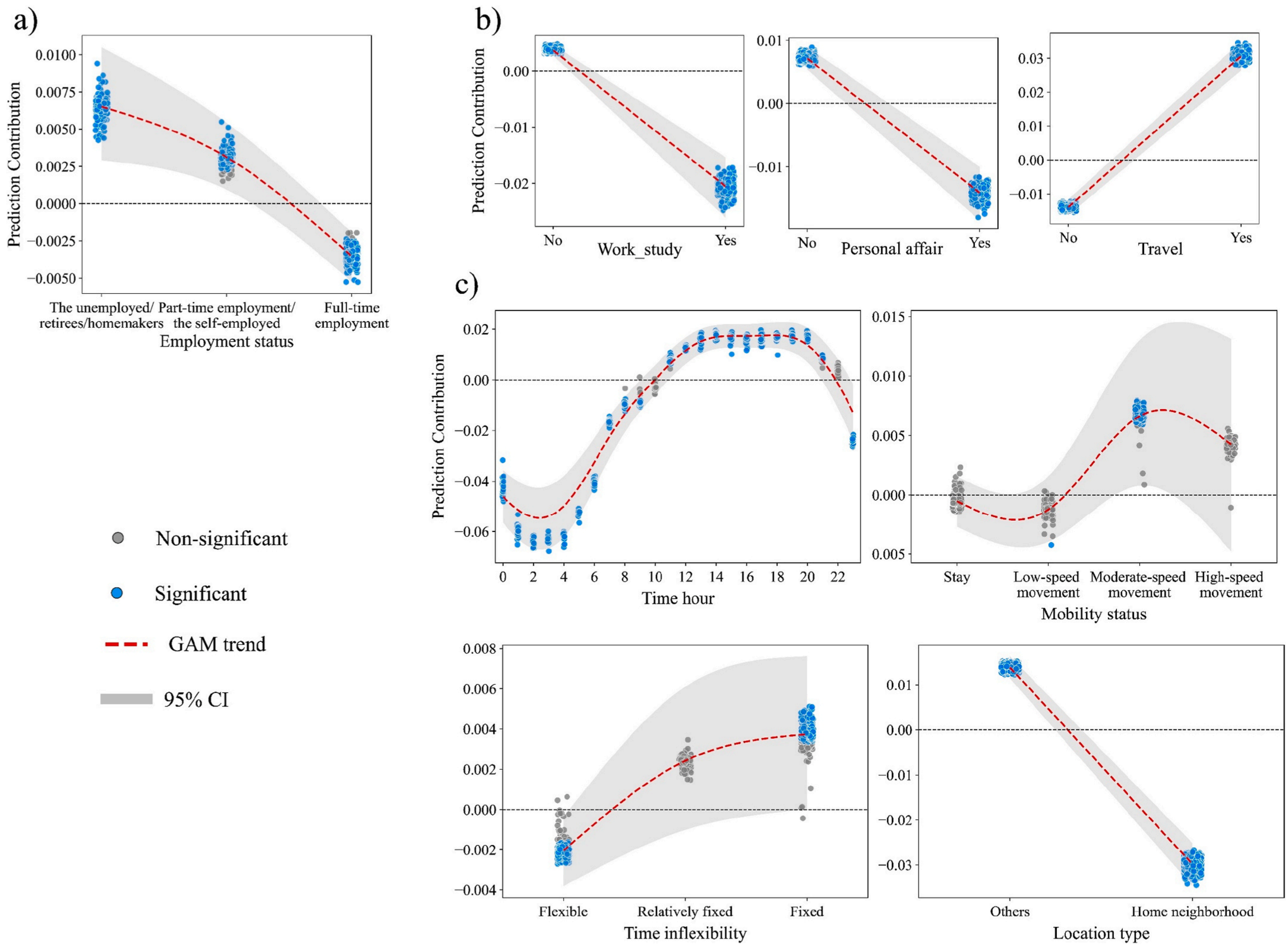

Fig. 9. SHAP-based partial dependence of socioeconomic status (panel a), activity-type contexts (panel b), and activity-temporality-state contexts (panel c) on compound exposure disadvantages, with 95% confidence intervals and generalized additive model (GAM) trend lines.

during travel, the exposure level is significantly higher. Fig. 9c illustrates the partial dependence relationships between activity-temporality-state contexts and compound exposure disadvantage. The partial dependence plots indicate that individuals encounter significantly higher levels of compound disadvantages from 11:00 to 21:00, with the lowest levels observed around 2:00 a.m. In terms of mobility status, moderate-speed movement (corresponding to cycling speeds) is associated with the most severe compound disadvantages, significantly exceeding those of other mobility statuses. The lowest levels occur during stay or low-speed movement (corresponding to walking speeds), followed by high-speed movement (corresponding to motorized travel). In addition, time inflexibility in daily mobility shows a significant association with compound disadvantages. Activities with more rigid schedules are linked to higher compound disadvantages, with fixed-schedule activities exhibiting significantly higher levels compared to others. Lastly, for activity location type, participants face significantly greater compound disadvantages when conducting activities outside their home neighborhoods.

Fig. 10a presents the partial dependence between activity-space function contexts and compound exposure disadvantages. Higher proportions of transportation, industry, and restaurant POIs in participants' activity spaces are associated with significantly elevated levels of experienced compound exposure disadvantage. Specifically, when the proportion of transportation POIs exceeds 20%, industry POIs over 10%, and restaurant POIs exceed approximately 30%, there is a rapid increase in individuals' compound exposure disadvantage. In contrast, higher proportions of residence, education and culture, and healthcare POIs in participants' activity spaces are associated with significantly lower levels of such exposure among participants. Proportions of residence POIs higher than 10% can reduce compound exposure disadvantage, with the best benefits observed at around 25%. For education, culture, and healthcare POIs, a proportion of 20% begins to reduce compound disadvantages, with marginal effects peaking at around 40%. The proportion of company POIs follows an inverted U-shaped curve, with an initial increase followed by a decrease, with intensified compound exposure disadvantage occurring within the 20% - 40% range. Fig. 10b shows the partial dependence between urban design contexts and compound exposure disadvantage. When the number of POIs within individuals' daily spaces exceeds 500 (within a 300 m buffer), it increases their compound exposure disadvantage. When the number of POIs is higher than about 2,000, the incremental disadvantage levels tend to plateau. A higher POI diversity is associated with a lower disadvantage level. When the Shannon index exceeds 2, the exposure disadvantage level declines markedly. For tree height, once the average height exceeds 5 m, the disadvantage level decreases significantly, with the marginal benefits leveling off with tree height higher than 8 m.

3.5. *The pairwise interactions between socioeconomic status, spatiotemporal contexts and compound disadvantages*

This section continues addressing RQ3 by examining whether and

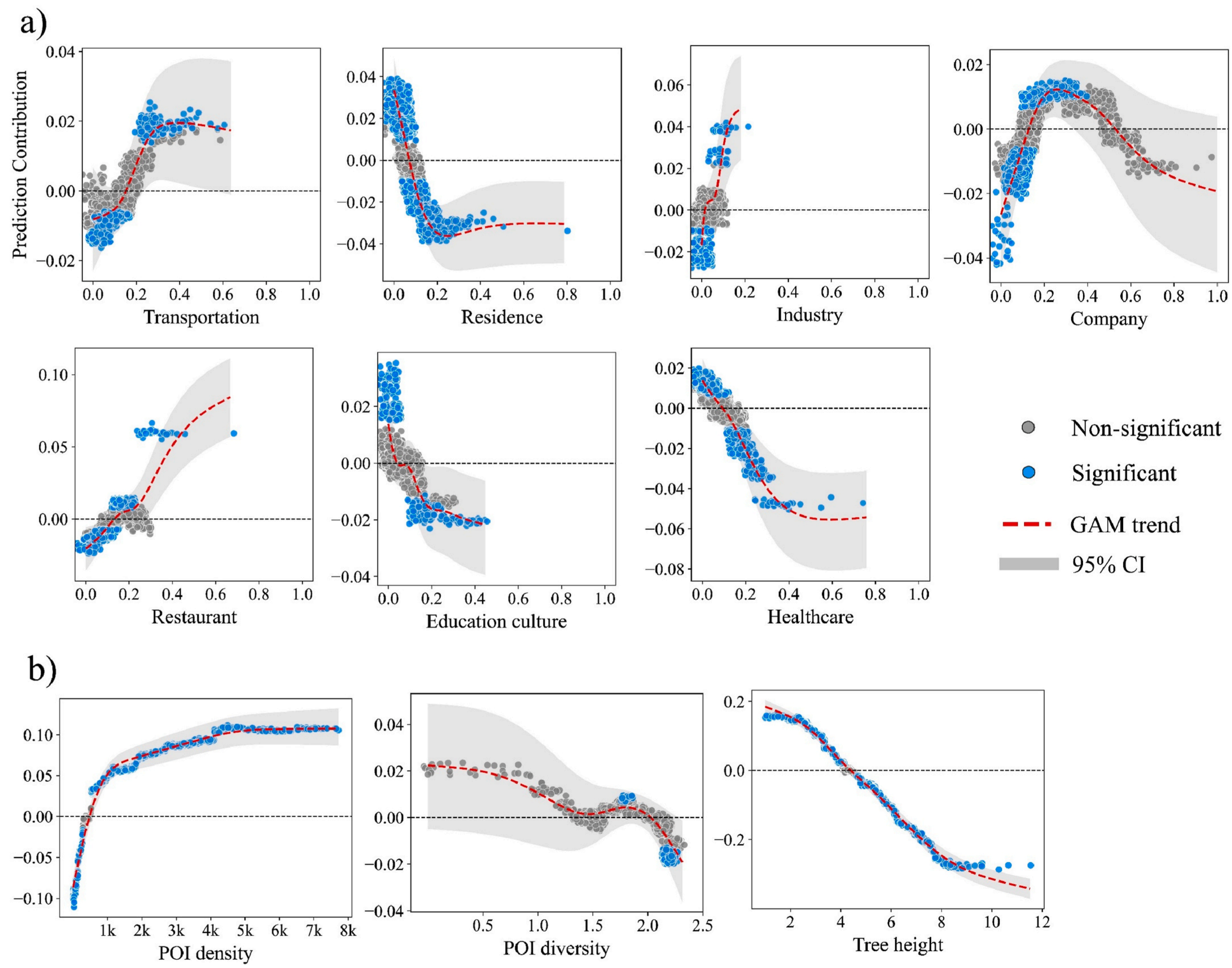

Fig. 10. SHAP-based partial dependence of activity-space function contexts (panel a) and urban design (panel b) on compound exposure disadvantages, with 95% confidence intervals and generalized additive model (GAM) trend lines.

how SES interacts with spatiotemporal contexts in shaping compound disadvantage patterns. We report all pairwise interactions showing pronounced patterns. Due to space constraints, the complete interaction matrix is available in our GitHub repository. By isolating the main effects, we investigated how SES interacts with the spatiotemporal contexts of daily mobility to predict compound disadvantage levels. Fig. 11 highlights the spatiotemporal contexts showing pronounced interaction patterns with SES. The findings show that, apart from the main effects, individuals from different income groups suffer varying levels of compound exposure disadvantage during work or study activities, with high-income groups experiencing lower disadvantage levels. The study reveals subtle differences in the variation of compound disadvantages across different employment statuses throughout the day. Compared to other employment statuses, full-time employees experienced lower compound disadvantages during the midday and afternoon periods. Gender also interacts with time of day and activity location types. Women experience markedly greater levels of compound disadvantage than men from midday to evening (12:00 - 23:00) and in spaces outside of their home neighborhoods.

Variables within spatiotemporal contexts also have pronounced interaction patterns. First, activity-type contexts interact with activity-temporality-state contexts, predicting compound disadvantage levels. The findings indicate that participants experience higher compound disadvantage levels during work or study activities when the location is in their home neighborhood compared to other location types. Moreover, participants experience higher compound disadvantage levels during travel activities in the morning period (7:00 - 10:00). Second, activity-temporality-state contexts interact with urban design contexts to predict compound disadvantage levels. In activity spaces outside the home neighborhood, greater tree height appears to more effectively reduce experienced compound disadvantages. Third, there are notable interactions between variables in the activity-temporality-state contexts. When activities are conducted within the home neighborhood, compound disadvantage levels are lower during early morning hours, yet higher during daytime and evening periods compared to other activity location types.

4. Discussion

4.1. Strength of the methodology

This study proposes a novel approach to explore the relationships between spatiotemporal contexts and compound exposure disadvantages from the perspective of people's daily mobility and activity. Despite substantial progress in measuring noise and greenery exposures (Collins et al., 2020; Hasegawa and Lau, 2022; James et al., 2016), most

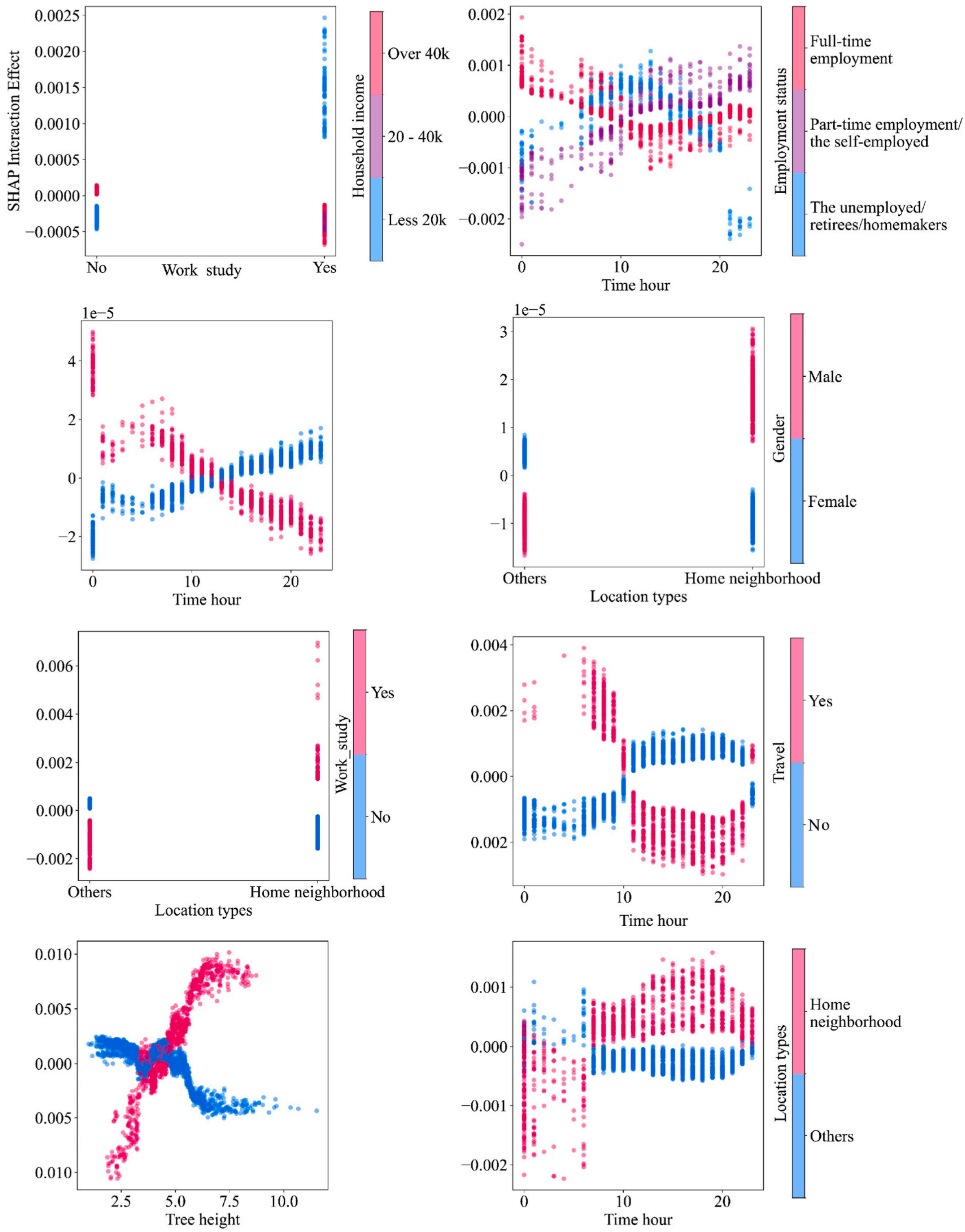

Fig. 11. SHAP-based interaction effects of socioeconomic status and spatiotemporal contexts on compound disadvantages.

studies still employed static measurement approaches (i.e., residence-based measurements). A few recent studies have applied spatiotemporal weighting or sampling methods to real-time data (Cai et al., 2023; Yu et al., 2024). Nevertheless, these efforts have not fully leveraged the potential of spatiotemporal information. Further, past studies have largely focused on isolated exposures to noise and greenery, overlooking their complex combinations at the individual level (Kan et al., 2023; Liu et al., 2023). This study represents the first attempt to employ an explainable machine learning framework that accounts for the dependence of time-sequence data to infer levels of compound exposure disadvantage from people's daily experienced spatiotemporal contexts.

The dynamic mobility-based measurement confirms the presence of the neighborhood effect averaging problem (NEAP) in both greenery and noise exposures. Noise exposure mostly demonstrates upward averaging, leading to residence-based estimates that underestimate daily noise exposure, a conclusion corroborated by Ma et al. (2020). Greenery exposure is marked by downward averaging, resulting in residence-based evaluations that overstate daily greenery exposure. Thus, residence-based measurements fail to accurately capture the extent of compound disadvantage experienced by individuals, as daily activities conducted outside the home neighborhood present greater levels of disadvantage. It indicates that residential spaces in Hong Kong are generally quieter (e.g., located away from major thoroughfares) or have localized green space, whereas external zones (e.g., commercial or transportation hubs) have higher noise levels and less greenery. For example, bustling districts like Central or Tsim Sha Tsui feature parks but are mostly dominated by traffic noise. Further, the findings show that noise and greenery exposures have a weaker negative relationship in mobility-based measurements than in residence-based measurements, implying that real compound exposure patterns are more complex. To further elucidate the compound disadvantage patterns from the perspective of daily mobility, this study employs a spatiotemporal threshold method to characterize individuals' daily activities as sequences of stay and move events. Based on these events, an indicator for computing compound disadvantages is proposed. Subsequently, an explainable machine learning framework utilizing GPBoost-SHAP is applied to uncover nonlinear relationships between spatiotemporal contexts and compound disadvantages. The GPBoost modeling results reveal that SES and spatiotemporal contexts account for 69.43% of the variance in compound exposure disadvantages.

4.2. Employment status as the predominant socioeconomic determinant of compound exposure disadvantages

Residual-based bootstraps and SHAP analyses indicate that employment status is the primary SES variable predicting compound exposure disadvantages. This may be elucidated through the mobility capital theory proposed by Kaufmann et al. (2004), indicating that daily mobility aligns with this framework. Full-time employees typically possess greater mobility capital, enabling them to proactively select or adapt to exposure conditions. The findings also show that activities outside the home neighborhood may increase disadvantage, but full-time employees frequently commute to environmentally optimized workplaces, such as high-end offices in central business districts, which may have noise barrier facilities or proximity to green spaces (e.g., Hong Kong Park), lowering overall compound exposure disadvantage levels. In contrast, the unemployed, retirees, and homemakers have more restricted mobility ranges, typically confined to the home neighborhood. Although the results show that disadvantages in the home neighborhood are generally lower, prolonged stays among these groups may amplify other factors, such as cumulative exposure due to a lack of activity diversity, or socioeconomic vulnerabilities (e.g., residing in high-density or main road-adjacent housing estates) that indirectly increase localized noise (e.g., neighborhood disturbances or traffic noise). A study on environmental inequalities in Hong Kong reveals that low-mobility groups tend to lack access to superior outdoor green spaces, resulting in an overall disadvantage (Marmot et al., 2024). Regarding activity type contexts, the results further corroborate that individuals experience significantly lower compound exposure disadvantages during work or study activities. Interaction effects also reveal that non-full-time employees encounter higher levels of compound exposure disadvantages during work periods (midday and afternoon). Assuming exposure to spaces with identical proportions of industry POIs, full-time employees exhibit lower compound exposure disadvantage levels, implying superior work environments for them.

4.3. Implications for individuals' daily activity arrangements

During travel activities, residents experience significantly higher levels of compound exposure disadvantages. Moreover, mobility status exhibits a pronounced nonlinear association with disadvantage levels. The lowest compound disadvantage level is at walking speeds, followed by motorized speeds, and the greatest disadvantage level is at cycling speeds. This pattern is consistent with Hong Kong's particular urban environment, in which pedestrian movement takes place mostly on enclosed footbridges that effectively block noise. Motorized vehicles also provide a somewhat enclosed setting. Interaction effects suggest that, in high-density spaces, individuals engaged in high-speed mobility experience fewer compound exposure disadvantages. In addition, travel activities conducted between 7:00 and 10:00 are associated with elevated levels of compound exposure disadvantages. This association may reflect the period's coincidence with the morning rush hour, during which traffic volumes surge, leading to significantly higher road noise, while typical travel routes often traverse urban roads or highways with minimal greenery.

The results indicate that residents experience significantly lower levels of compound exposure disadvantages during work or study activities. However, the type of work location is strongly associated with disadvantage levels. The interaction effects reveal that work or study at home appear to involve higher levels of such disadvantages. This phenomenon arises from the fact that work or study at home usually confines people to residential spaces for prolonged periods. While these spaces generally have lower overall noise levels (e.g., being away from major thoroughfares), they lack the specialized acoustic optimization that commercial office spaces typically feature. At the same time, these spaces often lack indoor green elements, such as views of outside spaces or indoor plants. Further, the lack of commuting-related mobility, which may limit encounters with outside greenery, contributes to the disadvantages. In Hong Kong, urban planning promotes green standards for commercial buildings (e.g., Leadership in Energy and Environmental Design (LEED)-certified workplaces), while there is substantial variety in noise mitigation and vegetation integration throughout residential neighborhoods. The study also identifies day and evening hours (11:00 - 21:00) as times when Hong Kong residents experience significant compound exposure disadvantages. The interaction effects also show that, during this period, disadvantages are more severe when activities are conducted in home neighborhoods than in other locations. This suggests that residents, particularly those in high-density neighborhoods, may benefit from conducting midday activities (approximately 11:00-15:00) in locations outside their home neighborhoods when feasible, such as shopping centers, workplaces, or neighborhoods with better environmental infrastructure.

4.4. Implications for urban spatial function planning

The findings also offer scientific guidance for urban spatial function planning. Based on the main effects from residual-based bootstraps and SHAP analyses, spatial density of POIs increases individuals' compound exposure disadvantages, whereas POI diversity reduces them, with evident threshold effects. When using a 300-m buffer radius as the fundamental spatial unit, the number of POIs should not exceed 500, because exceeding this threshold dramatically worsens compound exposure disadvantages. Similarly, the Shannon index determined using the POI type classifications in this study implies that the diversity levels should be more than 2. Tree height design also has optimal values, as the main effects indicate that average heights exceeding 5 m substantially reduce compound exposure disadvantages, with benefits plateauing above 8 m. Further, interaction effects show that higher trees, when placed outside of the home neighborhoods, provide greater benefits. Regarding ideal proportions for spatial function types (various POI categories), the main effects provide valuable insights. Specifically, the proportion of transportation POIs should not exceed 20%, industry POIs should not surpass 10%, and restaurant POIs should not exceed approximately 30%, as exceeding these leads to a rapid escalation in individuals' compound exposure disadvantages. Proportions of residential POIs should range from 10% to 25%. For education, culture, and healthcare POIs, a proportion of 20% begins to alleviate compound disadvantages, with marginal effects reaching their peak at around 40%. The proportion of company POIs should be less than 20% or greater than

40%. It should be noted that the thresholds identified above are derived from SHAP partial dependence plots based on observed associations in a cross-sectional study. They should be interpreted as descriptive patterns and hypothesis-generating insights rather than prescriptive planning standards.

4.5. Limitations and future work

Nonetheless, this study has shortcomings that offer avenues for future investigations. First, our stay and move events identification method uses a 30-min duration threshold to distinguish stay points from move events. While this threshold follows established conventions in mobility research and ensures a focus on meaningful activity episodes, it may exclude some shorter but potentially relevant activities. Second, the sample is approximately two-thirds female, which exceeds the Hong Kong population average. This imbalance may affect the generalizability of the findings, particularly those related to employment status and home-neighborhood exposure patterns, as female participants in Hong Kong tend to have higher rates of part-time employment, caregiving activities, and home-based routines, which could influence the observed relationships between mobility contexts and compound disadvantage. Future studies should aim for more balanced gender representation or apply post-stratification weighting to improve representativeness. Third, it focuses on overall greenery exposure, without accounting for more nuanced indicators such as greenery quality, morphology, and visibility. A few studies have preliminarily demonstrated that greenery quality and visibility appear to be critical (Zheng et al., 2025; Wang et al., 2026), suggesting that future investigations could expand the scope accordingly. Similarly, this study examines only the equivalent continuous A-weighted sound pressure levels (LAeq) for noise, whereas exposures to sound fluctuations, source diversity, and frequency patterns are also highly significant (Cai et al., 2023; Wang et al., 2024). Additionally, our use of POIs to characterize activity-space function contexts offers limited practical guidance for urban planning. Future work could substitute these with data more closely aligned with real-world planning practices. Finally, the seasonal and spatial ranges of the samples are somewhat limited. This study focuses on how daily spatiotemporal contexts relate to compound disadvantages rather than seasonal variations. While we controlled for day type and time of day, and Hong Kong exhibits limited seasonal greenery variation, we acknowledge that seasonal differences in exposures and activity patterns may exist. Particularly, while Hong Kong's subtropical climate reduces the magnitude of seasonal NDVI variation relative to temperate settings, the potential for seasonal overestimation remains. We also suggest that future studies could use time-matched NDVI values corresponding to each participant's actual survey season as a more temporally precise alternative. While this study used stratified sampling to recruit 800 individuals as broadly as feasible, the spatial ranges and location types they visited during the survey period remain constrained.

5. Conclusion

This study integrates data from mobile sensors, GPS, satellite imagery, POIs, and activity diaries to measure individuals' real-time noise exposure, greenery exposure, and spatiotemporal contexts of daily mobility. Methodologically, we employ a spatiotemporal threshold method to segment daily mobility into stay and move events and construct a compound disadvantage index at the event level. This event-based approach enables us to examine how compound exposures vary across different types of activities and mobility contexts, providing new insights into the temporal dynamics of compound disadvantage in daily life. Further, the study also represents the first effort to propose an analytical approach that integrates GPBoost, residual-based bootstraps, and SHAP to estimate the dependence and uncertainty in time-sequence data, thereby exploring nonlinear relationships between spatiotemporal contexts and compound exposure disadvantages. The main findings are as follows. First, comparisons between mobility-based and residence-based measurements confirm the presence of the NEAP for both noise and greenery exposures. Second, GPBoost modeling reveals that SES and spatiotemporal contexts collectively explain 69.43% of the variance in compound exposure disadvantages. Thus, individuals' compound exposure disadvantages in noise and greenery can be inferred from SES and spatiotemporal information. Third, employment status emerges as the primary SES variable predicting compound exposure disadvantages. Full-time employees experience lower disadvantages due to greater mobility capital and access to environmentally optimized workplaces, while unemployed individuals, retirees, and homemakers face higher disadvantages due to restricted mobility ranges and prolonged exposure in potentially suboptimal home neighborhoods. Fourth, the combined residual-based bootstrap and SHAP framework reveals several critical threshold effects and interaction patterns: POI density should not exceed 500 within 300-m buffers, tree heights should be above 5 m, and transportation/industry/restaurant POIs should remain below 20%/10%/30% respectively. The analyses also uncover a nonlinear mobility status effect, with cycling speeds associated with the highest disadvantages due to open exposure, while enclosed pedestrian footbridges and motorized vehicles provide protection. Pronounced SES interactions show that non-full-time employees experience significantly higher disadvantages during work periods (especially midday) and in high-density home neighborhoods. These specific findings offer actionable implications for individuals' daily activity arrangements and for urban spatial function planning.

Informed consent statement

Informed consent was obtained from all participants before data were collected from them.

Ethics approval statements

The survey was reviewed and approved by the Survey and Behavioral Research Ethics Committee of the Chinese University of Hong Kong (Protocol SBRE-19-123 approved on 8 January 2020; Protocol SBRE(R)-21-005 approved on 1 November 2021).

CRediT authorship contribution statement

Linsen Wang: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Mei-Po Kwan:** Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no financial and personal relationships with other people or organizations that can inappropriately influence our work, and there is no professional or other personal interest of any nature or kind in any product, service, and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.

Acknowledgments

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), a grant from the 1 + 1+1 CUHK-CUHK(SZ)-GDSTC Joint Collaboration Fund (4760974, 2025A0505000062), an Innovation and Technology Fund (ITF) (ITS/

018/24SC), 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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2026.119254.

Data availability

The data that has been used is confidential.

References

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: a next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2623–2631.
Basner, M., Babisch, W., Davis, A., Brink, M., Clark, C., Janssen, S., Stansfeld, S., 2014. Auditory and non-auditory effects of noise on health. The lancet 383 (9925), 1325–1332.
Cai, J., Kwan, M.P., Kan, Z., Huang, J., 2023. Perceiving noise in daily life: how real-time sound characteristics affect personal momentary noise annoyance in various activity microenvironments and times of day. Health Place 83, 103053.
Casey, J.A., Morello-Frosch, R., Mennitt, D.J., Fristrup, K., Ogburn, E.L., James, P., 2017. Race/ethnicity, socioeconomic status, residential segregation, and spatial variation in noise exposure in the contiguous United States. Environ. Health Perspect. 125 (7), 077017.
Cervero, R., Kockelman, K., 1997. Travel demand and the 3Ds: density, diversity, and design. Transp. Res., Part D Transp. Environ. 2 (3), 199–219.
Chau, C.K., Leung, T.M., Xu, J.M., Tang, S.K., 2018. Modelling noise annoyance responses to combined sound sources and views of sea, road traffic, and mountain greenery. J. Acoust. Soc. Am. 144 (6), 3503–3513.
Chu, S., Xu, W., Zhang, D., Lin, J., Liu, J., Liu, S., Hong, X.C., 2025. Urban Blue-green spaces and tranquility: a comprehensive review of noise reduction and sensory perception integration. J. Asian Architect. Build Eng. 1–22.
Collins, T.W., Nadybal, S., Grineski, S.E., 2020. Sonic injustice: disparate residential exposures to transport noise from road and aviation sources in the continental United States. J. Transport Geogr. 82, 102604.
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. Soc. Sci. Med. 94, 26–33.
De Vries, S., Verheij, R.A., Groenewegen, P.P., Spreeuwenberg, P., 2003. Natural environments—healthy environments? An exploratory analysis of the relationship between greenspace and health. Environ. Plann. 35 (10), 1717–1731.
Dzhambov, A.M., Dimitrova, D.D., 2015. Green spaces and environmental noise perception. Urban For. Urban Green. 14 (4), 1000–1008.
Gascon, M., Sánchez-Benavides, G., Dadvand, P., Martínez, D., Gramunt, N., Gotsens, X., et al., 2018. Long-term exposure to residential green and blue spaces and anxiety and depression in adults: a cross-sectional study. Environ. Res. 162, 231–239.
Gong, H., Chen, C., Bialostozky, E., Lawson, C.T., 2012. A GPS/GIS method for travel mode detection in New York City. Comput. Environ. Urban Syst. 36 (2), 131–139.
Graser, A., 2019. Movingpandas: efficient structures for movement data in python. GIForum 1, 54–68.
Hartig, T., Evans, G.W., Jamner, L.D., Davis, D.S., Gärling, T., 2003. Tracking restoration in natural and urban field settings. J. Environ. Psychol. 23 (2), 109–123.
Hartig, T., Mitchell, R., De Vries, S., Frumkin, H., 2014. Nature and health. Annu. Rev. Publ. Health 35 (1), 207–228.
Hasegawa, Y., Lau, S.K., 2022. Comprehensive audio-visual environmental effects on residential soundscapes and satisfaction: partial least square structural equation modeling approach. Landsc. Urban Plann. 220, 104351.
Horoshenkov, K.V., Khan, A., Benkreira, H., 2013. Acoustic properties of low growing plants. J. Acoust. Soc. Am. 133 (5), 2554–2565.
James, P., Hart, J.E., Banay, R.F., Laden, F., 2016. Exposure to greenness and mortality in a nationwide prospective cohort study of women. Environ. Health Perspect. 124 (9), 1344–1352.
Jeon, J.Y., Jo, H.I., 2020. Effects of audio-visual interactions on soundscape and landscape perception and their influence on satisfaction with the urban environment. Build. Environ. 169, 106544.
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.
Kaufmann, V., Bergman, M.M., Joye, D., 2004. Motility: mobility as capital. Int. J. Urban Reg. Res. 28 (4), 745–756.
Kim, J., Kwan, M.P., 2021. 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.
Klompmaker, J.O., Janssen, N.A., Bloemsma, L.D., Gehring, U., Wijga, A.H., van den Brink, C., et al., 2019. Associations of combined exposures to surrounding green, air pollution, and road traffic noise with cardiometabolic diseases. Environ. Health Perspect. 127 (8), 087003.
Kou, L., Kwan, M.P., Chai, Y., 2021. The effects of activity-related contexts on individual sound exposures: a time–geographic approach to soundscape studies. Environ. Plan. B Urban Anal. City Sci. 48 (7), 2073–2092.
Kwan, M.-P., 1999. Gender and individual access to urban opportunities: a study using space-time measures. Prof. Geogr. 51 (2), 211–227.
Kwan, M.P., 2012. The uncertain geographic context problem. Ann. Assoc. Am. Geogr. 102 (5), 958–968.
Kwan, M.P., 2013. Beyond space (as we knew it): toward temporally integrated geographies of segregation, health, and accessibility: Space–time integration in geography and GIScience. Ann. Assoc. Am. Geogr. 103 (5), 1078–1086.
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.
Lang, N., Jetz, W., Schindler, K., Wegner, J.D., 2023. A high-resolution canopy height model of the Earth. Nat. Ecol. Evol. 7 (11), 1778–1789.
Leslie, E., Coffee, N., Frank, L., Owen, N., Bauman, A., Hugo, G., 2007. Walkability of local communities: using geographic information systems to objectively assess relevant environmental attributes. Health Place 13 (1), 111–122.
Li, Z., 2022. Extracting spatial effects from machine learning model using local interpretation method: an example of SHAP and XGBoost. Comput. Environ. Urban Syst. 96, 101845.
Liang, S., Li, Q., Zhuo, L., Zou, D., Xu, Y., Zhou, S., 2025. Improving next location prediction with inferred activity semantics in mobile phone data. Int. J. Digit. Earth 18 (2), 2552880.
Liu, D., Kwan, M.P., Kan, Z., Liu, Y., 2023. 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., Wang, J., 2025. Analytically articulating the effect of buffer size on urban green space exposure measures. Int. J. Geogr. Inf. Sci. 39 (2), 255–276.
Liu, Y., Kwan, M.P., Wang, J., Cai, J., 2024. Confounding associations between green space and outdoor artificial light at night: systematic investigations and implications for urban health. Environ. Sci. Ecotechnol. 21, 100436.
Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., et al., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2 (1), 56–67.
Ma, J., Li, C., Kwan, M.P., Kou, L., Chai, Y., 2020. Assessing personal noise exposure and its relationship with mental health in Beijing based on individuals' space-time behavior. Environ. Int. 139, 105737.
Marmot, M., Castedo, A., Allen, J., Goldblatt, P., Woo, J., Yeoh, E.K., Wong, H., Chung, R., Lai, E., Chung, G., Kwan, M., Lee, R., 2024. The Environment and Health Inequalities in Hong Kong. Institute of Health Equity, London.
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.
Ministry of Ecology and Environment of China, 2020. The 2020 report on environmental noise pollution prevention and control in China. Beijing, China: Ministry of Ecology and Environment. Retrieved from. https://www.mee.gov.cn/hjzl/sthjzk/hjzywr/202006/t20200619_785111.shtml.
Mitchell, R., Popham, F., 2008. Effect of exposure to natural environment on health inequalities: an observational population study. Lancet 372 (9650), 1655–1660.
Münzel, T., Gori, T., Babisch, W., Basner, M., 2014. Cardiovascular effects of environmental noise exposure. Eur. Heart J. 35 (13), 829–836.
Ow, L.F., Ghosh, S., 2017. Urban cities and road traffic noise: reduction through vegetation. Appl. Acoust. 120, 15–20.
Shapley, L.S., 1953. Stochastic games. Proc. Natl. Acad. Sci. 39 (10), 1095–1100.
Sigrist, F., 2022. Gaussian process boosting. J. Mach. Learn. Res. 23 (232), 1–46.
Skakun, S., Wevers, J., Brockmann, C., Doxani, G., Aleksandrov, M., Batič, M., et al., 2022. Cloud mask intercomparison eXercise (CMIX): an evaluation of cloud masking algorithms for landsat 8 and Sentinel-2. Rem. Sens. Environ. 274, 112990.
Song, J., Zhou, S., Zou, D., Kwan, M.P., Cai, J., Lu, J., 2024. Non-linear associations between noise level and people's short-term noise annoyance in different activity contexts. Environ. Res. 260, 119772.
Sun, C., Shibuya, Y., Sekimoto, Y., 2024. Social segregation levels vary depending on activity space types: comparison of segregation in residential, workplace, routine and non-routine activities in Tokyo metropolitan area. Cities 146, 104745.
Wang, B., Gao, L., Juan, Z., 2017. A trip detection model for individual smartphone-based GPS records with a novel evaluation method. Adv. Mech. Eng. 9 (6), 1687814017705066.
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.
Wang, L., Kwan, M.P., Zhou, S., Liu, D., 2024. Assessing the affective quality of soundscape for individuals: using third-party assessment combined with an artificial intelligence (TPA-AI) model. Sci. Total Environ. 953, 176083.
Wang, L., Zhou, S., Kwan, M.P., Liu, D., Zhang, L., Song, J., 2026. Space-based planning implications for age-friendly cities: insights from spatial non-stationarity in multiple environment-health relationships. Cities 169, 106495.
Wang, L., Kwan, M.P., 2025. How greenery exposure influences noise perception across geographic contexts. Landsc. Urban Plann. 263, 105444.
World Health Organization, 2018. Environmental Noise Guidelines for the European Region World Health Organization. Copenhagen: Regional Office for Europe.
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.

Yu, C., Kwan, M.P., Liu, Y., 2026. Does activity context and its temporality matter in the stress-relieving effect of greenspace? Evidence based on individual mobile sensing. Cities 169, 106526.

Zhang, L., Zhou, S., Kwan, M.P., Chen, F., Dai, Y., 2020. The threshold effects of bus micro-environmental exposures on passengers' momentary mood. Transport. Res. Transport Environ. 84, 102379.

Zheng, L., Kwan, M.P., Liu, Y., 2025. Greenspace morphology and mental well-being: a mobility-based study on urban stress and emotion. Ecol. Indic. 178, 114090.