Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Green-space design and recreational running, walking and cycling: an ecological analysis of fitness-tracking trajectories in Nanchang, China

Siying Huang

huangsiying@ncnu.edu.cn

Nanchang Normal University, Economic and Technological Development Zone

Guanhong Xie

Nanchang Normal University, Economic and Technological Development Zone

Research Article

Keywords: Urban green space, Physical activity, Volunteered geographic information, GPS trajectories, Fitness-tracking applications, Spatial analysis, Landscape configuration, Built environment, Partial least squares structural equation modelling, China

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11526751/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Crowdsourced fitness-tracking trajectories show where recreational physical activity occurs at fine spatial resolution, but few studies have used green spaces themselves as the analytical unit. We examined how internal design, landscape configuration and surrounding urban context are associated with recreational running, walking and cycling across the same set of green spaces.

Methods

In this ecological cross-sectional study, we linked recreational activity trajectories recorded on three fitness platforms to 50 urban green-space polygons in Nanchang, China, yielding 729,670 site-level person-times. Site activity was expressed as mean daily person-times by activity type and compared with field counts (Spearman ρ = 0.78). We derived 97 candidate indicators from site geometry, path networks, Sentinel-2 imagery, landscape metrics, panoramic imagery and points of interest within a 1-km buffer. Partial least squares structural equation models related three formative constructs (internal spatial characteristics, landscape configuration and surrounding urban context) to activity. An exploratory model added a site-level Positive Response Index (PRI) derived from geotagged facial-expression images. Robustness checks included permutation tests with false-discovery-rate control, Moran's *I* and spatial error models.

Results

Internal spatial characteristics were positively associated with running (β = 0.37), walking (β = 0.44) and cycling (β = 0.38). Landscape configuration was associated with running (β = 0.25) and walking (β = 0.21) but not cycling. Surrounding urban context was associated only with walking (β = 0.26). The models explained 32%, 19% and 14% of between-site variance in running, walking and cycling, respectively. In the exploratory model, PRI was associated with walking, and the indirect products through PRI were 0.10 for internal spatial characteristics and 0.13 for surrounding urban context. Neither activity outcomes nor model residuals showed statistically supported spatial autocorrelation.

Conclusions

Within the same set of green spaces, design and contextual attributes showed activity-specific association patterns: internal route and site characteristics were associated with all three activities, whereas the surrounding urban context was associated only with walking. Site-level trajectory analysis provides a practical basis for activity-specific green-space assessment; longitudinal and individual-level studies can test whether these associations reflect effects of design.

Background

Urban green spaces are important settings for physical activity, and their health benefits are well documented [1–3]. Recreational running, walking and cycling are among the main ways in which people use these spaces, but each activity places different demands on the setting. Running and cycling depend on continuous, sufficiently long routes, whereas walking is more closely linked to nearby destinations, entrances and amenities [4, 5]. How a green space is designed, and how it is embedded in its surroundings, may therefore matter differently for each activity [6].

Research on green space and physical activity has usually represented exposure by availability, proximity, area or vegetation quantity around residential locations [7–10]. These measures cannot show whether a particular site is configured to support particular activities, and more greenness is not consistently associated with more physical activity [11]. Large-scale, passively collected location data now allow activity to be observed where it occurs. Volunteered geographic information (VGI), including GPS trajectories recorded by fitness-tracking applications, has been used to quantify visitor activity in parks and green spaces [12, 13] and to estimate walking at population scale [14]. Studies using trajectory data and street-view imagery have linked streetscape greenery, route environments and land use to walking [15–17], running [18–20] and cycling [21–24]; population-level trip records have also been used to map cyclists' exposure to greenery, noise and air pollution along their routes [25], and a recent meta-analysis in this journal identified park area, blue-space proximity and route characteristics as consistent correlates of jogging [26].

Three gaps limit what this literature can tell planners about green-space design. First, most trajectory-based studies use street segments, grid cells or neighbourhoods as the analytical unit, so green space enters the analysis as a street-level or areal exposure rather than as a designed place with internal routes, facilities and landscape structure [15, 26]. Park-level trajectory analyses remain few and have usually examined a single activity [27]. Second, few studies assess internal design, landscape configuration and surrounding urban context at the same green-space unit while distinguishing running, walking and cycling in the same set of sites [28, 29]. Third, crowdsourced trajectories raise methodological questions that are central to health geography but are rarely addressed together: platform users are a selective subset of visitors [30, 31]; associations may depend on the boundaries, spatial units and buffer distances used to define context, the modifiable areal unit problem (MAUP) and the uncertain geographic context problem (UGCoP) [32–35]; neighbouring sites may not be statistically independent; and individual movement data must be handled in ways that protect geoprivacy [36, 37].

Emotional responses are a further, less-studied link between place and activity. Affective responses to exercise vary with intensity [38], and the slower pace of walking may allow more sustained engagement with environmental cues [39, 40]. Geotagged social-media images analysed with automated facial-expression recognition provide place-referenced, if indirect, measures of emotional expression in green spaces [41–44], and Weibo posts have been combined with street-view imagery to relate streetscape

elements to expressed emotion [45]. These measures can be aggregated to the same spatial units as trajectory data.

In this study, we linked fitness-platform trajectories to 50 urban green spaces in Nanchang, China, and used the green space as the unit of analysis. We addressed three research questions:

RQ1. Which dimensions of green-space design and context (internal spatial characteristics, landscape configuration and surrounding urban context) are associated with site-level recreational running, walking and cycling?

RQ2. Which attributes show shared, and which show activity-specific, associations across the three activities?

RQ3. Is a site-level index of positive facial expression (the Positive Response Index, PRI) consistent with an indirect association between green-space characteristics and recreational activity?

The study makes two contributions. Substantively, it compares the same set of green spaces across three activities and shows where their associations converge and diverge. Methodologically, it shows how fitness-tracking trajectories can be aggregated to green-space polygons, checked against field counts and analysed with explicit attention to spatial dependence, spatial scale and geoprivacy. Because comparable trajectory data are available in many countries, for example through Strava Metro [30, 31], this site-level workflow can be applied by researchers elsewhere with different platforms and data infrastructures. Figure 1 summarizes the conceptual framework. Measured place configuration is linked to recorded recreational use through activity opportunities, and the exploratory PRI pathway is shown separately.

Methods

Study design and setting

We conducted an ecological cross-sectional study in which the urban green space was the unit of analysis. The study was conducted in Nanchang, the capital of Jiangxi Province in south-eastern China, with a population of approximately 6.3 million. Reporting follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline for cross-sectional studies [46], adapted to site-level units; the completed checklist is provided in Additional file 1.

Study sites and spatial units

Green spaces were eligible if they (a) were publicly accessible and had a defined boundary and (b) contained VGI-recorded activity on at least two of the three fitness platforms used in this study. Fifty green spaces met these criteria, ranging from large municipal parks to neighbourhood green spaces. Site areas ranged from 0.8 to 426.3 ha (median 18.6 ha; interquartile range (IQR) 6.2–52.1 ha). Each site was

represented by a single polygon, and this polygon was the spatial unit for all site-level measures. Because park and green-space boundary datasets differ in what they include, the delineation of these polygons affects how activity and exposure are assigned to sites [35]. Figure 2 shows the study sites and their internal path networks.

Data sources and spatial framework

Table 1 summarizes the data sources, their temporal coverage, spatial resolution and analytical uses (further details in Additional file 2: Table S1). Figure 3 shows how the data streams entered the analysis. All spatial layers were managed in the WGS 1984 geographic coordinate system (EPSG:4326), and spatial processing, including spatial joins and the calculation of distances, lengths and areas, was carried out in ArcGIS 10.6 (Esri, Redlands, CA, USA). Landscape metrics were computed in FRAGSTATS 4.2 [47]. Structural equation models were fitted in SmartPLS 3.0 [48]. Robustness and spatial-dependence analyses were conducted in Stata 19.0 (StataCorp, College Station, TX, USA), and semantic segmentation and the indirect-effect calculations were run in Python.

Table 1
Data sources and analytical uses.

Dataset	Period	Source	Use in this study
Fitness-platform trajectories (Keep, Codoon, Yuepao Circle)	Sep 2020–Dec 2022	Outdoor fitness platforms	Site-level activity volume for running, walking and cycling
Field observer counts	Nov 2021 (2 days)	Field survey at 5 sites	Validation of trajectory-derived counts
Geotagged facial-expression images	2022	Public Sina Weibo posts	Site-mean Positive Response Index (PRI)
Sentinel-2A imagery (10 m)	2020	Copernicus, via USGS	NDVI, NDWI; land cover for landscape metrics
Digital elevation model (GDEM, 30 m)	2021	Geospatial Data Cloud	Mean slope and elevation
Path network	2022	Baidu Maps, manually refined	Internal path-network indicators; panorama sampling points
Street-view and field panoramas	2021–2022	Baidu Street View; 360° camera survey	Panoramic visual indices (PSPNet), sampled every 50 m along paths
Points of interest	2020	Amap Open Platform API	Service, transport and development-intensity indicators (1-km buffer)
Housing prices	2022	Anjuke	Mean housing price (1-km buffer)
Population grid	2021	WorldPop	Population density (1-km buffer)

NDVI, Normalized Difference Vegetation Index; NDWI, Normalized Difference Water Index; PSPNet, Pyramid Scene Parsing Network; USGS, United States Geological Survey.

Trajectory data: acquisition, preprocessing and site assignment

We obtained VGI records from publicly available activity records on three fitness applications widely used in China: Keep, Codoon and Yuepao Circle. These applications record outdoor exercise trajectories, including route geometry, duration, distance, speed and activity type (running, walking or cycling). Because these applications are used to log exercise sessions, we treated the retained records as recreational activity rather than commuting or utilitarian trips. The records covered 1 September 2020 to 31 December 2022, a period of 852 calendar days.

Of the 1,423,357 records retrieved for the candidate green spaces, we excluded tracks shorter than 500 m, longer than 50 km or lasting less than 3 min, which removed implausibly short or long records and accidental recordings. We then spatially joined the retained trajectories to the green-space polygons in ArcGIS 10.6. A trajectory was assigned to every green-space polygon it intersected, so a trajectory passing through two or more green spaces was counted once for each of them. Each assignment contributed one VGI-recorded activity person-time to that site and activity type. The final dataset contained 729,670 site-level person-times: 477,304 for running (65.4%), 205,037 for walking (28.1%) and 47,329 for cycling (6.5%).

For green space *i* and activity type *m*, the outcome was the mean daily number of VGI-recorded person-times, $\mathcal{N}_{im}=N_{im}/T$, where N_{im} is the number of retained person-times and T = 852 days. Hereafter we refer to $\mathcal{N}_{im}$ as *activity volume*. It measures recorded use of the site, not the number of distinct individuals, and it is not normalized by site area or by the surrounding population.

Validation of trajectory-derived counts

To assess whether trajectory-derived counts reflected on-site use, trained observers counted runners, walkers and cyclists at five green spaces on one weekday and one weekend day in November 2021. We extracted VGI-recorded person-times for the same sites and time windows. The Spearman rank correlation between VGI-derived and observed counts was ρ = 0.78 ($p < 0.001$). VGI-derived person-times corresponded to 77% of observed counts (95% confidence interval (CI): 70%–84%).

Facial-expression data and the Positive Response Index

Automated facial-expression recognition can score emotional expressions in photographs [49]. Valence-based scoring builds on circumplex models of affect, in which positive and negative states occupy opposite poles of a continuous dimension [50, 51], and facial-expression analysis has previously been applied in green-space research with Asian samples [41]. We collected geotagged photographs posted publicly on Sina Weibo in 2022 whose locations fell within the 50 green-space polygons. Images were retained only if all five facial features (eyebrows, eyes, nose, mouth and ears) were clearly visible. Faces were rotated and cropped so that the nasal axis was vertical before scoring. The scoring software also estimates gender and age, but these attributes were not used in this analysis. Expressions were scored with the deep-learning model FireFACE v1.0, which assigns scores to multiple emotion categories [42]. For each image, PRI was calculated as

$$\mathrm{PRI} = S_{\mathrm{happiness}} - S_{\mathrm{sadness}} \quad (1)$$

where $S_{\mathrm{happiness}}$ and S_{sadness} are the model-derived expression scores. In total, 2,379 images (an average of 47.6 per site) contributed to the analysis, and the site-level mean PRI was used as an aggregate measure of positive emotional expression at each site.

Green-space indicators

The initial pool comprised 97 indicators in three dimensions: 35 describing internal spatial characteristics, 43 describing landscape configuration and 19 describing the surrounding urban context (definitions in Additional file 2: Tables S2–S4).

Internal spatial characteristics. These covered the internal path network, site morphology, ecological conditions, facilities and visual conditions. Path data were derived from Baidu Maps and manually refined in ArcGIS. All path-network and morphology indicators, including the maximum activity straight-line distance and activity line density, were derived from the path network and site geometry, not from the VGI trajectories that define the outcomes. Main entrances were identified as discrete public access points from map records and field verification [52]. NDVI and NDWI were calculated from Sentinel-2A imagery, and slope and elevation from a 30-m digital elevation model (DEM). Visual indices were calculated from panoramic images at sampling points spaced 50 m along the internal path network. Images came from Baidu Street View and were supplemented by field panoramas where street-view coverage was missing. Each panorama was stitched from images in eight horizontal directions and segmented with the Pyramid Scene Parsing Network (PSPNet) [53]. Site-level visual indices are means over all sampling points within a site. The visual-data workflow and an example segmentation are shown in Additional file 2: Figures S1–S2.

Landscape configuration. Land cover within each site was classified into forest, shrub, grassland, built-up and water from Sentinel-2A imagery. Class- and landscape-level metrics were computed in FRAGSTATS 4.2 [47] with each site polygon treated as a separate landscape.

Surrounding urban context. These indicators describe land use, points of interest (POIs), public transport, housing, enterprises and population within a 1-km Euclidean buffer around each site. A distance of 1 km corresponds to a walk of roughly 10–15 minutes. Because associations with contextual indicators can vary with buffer size [10, 34], scale effects are considered below.

Statistical analysis

Indicator screening and collinearity. Because the 97 candidate indicators outnumbered the 50 sites, indicators were prescreened before modelling. We computed Pearson correlations between each indicator and the three activity outcomes and carried forward indicators correlated with at least one outcome at $p < 0.05$ (two-sided). The Panoramic Building View Index (PBVI) did not meet this threshold but was carried forward on conceptual grounds, so that the visual environment was represented in the internal-spatial construct. The screening correlations served variable reduction rather than hypothesis testing and were not adjusted for the 291 indicator–outcome comparisons. Multicollinearity among the screened indicators was then assessed with the variance inflation factor (VIF), and indicators with VIF ≥ 5.0 were removed. This procedure retained 46 indicators: 21 internal-spatial, 17 landscape-configuration and 8 surrounding-urban-context indicators (Additional file 2: Table S5). The maximum VIF among the retained indicators was 4.679.

Partial least squares structural equation modelling. We fitted PLS-SEM models because they can combine formatively specified constructs with single-indicator measures in one structural model and are commonly used with small samples and many indicators [54]. Internal spatial characteristics, landscape configuration and surrounding urban context were specified as formative constructs. Their indicators are not interchangeable and jointly define each dimension, rather than reflecting a single underlying factor. Indicator retention followed recommendations for formative measurement models [54, 55]. Indicators with significant outer weights in the bootstrap analysis were retained. Indicators with non-significant outer weights were retained if their outer loadings were at least 0.70 and they were theoretically relevant, because they made a substantial absolute contribution to the construct. Indicators with low loadings and limited theoretical support were removed. Indicators labelled “positive orientation” in Additional file 2 (Tables S5–S6) were entered with their sign aligned to the other indicators of the same construct. Path coefficients, bootstrap standard errors, *t* statistics and *p* values were obtained from 5,000 bootstrap resamples. We report R^2 and adjusted R^2 for each outcome. With $n = 50$, the minimum path coefficient detectable with 80% power at $\alpha = 0.05$ is about 0.35 according to the inverse square root method [56].

Exploratory indirect-association model with PRI. For each activity, we fitted a model in which the place-characteristic constructs were related to site-level PRI (path *a*), PRI was related to activity volume (path *b*), and the conditional direct association was retained (path *c*'). Following established guidance on indirect effects [57], we calculated the indirect product *ab* from the estimated paths and expressed it as a share of the corresponding walking path in the direct model (the total association without PRI). Because the *a* and *b* paths draw on different people (image posters and app users) observed over largely different periods, *ab* is interpreted as a site-level indirect association rather than as individual-level psychological mediation. Approximate 95% CIs for *ab* were calculated with the delta method and, because the sampling distribution of a product is skewed, with the Monte Carlo method (1,000,000 draws) [58]. Both assume zero covariance between the *a* and *b* estimates.

Robustness checks. As a sensitivity check, standardized construct scores exported from SmartPLS were entered into nine site-level ordinary least squares regressions linking each of the three constructs to each of the three outcomes. We used heteroskedasticity-consistent (HC3) standard errors [59], obtained empirical two-sided *p* values by permuting each outcome across sites 10,000 times while holding construct scores fixed, and controlled the false-discovery rate across the nine tests with the Benjamini–Hochberg (BH) procedure [60].

Spatial dependence. Neighbouring green spaces may share unmeasured influences, such as catchment populations or riverside location, which could violate the independence assumption. We therefore tested for spatial autocorrelation in Stata 19.0. Spatial weights were defined as row-standardized inverse distances between green-space centroids, with zero diagonal elements. We calculated global Moran’s *I* for each activity outcome and for the residuals of each non-spatial regression, with empirical *p* values from 10,000 random permutations and BH correction across the nine residual tests. As a secondary specification, we estimated spatial error models by maximum likelihood (Stata spregress, ml, with a spatially autoregressive error term) [61] and examined whether coefficient direction or residual

dependence changed between non-spatial and spatial-error specifications. Because spatial models fitted to a small number of areas can yield unstable estimates [62], the spatial error models were treated as a secondary check rather than as the primary specification.

Spatial scale, zoning and geographic context. Because the analytical units are real, functionally defined green spaces rather than arbitrary zones, the zoning component of the MAUP [32] mainly arises through how site boundaries are delineated [35]. The scale component arises chiefly through the 1-km buffer used for contextual indicators and through the rule used to assign trajectories to sites, both of which reflect the UGCoP [33]. Both were fixed before analysis (a 1-km buffer and assignment to every intersected polygon).

Geoprivacy and data protection

Individual movement data can reveal home locations and daily routines, and their use in health research requires explicit safeguards [36, 37]. Trajectories were used only to produce counts aggregated to the 50 green spaces and three activity types. No individual trajectory, start or end point, or user-level attribute was mapped, analysed or released, and every published value is a site-level aggregate. Facial-expression images were processed automatically. Only per-image expression scores were retained for aggregation to site means, and no facial image is reproduced in this article.

Results

Site characteristics and spatial distribution of activity

Table 2 summarizes the 50 green spaces. Mean daily activity volume per site ranged from 0.6 to 150.5 person-times (median 29.1; IQR 18.0–56.3). Running dominated (median 25.7 per day), followed by walking (median 2.9) and cycling (median 0.0); no cycling was recorded at 25 sites. Site-mean PRI ranged from 4.5 to 39.9.

Table 2
Site area, activity volume and facial-expression scores at the 50 green spaces.

Variable	Median (IQR)	Range
Site area (ha)	18.6 (6.2–52.1)	0.8–426.3
Activity volume (mean daily VGI-recorded person-times per site)		
All activities	29.1 (18.0–56.3)	0.6–150.5
Running	25.7 (15.5–50.9)	0.4–118.2
Walking	2.9 (2.3–5.1)	0.2–31.5
Cycling[a]	0.0 (0.0–0.8)	0.0–27.7
Facial-expression scores (site means)		
Positive Response Index (PRI)	19.0 (13.9–27.1)	4.5–39.9
Happiness score	32.9 (28.8–38.6)	21.9–48.9
Sadness score	13.2 (11.3–15.5)	8.2–19.2

IQR, interquartile range; VGI, volunteered geographic information.

[a]No cycling was recorded at 25 sites.

Figure 4 maps activity volume by activity type. High-volume sites were distributed along the north-east–south-west axis of the Gan River and included the major municipal parks and university-affiliated green spaces in central Nanchang, which recorded activity in all three categories.

Exploratory indicator screening

Pearson screening identified attributes that were positively correlated with all three activities. These spanned site extent (site area and perimeter), path provision (road count and road area), network structure (road-line and intersection density, maximum activity straight-line distance) and landscape structure (Landscape Shape Index and landscape-level number of patches) (Fig. 5).

Other attributes differed by activity. Running was correlated with forest- and grass-patch numbers and sports-facility condition. Walking was correlated with vegetation and water indices, mean water-patch area and the density of nearby services. Cycling was correlated with parking capacity and the number of built-up and water patches. For example, forest-patch number was correlated with running (r = 0.695), mean water-patch area with walking (r = 0.733) and parking capacity with cycling (r = 0.723). The Largest Patch Index was inversely correlated with running for forest patches and with cycling for built-up and water patches. After formative measurement assessment, 25 indicators were retained across the three dimensions (Fig. 6; Additional file 2: Table S6).

Construct-level associations

Internal spatial characteristics were positively associated with running (β = 0.372, p = 0.003), walking (β = 0.437, p = 0.002) and cycling (β = 0.377, p = 0.003) (Table 3; full estimates in Additional file 3). Landscape configuration was positively associated with running (β = 0.254, p = 0.030) and walking (β = 0.209, p = 0.039) but not with cycling (β = 0.016, p = 0.962). Surrounding urban context was positively associated with walking (β = 0.258, p = 0.021) but not with running (β = 0.060, p = 0.586) or cycling (β = 0.088, p = 0.736). The models explained 31.6% of between-site variance in running, 19.0% in walking and 14.2% in cycling (adjusted R^2 0.288, 0.156 and 0.125, respectively). Three of the five statistically supported paths (those for landscape configuration and for surrounding urban context) were smaller than the minimum coefficient detectable with 80% power at n = 50 (about 0.35).

Table 3
Construct-level associations and exploratory PRI paths for recreational activity.

Path	β	SE	*t*	*p* or 95% CI
Panel A. Construct-level associations (PLS-SEM, n = 50 sites)				
Internal spatial characteristics → running	0.372	0.126	2.948	0.003
Internal spatial characteristics → walking	0.437	0.144	3.040	0.002
Internal spatial characteristics → cycling	0.377	0.127	2.960	0.003
Landscape configuration → running	0.254	0.117	2.173	0.030
Landscape configuration → walking	0.209	0.101	2.064	0.039
Landscape configuration → cycling	0.016	0.330	0.048	0.962
Surrounding urban context → running	0.060	0.110	0.545	0.586
Surrounding urban context → walking	0.258	0.112	2.310	0.021
Surrounding urban context → cycling	0.088	0.260	0.337	0.736
R^2 (adjusted R^2): running 0.316 (0.288); walking 0.190 (0.156); cycling 0.142 (0.125)				
Panel B. Exploratory PRI model for walking				
Internal spatial characteristics → PRI (*a*)	0.264	0.107	2.467	0.014
Surrounding urban context → PRI (*a*)	0.332	0.121	2.753	0.006
PRI → walking (*b*)	0.382	0.113	3.382	0.001
Internal spatial characteristics → walking (*c*')	0.254	0.089	2.858	0.004
Surrounding urban context → walking (*c*')	0.355	0.163	2.177	0.030
Internal spatial characteristics → PRI → walking (*ab*)	0.101	–	–	D: 0.002, 0.200
				MC: 0.016, 0.217
Surrounding urban context → PRI → walking (*ab*)	0.127	–	–	D: 0.010, 0.244
				MC: 0.027, 0.263

Standardized path coefficients; SE, *t* and *p* from 5000 bootstrap resamples. Indirect products (*ab*) were computed from the estimated *a* and *b* paths and correspond to 23.1% (internal spatial characteristics) and 49.2% (surrounding urban context) of the corresponding direct-model walking paths in Panel A. D, delta-method CI; MC, Monte Carlo CI (1000000 draws); both assume zero covariance between the *a* and *b* estimates (Additional file 4). CI, confidence interval; PLS-SEM, partial least squares structural equation modelling; PRI, Positive Response Index; SE, standard error.

Exploratory indirect associations through PRI

Statistically supported constituent paths were found only in the walking model. Internal spatial characteristics (a = 0.264, p = 0.014) and surrounding urban context (a = 0.332, p = 0.006) were positively associated with site-mean PRI, and PRI was positively associated with walking volume (b = 0.382, p = 0.001). Conditional direct associations with walking remained positive for internal spatial characteristics (c'=0.254, p = 0.004) and surrounding urban context (c'=0.355, p = 0.030) (Fig. 7).

The indirect product through PRI was 0.101 for internal spatial characteristics (delta-method 95% CI 0.002–0.200; Monte Carlo 95% CI 0.016–0.217) and 0.127 for surrounding urban context (delta-method 0.010–0.244; Monte Carlo 0.027–0.263). Relative to the corresponding walking paths in the direct model (β = 0.437 and 0.258), these products correspond to 23.1% and 49.2%, respectively. Because the PRI model was estimated separately, the decomposition is not additive: for surrounding urban context, the conditional direct association in the PRI model (c'=0.355) exceeds the direct-model path. Both sets of confidence intervals excluded zero, although their lower bounds lay close to it. No indirect pattern was statistically supported for running or cycling. Additional file 4 reproduces the indirect products and both sets of confidence intervals from the estimates in Additional file 3.

Sensitivity and spatial-dependence analyses

The construct-score regressions with HC3 standard errors were directionally consistent with the PLS-SEM estimates. Internal spatial characteristics were positively associated with all three activities, landscape configuration with running and walking but not cycling, and surrounding urban context with walking, with weaker associations for running and cycling. After BH correction across the nine tests, the strongest associations remained statistically supported and the weaker ones were attenuated.

Global Moran's I for the activity outcomes was 0.019 for running (Z = 1.524, p = 0.064), − 0.014 for walking (Z = 0.388, p = 0.349) and − 0.006 for cycling (Z = 0.811, p = 0.209), using the inverse-distance weights. The expected value under the null hypothesis is $-1/(n-1) = -0.020$. None of the outcomes showed statistically supported global spatial autocorrelation. Residual Moran's I statistics from the nine non-spatial models were small, and none remained statistically significant after BH correction. Moran's I for the innovations of the spatial error models was also close to zero. These diagnostics indicate that residual spatial dependence did not materially affect the site-level associations.

Discussion

Principal findings

Using 729,670 site-level person-times derived from fitness-tracking trajectories across 50 green spaces, we found that design and contextual attributes showed different association patterns for recreational running, walking and cycling within the same set of sites. Internal spatial characteristics, chiefly site extent, path provision and network structure, were associated with all three activities. Landscape configuration was associated with running and walking, and the surrounding urban context was associated only with walking. In an exploratory model, a site-level index of positive facial expression

(PRI) was associated with walking and was consistent with a modest indirect association between site characteristics and walking. Neither activity volume nor model residuals showed material spatial autocorrelation. These patterns support assessing green spaces according to how well their spatial conditions match particular activities, rather than by greenness or area alone.

Activity-specific associations of green-space design and context

The results extend research on green space and physical activity beyond availability and greenness. Internal paths and facilities describe opportunities for movement within a site, landscape configuration describes the spatial arrangement through which activity occurs, and the surrounding context describes how a site is embedded in everyday destinations and access networks. Previous work has linked park size and features to park-based activity [7] and has emphasized that aesthetics, accessibility, land-use mix, safety and size jointly shape the use of public open space [6, 63]. Our study adds activity-specific evidence by evaluating these dimensions at the same green-space unit.

Route provision, site morphology and internal network structure were common correlates of all three activities. This is consistent with street-level evidence linking running to streetscape qualities and land-use mix in Shanghai [18], with objective and perceived street environments as complementary predictors of running [20], with park-level evidence linking internal park environments to jogging [27], and with the meta-analytic finding that park area, blue-space proximity and route environments are consistent correlates of jogging [26]. Sites with similar amounts of vegetation may therefore offer different opportunities for recreation, depending on their routes, patch structure and entrances, which is consistent with longitudinal evidence that greenness alone does not reliably predict physical activity [11].

Walking was the only activity associated with the surrounding urban context. This is consistent with evidence that destinations and route attributes are among the most consistent correlates of adult walking [4, 64, 65], and that wayfinding and perceived safety shape comfortable use of green space for walking [66, 67]. Running and cycling may depend more on route continuity, site extent and internal network structure. This interpretation agrees with studies reporting distinct built-environment correlates for cycling [22, 23, 68] and for running and cycling around blue–green spaces [28]. Associations with street-level greenery also vary across activities and spatial units [15, 17, 19, 29], which supports evaluating spatial configuration alongside the quantity of greenery.

PRI as an exploratory indicator of emotional context

In the walking model, internal spatial characteristics and the surrounding urban context were positively associated with PRI, which in turn was associated with walking volume. A plausible interpretation is that walking, being slower, allows more sustained engagement with environmental cues and nearby services [39, 40]. This would be consistent with evidence that affective responses to physical activity vary with intensity [38]. Element-level studies associate grass, trees, water and sky with positive emotional responses, with patterns that vary across environmental contexts [69, 70]. Studies using geotagged

sentiment and facial expressions similarly show that emotional information can be referenced to specific places [43–45, 71]. The PRI pathway is analogous to studies treating walking or social processes as intermediate steps between environments and wellbeing [72].

PRI and activity volume were derived from different people (social-media posters and app users) over largely different periods, and facial expressions in posted photographs are a selective and partly performative signal. The PRI pathway is therefore best understood as a site-level link between emotional context and walking rather than as individual-level mediation. Its contribution is to add an emotional dimension to the site-level assessment of walking environments.

Methodological lessons for trajectory-based green-space research

Four aspects of the analysis may be useful for other studies that use crowdsourced trajectories to evaluate green spaces. First, using the green-space polygon as the analytical unit links outcomes directly to the units that planners design and manage. This avoids treating green space as a street-segment or grid-cell attribute, but it makes results depend on how boundaries are delineated and how trajectories are assigned to sites [35]. Second, checking trajectory-derived counts against field counts, even at a few sites, gives an empirical basis for using platform data as a proxy for on-site use [12, 13]. Rank-order agreement does not, however, remove the platform-user bias documented for crowdsourced activity data [30, 31]. Third, explicit tests for spatial dependence, both of outcomes and of model residuals, are needed when the units are neighbouring sites in one city. In our data, these tests did not indicate material dependence, but this cannot be assumed in denser site networks. Fourth, aggregating trajectories to site-level counts before analysis is a simple and effective geoprivacy safeguard [36, 37]. It also means that individual-level questions, such as who uses which site and from where, cannot be addressed with these data.

Transferability to other settings

The workflow does not depend on Chinese platforms. Researchers elsewhere can apply it to any source that provides geolocated activity records and green-space boundaries, such as Strava Metro or other fitness-app data [31]. Its core steps are: harmonize coordinate systems; assign trajectories to green-space polygons with an explicit and reported rule; aggregate to site-level activity counts before analysis; check counts against field observations; derive design indicators independently of the trajectories; and test residual spatial dependence. Three elements are context-specific. Platform user profiles differ between countries [30]. Coordinate obfuscation (GCJ-02/BD-09) is specific to Chinese map and platform data and must be corrected before overlay. And the facial-expression measure relied on a local social-media platform, whose role would have to be filled by other georeferenced emotion data elsewhere. Testing whether the activity-specific patterns found here, a shared role for internal routes and a walking-specific role for the surrounding context, recur in cities with different climates, urban forms and park systems would show how general they are.

Implications for planning and management

The findings support activity-specific site diagnosis. Audits can examine route provision and continuity, sports-facility condition, patch arrangement, and connections between entrances and nearby services. For running and cycling, evaluations can focus on route provision, continuity and internal network structure. Walking-oriented assessments can also examine nearby daily-living and leisure services, water-related landscape features and entrance connections. Additional file 2: Table S7 distinguishes indicators that are more readily modifiable from those requiring longer-term planning and those that are largely contextual. This helps to separate maintenance and facility upgrades from longer-term changes in site and landscape configuration. Such assessments complement existing green-space quality audit tools [73] and visitor feedback on comfort, safety and perceived quality [6, 74]. Figure 8 summarizes the attributes and evaluation uses for each activity. These uses support the comparison of sites and the prioritization of evaluation; estimating the effects of specific interventions requires longitudinal or natural-experiment designs.

Strengths and limitations

The study's strengths include a large trajectory dataset covering three activity types, the same set of sites for all activities, indicators derived independently of the outcome trajectories, field validation of the trajectory-derived counts, and explicit checks for spatial dependence. Several limitations should be considered.

First, the design is ecological and cross-sectional. Associations at site level may not hold for individuals (the ecological fallacy) [75], and the analysis cannot establish temporal order or individual-level mediation. Second, platform users are not representative of all visitors [30], and the data lack age, sex, socioeconomic and disability information, so differences between population groups could not be assessed. Third, the outcome is total recorded use and was not normalized by site area or by the population living nearby. Larger and more central sites may therefore record more activity partly because of their size and catchment, although site area was included among the internal-spatial indicators.

Conclusions

Fitness-tracking trajectories aggregated to green-space polygons showed that design and contextual attributes are associated with recreational running, walking and cycling in activity-specific ways. Internal spatial characteristics provided a shared basis for all three activities. Landscape configuration was additionally associated with running and walking, and the surrounding urban context with walking only. A site-level index of positive facial expression added an emotional dimension to the spatial associations with walking. Methodologically, the study shows that site-level trajectory analysis, combined with field validation, explicit tests for spatial dependence and aggregation for geoprivacy, can inform activity-specific green-space assessment. Longitudinal, multi-city and individual-level studies are the next step in testing whether these associations reflect effects of green-space design.

Abbreviations

BH: Benjamini–Hochberg

CI: Confidence interval

DEM: Digital elevation model

GPS: Global Positioning System

HC3: Heteroskedasticity-consistent standard errors, type 3

IQR: Interquartile range

MAUP: Modifiable areal unit problem

NDVI: Normalized Difference Vegetation Index

NDWI: Normalized Difference Water Index

PBVI: Panoramic Building View Index

PLS-SEM: Partial least squares structural equation modelling

POI: Point of interest

PRI: Positive Response Index

PSPNet: Pyramid Scene Parsing Network

STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

UGCoP: Uncertain geographic context problem

VGI: Volunteered geographic information

VIF: Variance inflation factor

Declarations

Ethics approval and consent to participate

The study was reviewed and approved by the Academic Ethics Committee of the College of Landscape Architecture and Art, Jiangxi Agricultural University (protocol ES-2022-06-01). Under the approved conditions, facial-expression data were collected only from publicly accessible Sina Weibo posts; content with restricted access or from non-public media, for which individual consent would have been

required, was not used, so individual consent was not required for the publicly posted images. Facial data were stored securely, used only for scientific research and not disclosed to any person or organization unrelated to the work, and no individual's personal information is reported. Fitness-platform trajectories were analysed only as counts aggregated to green-space sites, and the field counts used for validation were anonymous counts of passers-by.

Availability of data and materials

Additional file 2 documents the data sources and the definition, screening and retention of all indicators; Additional file 3 contains the 50-site path estimates; and Additional file 4 contains the Python script (Python ≥3.8, operating-system independent, requires NumPy) that reproduces the reported indirect products and confidence intervals. The raw fitness-platform trajectories (Keep, Codoon and Yuepao Circle) and the facial-expression images (Sina Weibo) are third-party data that cannot be redistributed under the platforms' terms of use and the applicable privacy and ethical requirements.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was supported by the Jiangxi Early-Career Youth Talent Development Project (grant number S202520697), "Mechanisms Linking Multisensory Urban Greenway Environments to Exercise Adherence Based on Digital Intelligence Assessment".

Authors' contributions

SH conceptualized the study, developed the methodology, conducted the formal analysis and investigation, curated the data, prepared the visualizations and wrote the original draft. GX contributed to the conceptualization, supervised the work, administered the project and reviewed and edited the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank the trained observers who conducted the field counts used to validate the trajectory data.

Authors' information

SH teaches at Nanchang Normal University (ORCID: 0000-0002-3797-3397). She received her PhD in Landscape Architecture and Art from Jiangxi Agricultural University in 2024. Her research focuses on digital landscape research and recreational activity in urban green spaces.

References

1. Hartig T, Mitchell R, de Vries S, Frumkin H. Nature and health. Annu Rev Public Health. 2014;35:207–28. https://doi.org/10.1146/annurev-publhealth-032013-182443.
2. James P, Banay RF, Hart JE, Laden F. A review of the health benefits of greenness. Curr Epidemiol Rep. 2015;2:131–42. https://doi.org/10.1007/s40471-015-0043-7.
3. Sallis JF, Floyd MF, Rodríguez DA, Saelens BE. Role of built environments in physical activity, obesity, and cardiovascular disease. Circulation. 2012;125:729–37. https://doi.org/10.1161/CIRCULATIONAHA.110.969022.
4. Sugiyama T, Neuhaus M, Cole R, Giles-Corti B, Owen N. Destination and route attributes associated with adults' walking: a review. Med Sci Sports Exerc. 2012;44:1275–86. https://doi.org/10.1249/MSS.0b013e318247d286.
5. Koohsari MJ, Mavoa S, Villanueva K, Sugiyama T, Badland H, Kaczynski AT, et al. Public open space, physical activity, urban design and public health: concepts, methods and research agenda. Health Place. 2015;33:75–82. https://doi.org/10.1016/j.healthplace.2015.02.009.
6. Koohsari MJ, Oka K, Nakaya T, Talen E, Kaczynski AT. Revisiting public open space and physical activity: what we know, what we still don't, and emerging issues–-a decade later. Health Place. 2025;94:103465. https://doi.org/10.1016/j.healthplace.2025.103465.
7. Kaczynski AT, Potwarka LR, Saelens BE. Association of park size, distance, and features with physical activity in neighborhood parks. Am J Public Health. 2008;98:1451–6. https://doi.org/10.2105/AJPH.2007.129064.
8. Mytton OT, Townsend N, Rutter H, Foster C. Green space and physical activity: an observational study using Health Survey for England data. Health Place. 2012;18:1034–41. https://doi.org/10.1016/j.healthplace.2012.06.003.
9. Jansen FM, Ettema DF, Kamphuis CBM, Pierik FH, Dijst MJ. How do type and size of natural environments relate to physical activity behavior? Health Place. 2017;46:73–81. https://doi.org/10.1016/j.healthplace.2017.05.005.
10. Labib SM, Lindley S, Huck JJ. Spatial dimensions of the influence of urban green-blue spaces on human health: a systematic review. Environ Res. 2020;180:108869. https://doi.org/10.1016/j.envres.2019.108869.
11. van Beek JFE, Malisoux L, Klein O, Bohn T, Tharrey M, Van Lenthe FJ, et al. Longitudinal study of changes in greenness exposure, physical activity and sedentary behavior in the ORISCAV-LUX cohort study. Int J Health Geogr. 2024;23:14. https://doi.org/10.1186/s12942-024-00374-7.

12. Heikinheimo V, Di Minin E, Tenkanen H, Hausmann A, Erkkonen J, Toivonen T. User-generated geographic information for visitor monitoring in a national park: a comparison of social media data and visitor survey. ISPRS Int J Geo-Inf. 2017;6:85. https://doi.org/10.3390/ijgi6030085.
13. Filazzola A, Xie G, Barrett K, Dunn A, Johnson MTJ, MacIvor JS. Using smartphone-GPS data to quantify human activity in green spaces. PLoS Comput Biol. 2022;18:e1010725. https://doi.org/10.1371/journal.pcbi.1010725.
14. Nagata S, Nakaya T, Hanibuchi T, Nakaya N, Hozawa A. Development of a method for walking step observation based on large-scale GPS data. Int J Health Geogr. 2022;21:10. https://doi.org/10.1186/s12942-022-00312-5.
15. Lu Y, Sarkar C, Xiao Y. The effect of street-level greenery on walking behavior: evidence from Hong Kong. Soc Sci Med. 2018;208:41–9. https://doi.org/10.1016/j.socscimed.2018.05.022.
16. Wang R, Liu Y, Lu Y, Yuan Y, Zhang J, Liu P, et al. The linkage between the perception of neighbourhood and physical activity in Guangzhou, China: using street view imagery with deep learning techniques. Int J Health Geogr. 2019;18:18. https://doi.org/10.1186/s12942-019-0182-z.
17. Sakamoto S, Kogure M, Hanibuchi T, Nakaya N, Hozawa A, Nakaya T. Effects of greenery at different heights in neighbourhood streetscapes on leisure walking: a cross-sectional study using machine learning of streetscape images in Sendai City, Japan. Int J Health Geogr. 2023;22:29. https://doi.org/10.1186/s12942-023-00351-6.
18. Guo H, Zhang S, Liu Y, Lin R, Liu J. Building running-friendly cities: effects of streetscapes on running using 9.73 million fitness tracker data in Shanghai, China. BMC Public Health. 2024;24:2251. https://doi.org/10.1186/s12889-024-19605-4.
19. Yang W, Chen H, Li J, Guo W, Fei J, Li Y, et al. How does visual environment affect outdoor jogging behavior? Insights from large-scale city images and GPS trajectories. Urban Urban Green. 2024;95:128291. https://doi.org/10.1016/j.ufug.2024.128291.
20. Dong L, Jiang H, Li W, Qiu B, Wang H, Qiu W. Assessing impacts of objective features and subjective perceptions of street environment on running amount: a case study of Boston. Landsc Urban Plan. 2023;235:104756. https://doi.org/10.1016/j.landurbplan.2023.104756.
21. Sun Y, Du Y, Wang Y, Zhuang L. Examining associations of environmental characteristics with recreational cycling behaviour by street-level Strava data. Int J Environ Res Public Health. 2017;14:644. https://doi.org/10.3390/ijerph14060644.
22. Ito K, Bansal P, Biljecki F. Examining the causal impacts of the built environment on cycling activities using time-series street view imagery. Transp Res Part Policy Pract. 2024;190:104286. https://doi.org/10.1016/j.tra.2024.104286.
23. Gao M, Fang C. Deciphering urban cycling: analyzing the nonlinear impact of street environments on cycling volume using crowdsourced tracker data and machine learning. J Transp Geogr. 2025;124:104179. https://doi.org/10.1016/j.jtrangeo.2025.104179.
24. Gao M, Fang C. Pedaling through the cityscape: unveiling the association of urban environment and cycling volume through street view imagery analysis. Cities. 2025;156:105573.

https://doi.org/10.1016/j.cities.2024.105573.

25. Willberg E, Poom A, Helle J, Toivonen T. Cyclists' exposure to air pollution, noise, and greenery: a population-level spatial analysis approach. Int J Health Geogr. 2023;22:5. https://doi.org/10.1186/s12942-023-00326-7.
26. Huang D, Xu W, Helbich M, Yang X. Associations between the built environment and jogging behaviour based on voluntarily contributed tracking data: a systematic review and meta-analysis. Int J Health Geogr. 2025;24:42. https://doi.org/10.1186/s12942-025-00428-4.
27. Liu Y, Hu J, Yang W, Luo C. Effects of urban park environment on recreational jogging activity based on trajectory data: a case of Chongqing, China. Urban Urban Green. 2022;67:127443. https://doi.org/10.1016/j.ufug.2021.127443.
28. Zhang H, Nijhuis S, Newton C, Tao Y. Healthy urban blue space design: exploring the associations of blue space quality with recreational running and cycling using crowdsourced data. Sustain Cities Soc. 2024;117:105929. https://doi.org/10.1016/j.scs.2024.105929.
29. Mao Y, Xia T, Hu F, Hu J, He Y, Yan J, et al. Greener is not always better: exploring the non-linear relationships between three-dimensional green and gray spaces exposure and various physical activities. Build Environ. 2025;272:112654. https://doi.org/10.1016/j.buildenv.2025.112654.
30. Venter ZS, Gundersen V, Scott SL, Barton DN. Bias and precision of crowdsourced recreational activity data from Strava. Landsc Urban Plan. 2023;232:104686. https://doi.org/10.1016/j.landurbplan.2023.104686.
31. Lee K, Sener IN. Strava Metro data for bicycle monitoring: a literature review. Transp Rev. 2021;41:27–47. https://doi.org/10.1080/01441647.2020.1798558.
32. Openshaw S. The modifiable areal unit problem. No. 38 in Concepts and Techniques in Modern Geography. Norwich: Geo Books; 1984.
33. Kwan MP. The uncertain geographic context problem. Ann Assoc Am Geogr. 2012;102:958–68. https://doi.org/10.1080/00045608.2012.687349.
34. James P, Berrigan D, Hart JE, Hipp JA, Hoehner CM, Kerr J, et al. Effects of buffer size and shape on associations between the built environment and energy balance. Health Place. 2014;27:162–70. https://doi.org/10.1016/j.healthplace.2014.02.003.
35. Fry D, Roman LA, Kondo MC. Comparing mapped park and greenspace boundaries in Philadelphia: implications for exposure assessment in health studies. Int J Health Geogr. 2024;23:20. https://doi.org/10.1186/s12942-024-00370-x.
36. Kamel Boulos MN, Kwan MP, El Emam K, Chung ALL, Gao S, Richardson DB. Reconciling public health common good and individual privacy: new methods and issues in geoprivacy. Int J Health Geogr. 2022;21:1. https://doi.org/10.1186/s12942-022-00300-9.
37. Wang J, Kwan MP. Daily activity locations k-anonymity for the evaluation of disclosure risk of individual GPS datasets. Int J Health Geogr. 2020;19:7. https://doi.org/10.1186/s12942-020-00201-9.

38. Ekkekakis P, Hall EE, Petruzzello SJ. Variation and homogeneity in affective responses to physical activity of varying intensities: an alternative perspective on dose–response based on evolutionary considerations. J Sports Sci. 2005;23:477–500. https://doi.org/10.1080/02640410400021492.
39. Kaplan S. The restorative benefits of nature: toward an integrative framework. J Environ Psychol. 1995;15:169–82. https://doi.org/10.1016/0272-4944(95)90001-2.
40. Johansson M, Hartig T, Staats H. Psychological benefits of walking: moderation by company and outdoor environment. Appl Psychol Health Well Being. 2011;3:261–80. https://doi.org/10.1111/j.1758-0854.2011.01051.x.
41. Guan H, Wei H, Hauer RJ, Liu P. Facial expressions of Asian people exposed to constructed urban forests: accuracy validation and variation assessment. PLoS ONE. 2021;16:e0253141. https://doi.org/10.1371/journal.pone.0253141.
42. Wang L, Zhou C. Evaluation of perceptions using facial expression scores on ecological service value of blue and green spaces in 61 parks in Guizhou. Sustainability. 2024;16:4108. https://doi.org/10.3390/su16104108.
43. Song X, Wu H, Jiang W, Zhi J, Xia X, Long Y, et al. Using geotagged facial expressions to visualize and characterize different demographic groups' emotion in theme parks. Sci Rep. 2024;14:20983. https://doi.org/10.1038/s41598-024-69555-5.
44. Syamili MS, Takala T, Korrensalo A, Tuittila ES. Happiness in urban green spaces: a systematic literature review. Urban Urban Green. 2023;86:128042. https://doi.org/10.1016/j.ufug.2023.128042.
45. Yang X, Shen J. Examining streetscape visuals and emotional responses through social media and street view image analysis. Int J Health Geogr. 2026;25:15. https://doi.org/10.1186/s12942-025-00440-8.
46. von Elm E, Altman DG, Egger M, Pocock SJ, Gtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370:1453–7. https://doi.org/10.1016/S0140-6736(07)61602-X.
47. McGarigal K, Cushman SA, Ene E. FRAGSTATS v4: spatial pattern analysis program for categorical and continuous maps. Amherst: University of Massachusetts; 2012. Software. Available from: https://www.fragstats.org. Accessed 18 Oct 2026.
48. Ringle CM, Wende S, Becker JM. SmartPLS 3. Bönningstedt: SmartPLS GmbH; 2015. Software. Available from: https://www.smartpls.com. Accessed 18 Oct 2026.
49. Kopalidis T, Solachidis V, Vretos N, Daras P. Advances in facial expression recognition: a survey of methods, benchmarks, models, and datasets. Information. 2024;15:135. https://doi.org/10.3390/info15030135.
50. Russell JA. A circumplex model of affect. J Pers Soc Psychol. 1980;39:1161–78. https://doi.org/10.1037/h0077714.
51. Mauss IB, Robinson MD. Measures of emotion: a review. Cogn Emot. 2009;23:209–37. https://doi.org/10.1080/02699930802204677.

52. Wang Y, Yang J, Hong Y, Ma G, Fu Y, Lin Y, et al. From equity to renewal: multi-factor evaluation of urban park accessibility in community life circles, Taijiang District, Fuzhou. Front Environ Sci. 2025;13:1727205. https://doi.org/10.3389/fenvs.2025.1727205.
53. Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE; 2017. pp. 2881–90.
54. Hair JF, Hult GTM, Ringle CM, Sarstedt M. A primer on partial least squares structural equation modeling (PLS-SEM). 3rd ed. Thousand Oaks: SAGE; 2022.
55. Becker JM, Klein K, Wetzels M. Hierarchical latent variable models in PLS-SEM: guidelines for using reflective-formative type models. Long Range Plann. 2012;45:359–94. https://doi.org/10.1016/j.lrp.2012.10.001.
56. Kock N, Hadaya P. Minimum sample size estimation in PLS-SEM: the inverse square root and gamma-exponential methods. Inf Syst J. 2018;28:227–61. https://doi.org/10.1111/isj.12131.
57. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behav Res Methods. 2008;40:879–91. https://doi.org/10.3758/BRM.40.3.879.
58. Preacher KJ, Selig JP. Advantages of Monte Carlo confidence intervals for indirect effects. Commun Methods Meas. 2012;6:77–98. https://doi.org/10.1080/19312458.2012.679848.
59. Long JS, Ervin LH. Using heteroscedasticity consistent standard errors in the linear regression model. Am Stat. 2000;54:217–24. https://doi.org/10.1080/00031305.2000.10474549.
60. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Stat Methodol. 1995;57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
61. Anselin L. Spatial econometrics: methods and models. Dordrecht: Kluwer Academic; 1988.
62. Aswi A, Cramb S, Duncan E, Mengersen K. Evaluating the impact of a small number of areas on spatial estimation. Int J Health Geogr. 2020;19:39. https://doi.org/10.1186/s12942-020-00233-1.
63. Okenwa B, Munar Bauzá M. Experts' insights on physical environmental factors impacting the use of urban public open spaces for physical activity. Wellbeing Space Soc. 2025;9:100308. https://doi.org/10.1016/j.wss.2025.100308.
64. Ewing R, Cervero R. Travel and the built environment: a meta-analysis. J Am Plann Assoc. 2010;76:265–94. https://doi.org/10.1080/01944361003766766.
65. Cerin E, Nathan A, van Cauwenberg J, Barnett DW, Barnett A. The neighbourhood physical environment and active travel in older adults: a systematic review and meta-analysis. Int J Behav Nutr Phys Act. 2017;14:15. https://doi.org/10.1186/s12966-017-0471-5.
66. Ryan DJ, Hill KM. Public perceptions on the role of wayfinding in the promotion of recreational walking routes in greenspace–-cross-sectional survey. Wellbeing Space Soc. 2022;3:100111. https://doi.org/10.1016/j.wss.2022.100111.

67. Wang R, Jiang Y, Liu D, Peng H, Cao M, Yao Y. Is perceived safety a prerequisite for the relationship between green space availability, and the use and perceived comfort of green space? Wellbeing Space Soc. 2025;8:100247. https://doi.org/10.1016/j.wss.2025.100247.
68. Winters M, Brauer M, Setton EM, Teschke K. Built environment influences on healthy transportation choices: bicycling versus driving. J Urban Health. 2010;87:969–93. https://doi.org/10.1007/s11524-010-9509-6.
69. Ho LC, Wei YT, Li D, Chiang YC. Revealing emotional responses to urban environmental elements through street view data and deep learning. Environ Plan B Urban Anal City Sci. 2026;53:647–72. https://doi.org/10.1177/23998083251348280.
70. Cheng Y, Browning MHEM, Zhao B, Qiu B, Wang H, Zhang J. How can urban green space be planned for a 'happy city'? Evidence from overhead- to eye-level green exposure metrics. Landsc Urban Plan. 2024;249:105131. https://doi.org/10.1016/j.landurbplan.2024.105131.
71. Sun P, Lu W, Jin L. How the natural environment in downtown neighborhood affects physical activity and sentiment: using social media data and machine learning. Health Place. 2023;79:102968. https://doi.org/10.1016/j.healthplace.2023.102968.
72. Ross A, Godwyll J, Searle M. Walking mediates associations between the neighborhood environment and flourishing. Wellbeing Space Soc. 2021;2:100014. https://doi.org/10.1016/j.wss.2020.100014.
73. Park H, Brown CD, Pearson AL. A systematic review of audit tools for evaluating the quality of green spaces in mental health research. Health Place. 2024;86:103185. https://doi.org/10.1016/j.healthplace.2024.103185.
74. Zhang P, Lin Y, Zhao B, Ma X, Luo Q, La C. How perceived environmental quality influences physical activity in urban green spaces? A moderated mediation model. Front Public Health. 2025;13:1670171. https://doi.org/10.3389/fpubh.2025.1670171.
75. Robinson WS. Ecological correlations and the behavior of individuals. Am Sociol Rev. 1950;15:351–7. https://doi.org/10.2307/2087176.

Figures

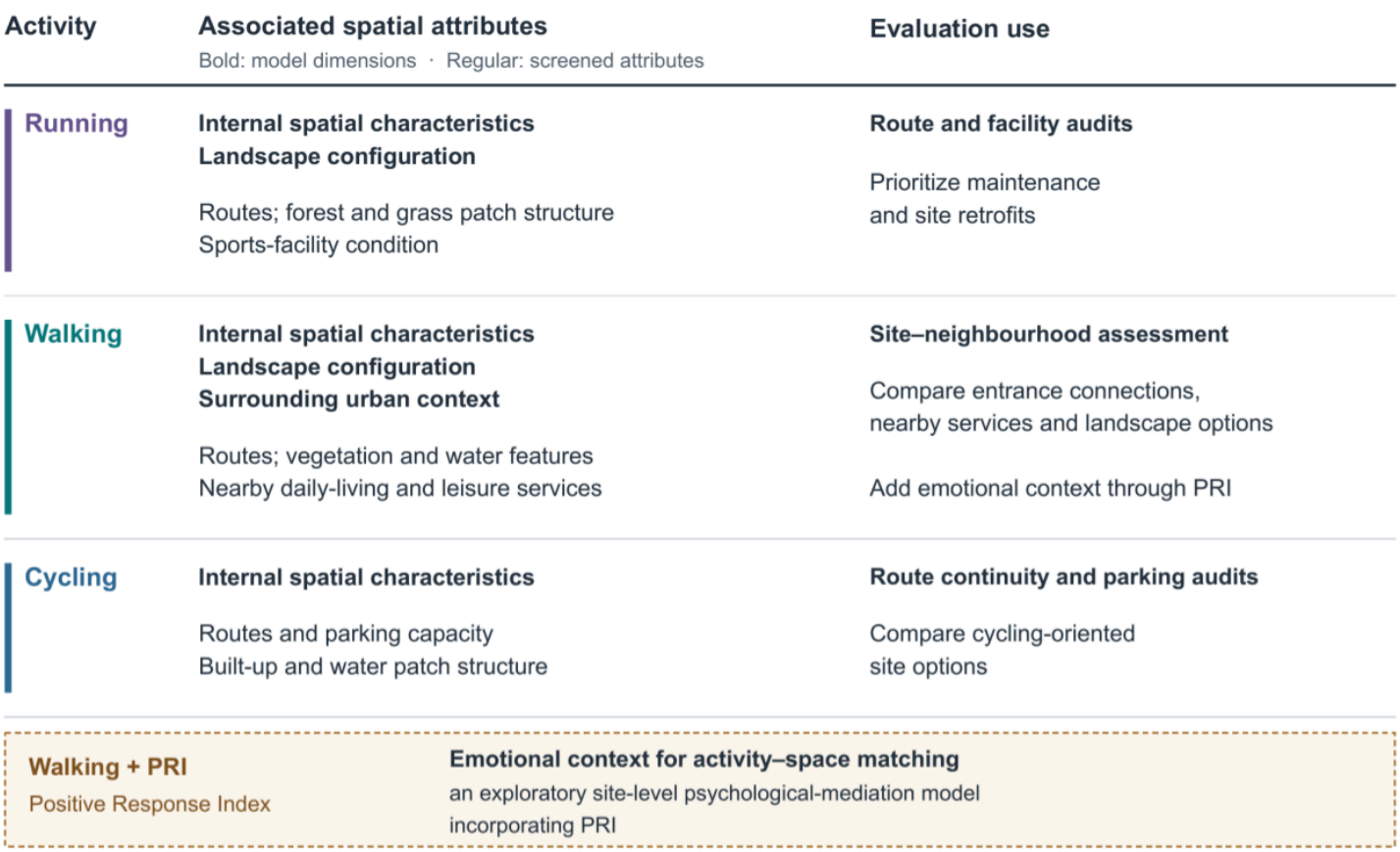


Figure 1

Activity-specific spatial associations and their uses in green-space evaluation. Bold text identifies construct-level associations from the PLS-SEM models; regular text lists selected attributes from the exploratory screening. The shaded band indicates the exploratory contribution of the Positive Response Index (PRI) to walking-oriented assessment. PLS-SEM = partial least squares structural equation modelling.

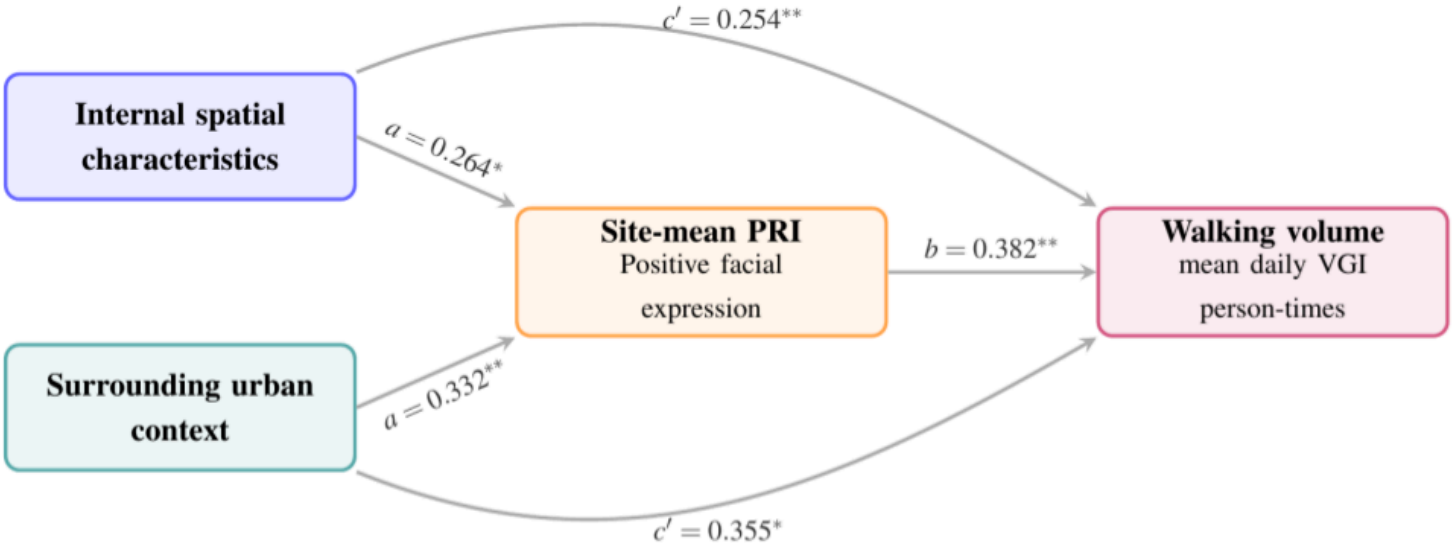


Figure 2

Exploratory PRI model for recreational walking across 50 green spaces. Values are standardized path coefficients: *a*, construct to PRI; *b*, PRI to walking volume; *c*', conditional direct association. * $p<0.05$; ** $p<0.01$. PRI = Positive Response Index; VGI = volunteered geographic information.

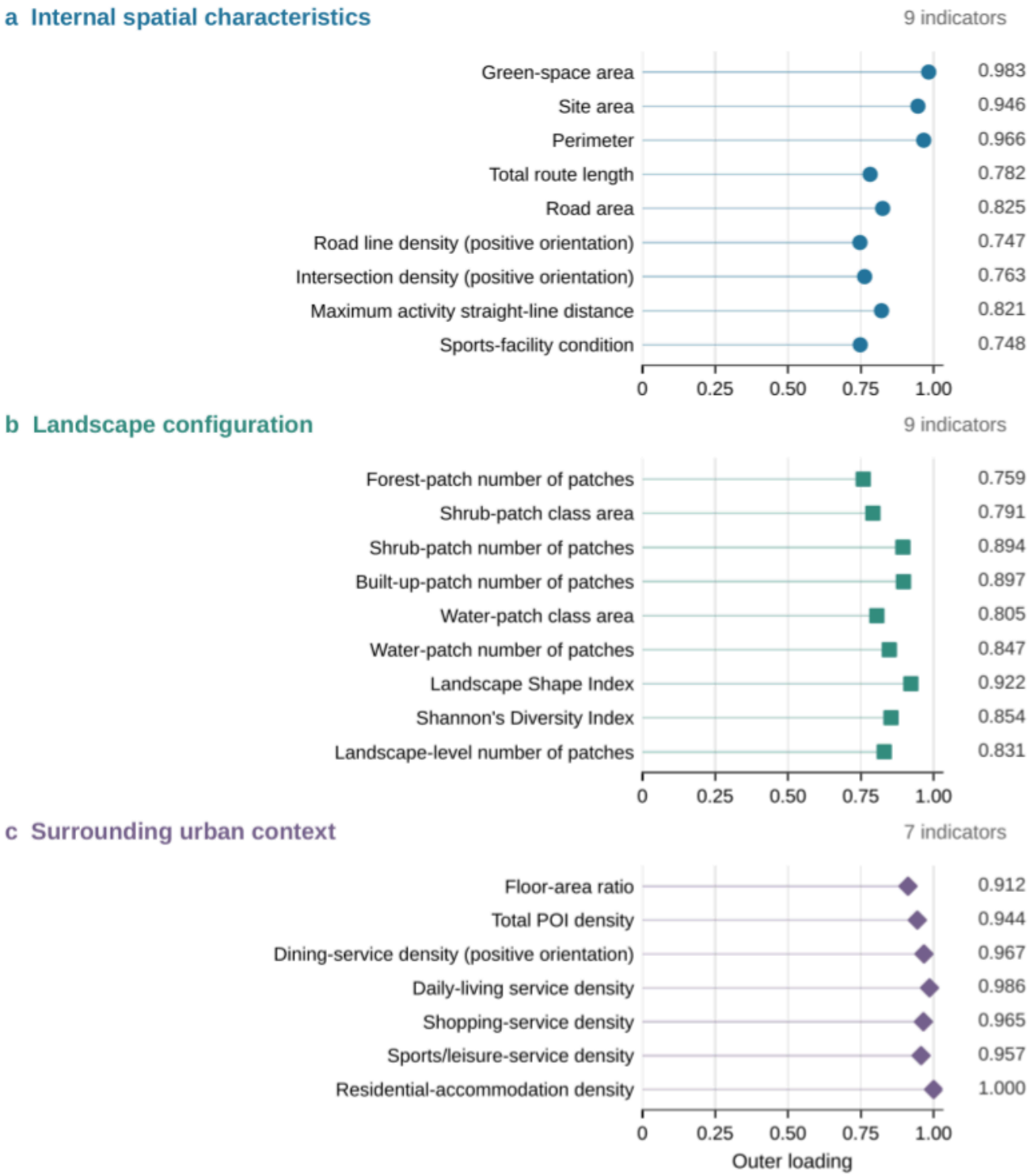


Figure 3

Indicators retained in the formative measurement model and their outer loadings. Panels: (a) internal spatial characteristics (9 indicators), (b) landscape configuration (9) and (c) surrounding urban context (7). “Positive orientation” marks sign-aligned indicators.

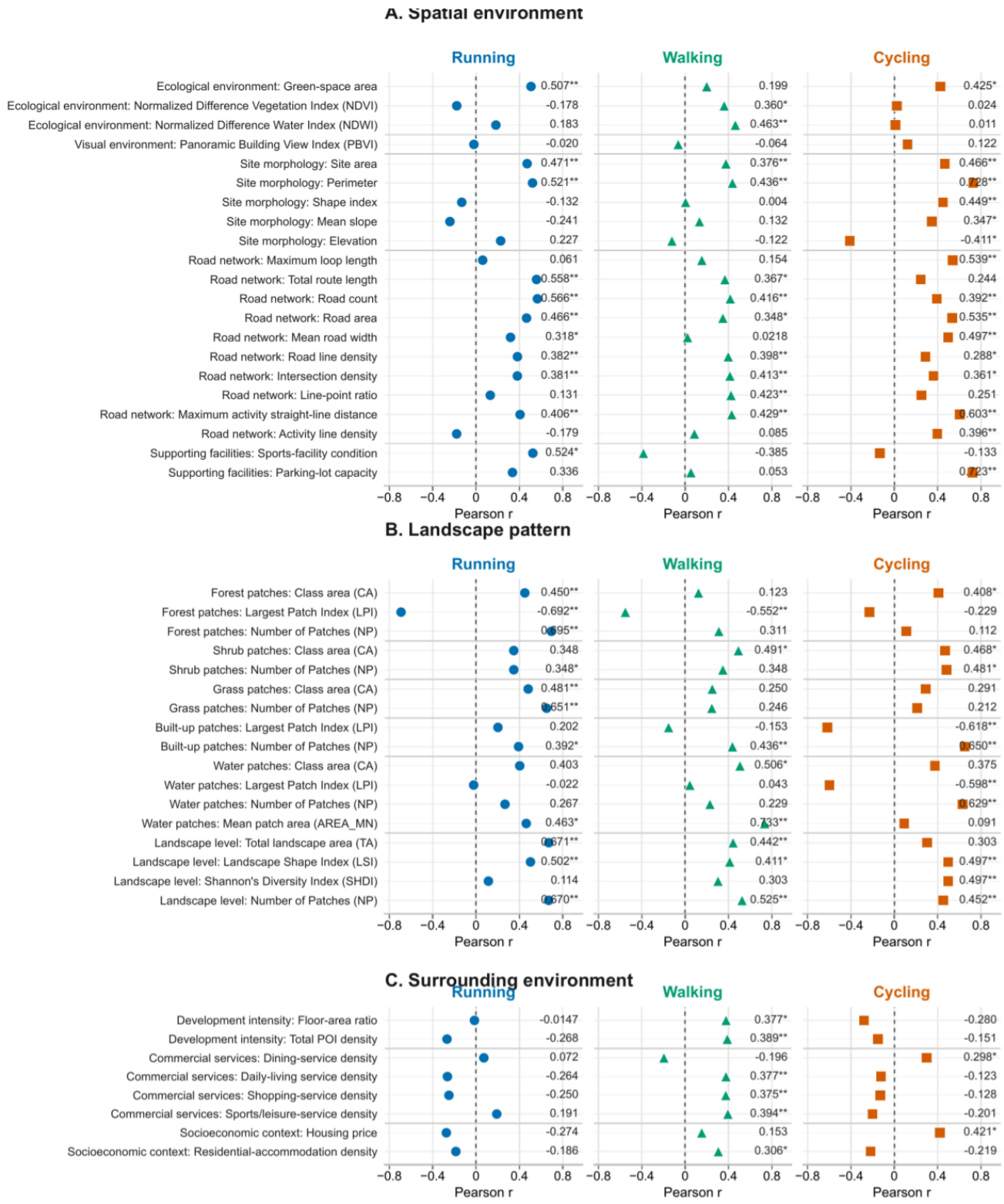


Figure 4

Exploratory screening correlations between retained indicators and activity volume. Pearson correlations between the 46 indicators retained after screening and collinearity filtering and site-level activity volume for running, walking and cycling (*n*=50 sites). Point labels are Pearson *r* values; * *p*<0.05, ** *p*<0.01 (unadjusted). Panel A, internal spatial characteristics (labelled “Spatial environment” in the

graphic); panel B, landscape configuration (“Landscape pattern”); panel C, surrounding urban context (“Surrounding environment”).

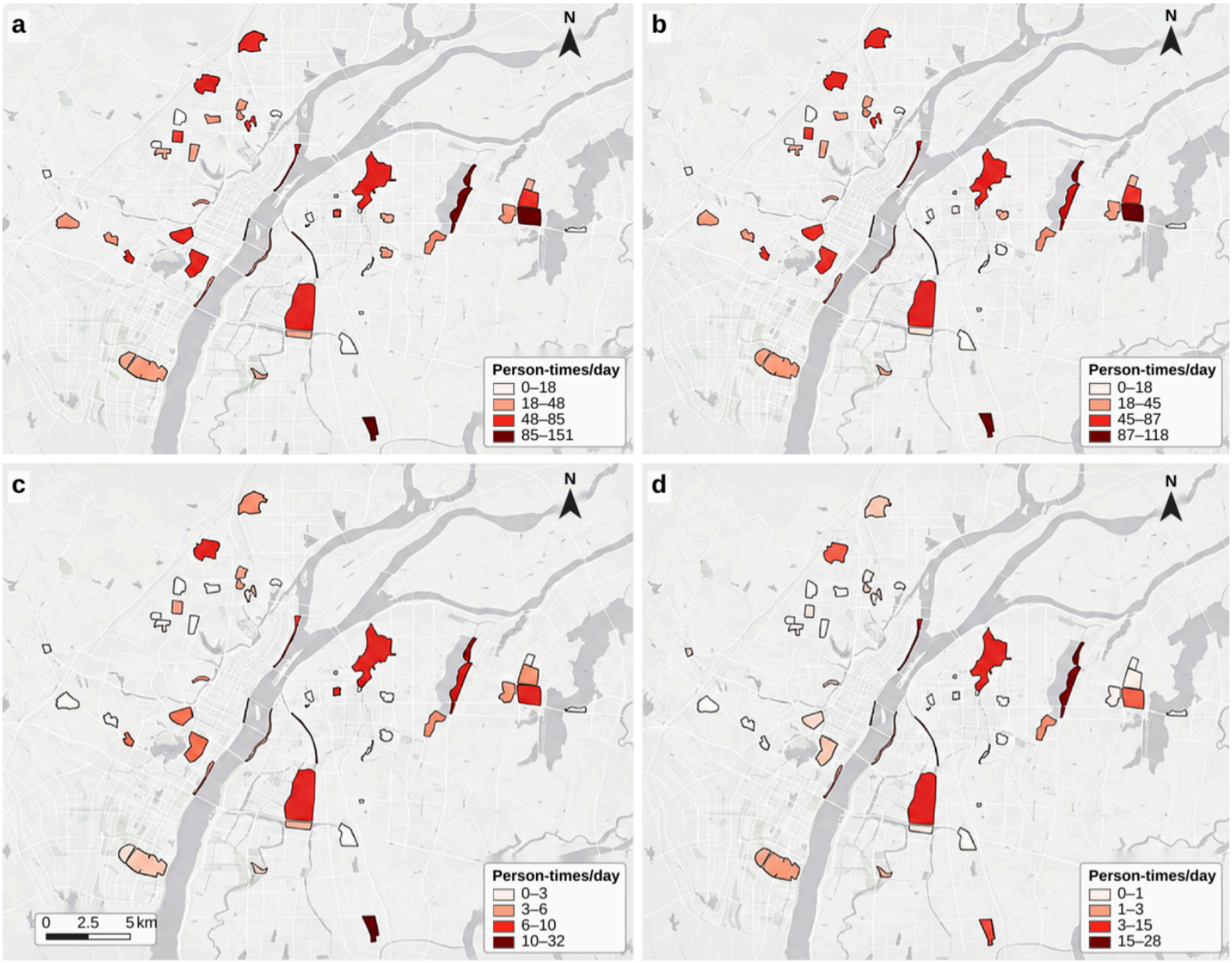


Figure 5

Activity volume at the 50 green spaces, September 2020–December 2022. Activity volume is the mean daily number of VGI-recorded person-times for (a) all activities, (b) recreational running, (c) recreational walking and (d) recreational cycling. Class breaks differ between panels. VGI = volunteered geographic information.

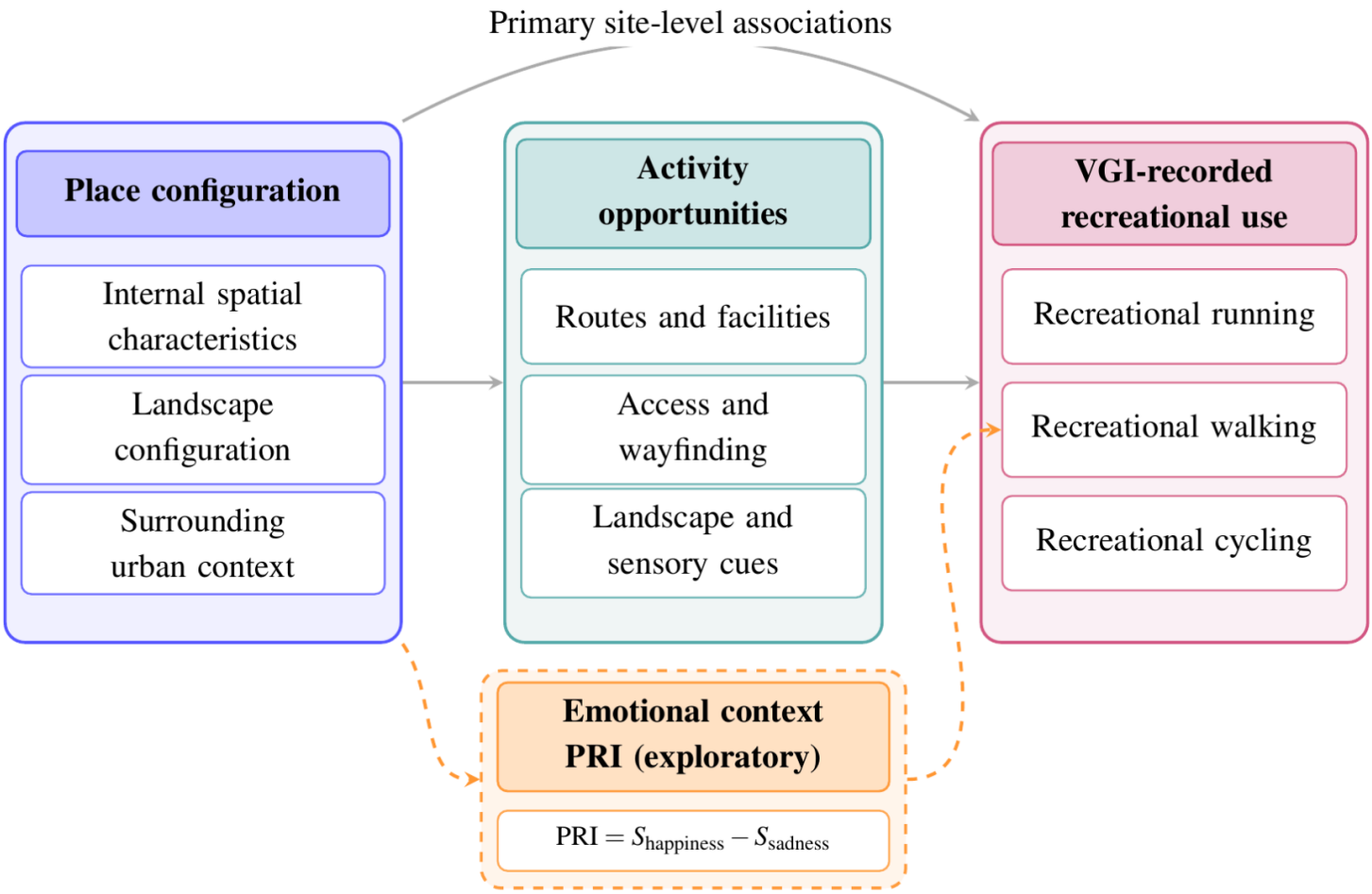


Figure 6

Conceptual framework. Measured place configuration (left) is linked to VGI-recorded recreational use (right) through activity opportunities (centre). Solid arrows show the site-level associations tested in the primary models; dashed orange arrows show the exploratory pathway through the Positive Response Index (PRI). VGI = volunteered geographic information.

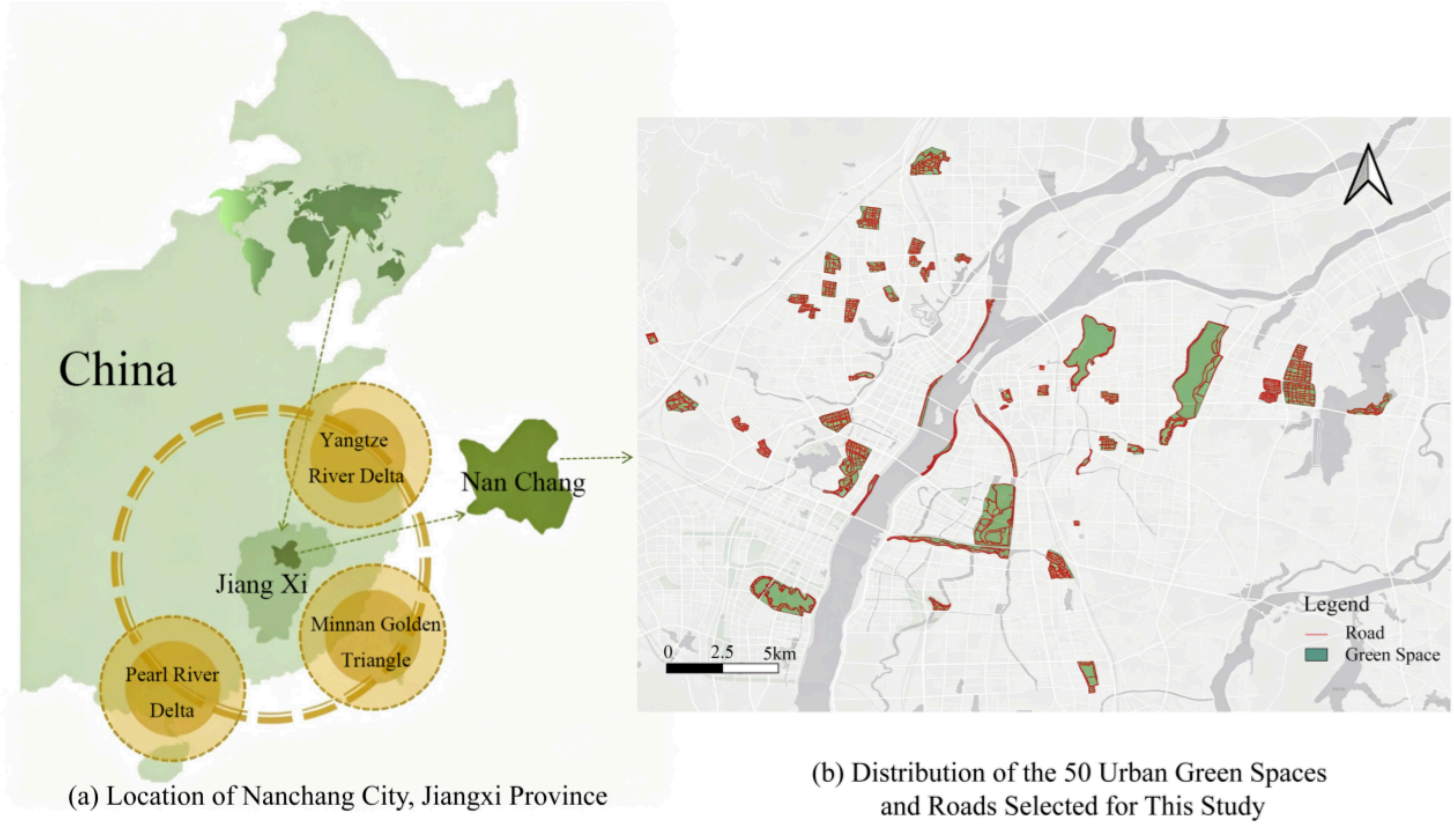


Figure 7

Study area. (a) Location of Nanchang in Jiangxi Province, China. (b) The 50 urban green spaces analysed and their internal path networks.

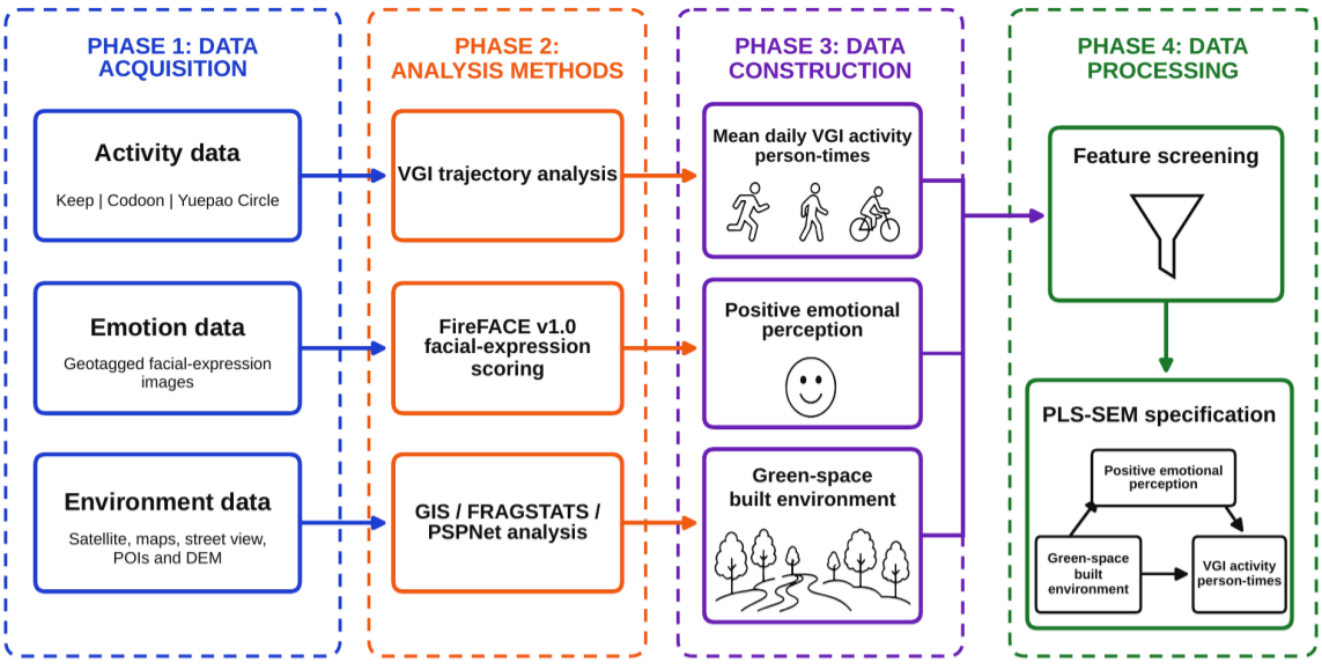


Figure 8

Data acquisition and analytical workflow. DEM = digital elevation model; PLS-SEM = partial least squares structural equation modelling; POI = point of interest; PSPNet = Pyramid Scene Parsing Network; VGI = volunteered geographic information.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile4.py
- Additionalfile3.csv
- Additionalfile1.pdf
- Additionalfile2.pdf