Fang et al. Computational Urban Science (2024) 4:30
https://doi.org/10.1007/s43762-024-00141-1

Computational Urban Science

ORIGINAL PAPER

Open Access

Investigating spatial and temporal characteristics and elements influencing running among residents of Nanchang city in green open areas

Tianwei Fang[1], Linguo Zhou[2,3], Zhenrao Cai[2,3], Zhijin Tan[2,3], Chao Chen[2,3], Jiali Zheng[2,3] and Chaoyang Fang[2,3*]

Abstract

Understanding the exercise behaviors of urban residents in green open spaces is crucial for optimizing future urban spatial planning and improving residents' quality of life. This study provides an in-depth analysis of residents' running behaviors across 87 green open spaces in Nanchang City. Running data collected from the exercise platform Keep, coupled with urban geographic data and field research, reveals the spatial distribution and temporal variations of running activities, including running flow, distance, time, and pace. A set of indicator systems, such as total area, landscape form index, surrounding residential land area, and road conditions, is used to analyze its relationship with the environmental elements of green open spaces. The research results show that the running flow, running distance, and running time in green open spaces in Nanchang City are significantly influenced by different objective elements, while running pace shows little correlation with these environmental elements. Specifically, the study highlights the significant impact of elements such as total area and road length, as well as the number of nearby bus stops and population density, on residents' running behaviors in Nanchang City's green open spaces.

Keywords Green open space, Running flow, Spatial and temporal characteristics, Running distance, Ridge regression model, Nanchang

1 Introduction

Prolonged sitting poses significant health risks, while a lack of physical activity contributes to chronic diseases and premature death (Bankoski et al., 2011; Gao et al., 2024; Piercy et al., 2018). Running can help counteract the harmful effects of prolonged sitting, increase physical activity, and benefit public health by preventing many non-communicable diseases and extending lifespan (Lee et al., 2012; Theodoratou et al., 2016). Studies show that runners have a lower risk of mortality from all causes and cardiovascular diseases compared to those who have never run (Lee et al., 2014). But it has been indicated by previous studies that running in some types of spaces can bring more emotional benefits than in other types. For example, running in forests and parks can relieve stress and improve headaches, thereby improving athletic performance. And running in such green open spaces has a significantly better stress relieving effect than running in cities (Bodin et al.,2003; Hansmann et al., 2007).

Green open spaces are crucial for urban residents to engage in outdoor activities, particularly sports like running (Liu et al.,2021). These spaces are part of the urban environment, distinct from man-made structures, and

*Correspondence:
Chaoyang Fang
fcy@jxnu.edu.cn
[1] Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China
[2] School of Geography and Environment, Jiangxi Normal University, Nanchang 330022, China
[3] Key Laboratory of Poyang Lake Wetland and Watershed Research, Ministry of Education, Jiangxi Normal University, Nanchang 330022, China

consist of areas not covered by building spaces that offer social, economic, and environmental benefits through increased human activity within the city (De Jong et al., 2012). Examples of green open spaces include freely accessible parks, squares, waterfront greenways, sports activity plots, and other areas designated for entertainment and leisure. Therefore, studying the spatial and temporal characteristics of running in green open spaces is of practical significance.

Green open spaces, in particular, offer both essential ecological benefits and valuable services for urban residents (Abdelkarim et al., 2023; Petrunoff et al., 2021; Xue et al., 2017). These areas can attract people to run and improve the efficiency of their running activities (Dinnie et al., 2013; Zhang et al., 2021). As one of China's metropolises, Nanchang boasts a large population and a variety of green open spaces, meeting its residents' strong demand for running venues.

Various elements within green open spaces affect the frequency and performance of runners. Runners typically prefer comfortable running surfaces, green and water environments, favorable weather with suitable temperatures between 20 ℃ and 30 ℃, bright lighting, smooth roads, well-developed infrastructure, and wide streets (Borgers et al., 2016; Deelen et al., 2019; Gao et al., 2016; Huang et al., 2022; Jiao et al., 2017; Tian et al., 2022). However, most current research is focused on the specific spatial environments of cities in Western countries, with relatively little attention given to Chinese cities. Moreover, different data sources and research contexts can yield different conclusions. For example, a GPS-based study in Chengdu, China, found that reduced sky visibility decreased running flow, which contrasts with findings from a similar study in London, UK (Jiang et al., 2022; Schuurman et al., 2021).

In addition, while there is a substantial body of work on walking and cycling, the study of running as a specific form of exercise within urban green spaces remains limited, particularly concerning the spatial and temporal dynamics and the nuanced interplay with environmental elements (Christiansen et al., 2016; Kaczynski et al., 2014). This study will first analyze the spatial and temporal characteristics of residents running in green open spaces in Nanchang City. It will then identify the internal and external environmental indicators of these spaces and the indicators of residents' running status, analyzing their relationship. Based on this analysis, the influence of spatial environmental elements in green open spaces on the spatial and temporal characteristics of running will be determined.

This article first processes the obtained running record data in Nanchang City to obtain running flow, running distance, running time, and running pace data for each green open space. Then it analyzes their temporal and spatial distribution characteristics. Finally, correlation analysis, multilinearity test, and Ridge regression model are used to analyze the relationship between these running data and internal and external environmental elements.

Through this study, the indicator system of the correlation between urban environment and running can be improved. Based on existing research findings, this article selects indicators that affect running from both objective and subjective perspectives of green open space environment. The research results enrich the literature on the relationship between spatial environment and human behavior. In addition, the theoretical framework of the impact mechanism of running in green open spaces can be expanded. Based on existing research theories, the impact relationship between green open space environment and running performance and experience has been sorted out. Based on multi-source data, a regression model is used to analyze the impact of green open space environment on running performance.

2 Research area and data

2.1 Overview of the research area

Nanchang City is located in a subtropical monsoon humid climate zone, characterized by a warm climate with abundant rainfall and sunshine (Zhu et al., 2023). Spring and autumn seasons are short, while summer and winter are long. This city features numerous lakes, including Yaohu Lake, Meihu Lake, Donghu Lake, Xihu Lake, Nanhu Lake, and Beihu Lake, creating a landscape interwoven with rivers and surrounded by lakes. The urban terrain is flat and gently undulating, making it ideal for running activities. The research area for this study includes all districts of Nanchang City and the contiguous built-up area of Nanchang County. Within this range, green open spaces with running trajectories from 2019 to 2023 were selected using data from the exercise platform Keep (https://www.keep.com/).For all the green open spaces, 87 green open spaces with an average annual running flow of over 100 records over the five years were selected as the research objects. In total, the selected green open spaces in Nanchang City were identified for the study, as shown in Fig. 1.

2.2 Indicator selection

2.2.1 Selection of internal environmental element indicators for green open spaces

The physical environmental elements within green open spaces are the core attraction for runners, and the quality of these spaces is mainly reflected in their ecosystem services and green infrastructure (Zhu et al., 2017).

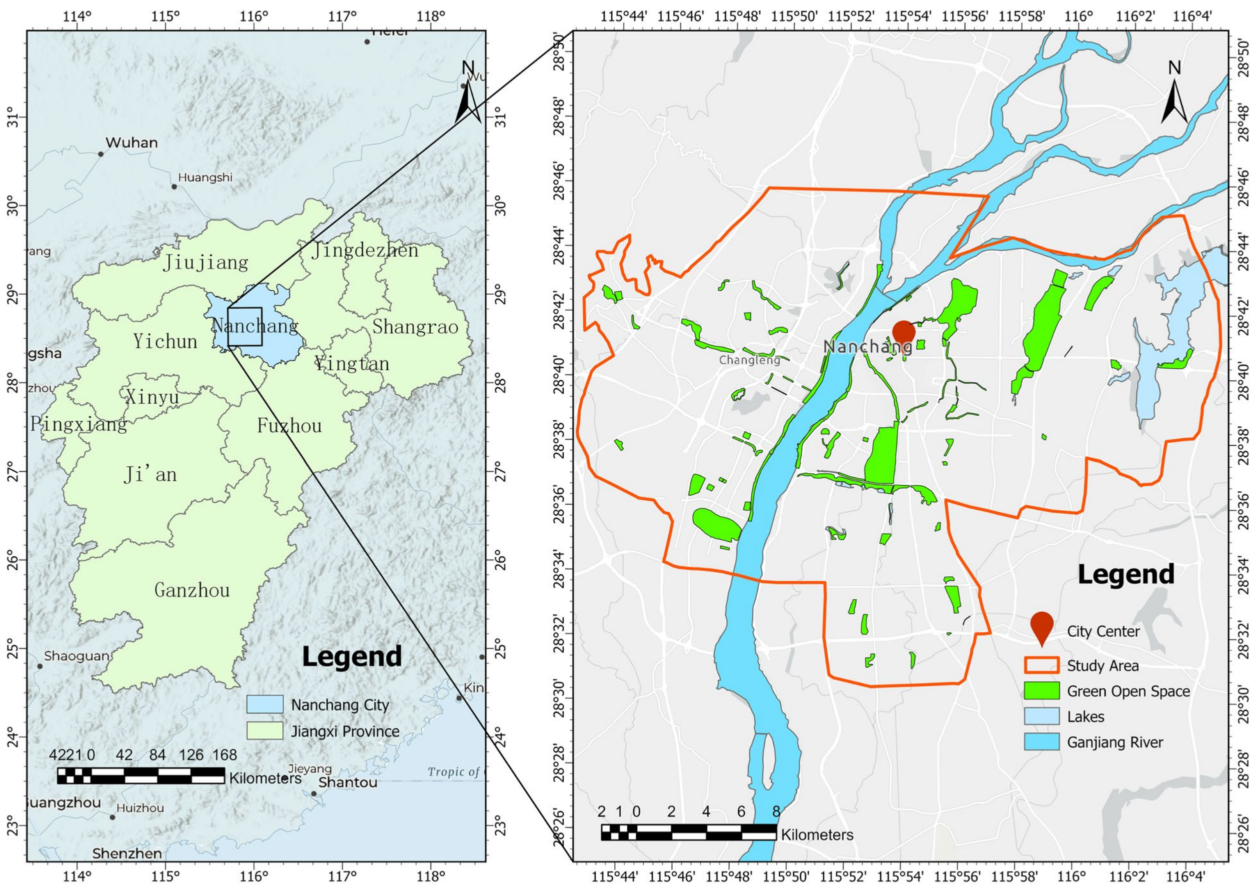

Fig. 1 The study area and green open spaces in Nanchang City

Research has shown that large parks can promote walking and running, making the area of green open spaces a key indicator (Murtagh et al.,2012). Vegetation, particularly, trees, is an important component, with studies demonstrating a positive correlation between tree presence and the frequency of physical activity (Akpinar et al.,2017). Different environmental types also impact running differently (Zhong et al., 2022). Moreover, elements that positively impact running include relatively long road sections and higher safety (Jiang et al., 2022). Comfortable roads, such as synthetic surface tracks, help runners focus, maintain speed, reduce injury risk, and enhance running performance, making them essential for the indicator system (Borgers et al., 2016). Health facilities are another important aspect of green infrastructure, with research indicating that better park infrastructure correlates with higher usage rates, which needs to be analyzed and validated (Coombes et al., 2010).

To quantify the impact of vegetation on running, the indicator normalized vegetation index (NDVI) is used, which is usually used to assess vegetation cover, growth, and drought conditions, and land cover changes. This article uses the average NDVI of all pixels in a green open space to represent the NDVI of this green open space. NDVI is calculated using the following formula:

$$NDVI = \frac{B_{NIR} - B_R}{B_{NIR} + B_R}$$

In the equation, B_{NIR} represents near infrared band reflectance of remote sensing images, B_R represents red band reflectance of remote sensing images.

To quantify the impact of different kinds of environments on running, landscape shape index (LSI), and presence of blue landscapes are used. LSI, which is usually used to assess landscape shape, ecological connectivity, habitat quality, and land cover changes, is calculated using the following formula:

$$LSI = \frac{2\sqrt{\pi S}}{C}$$

In the equation, S represents the area of green open space, C represents the Perimeter of green open space. For presence of blue landscapes, if there is any blue landscape of river or lake in or on the boundary of a green open space, the presence of blue landscapes will be marked as 1, or it will be marked as 0.

In summary, this study selects the areas of green open spaces, NDVI, LSI, and presence of blue landscapes as indicators to characterize natural environmental characteristics. Indicators of green infrastructure include road width, road length, road density, presence of synthetic surface tracks, and number of public toilets.

2.2.2 Selection of external environmental indicators for green open spaces

Green open spaces are an important component of urban areas, and the surrounding environmental characteristics significantly influence them. High-density building environments and transportation infrastructure impact physical activities like walking and cycling (Christiansen et al., 2016; Kaczynski et al., 2014). The high-accessibility road network and dense commercial land that usually exist in high-density building environments is positively correlated with running and walking (Sugiyama et al., 2019; Zhang et al., 2024). However, another element the environments, overwhelming architectural environment, has a negative impact on running performance (Harden et al., 2024). Therefore, it is necessary to analyze their detailed impact on running performance. Researchers have differing opinions on how population density affects running. Ettema (Ettema, 2016) suggests that obstacles in public transportation in densely populated urban areas can reduce the enjoyment of running, while Huang et al. (Huang et al., 2022) found that running satisfaction is not related to population density. Furthermore, more public transportation nodes and facilities can promote running.

In summary, this study selects the density of four elements in the buffer zone—population, bus stops, intersections, and road length—as well as the surrounding building, residential and commercial land ratio, distance to the city center, and average nighttime lighting, which represents the average nighttime light data within each green open space, can comprehensively reflect information such as urbanization level, economic development, and population, to characterize the external environmental elements of green open spaces. And this study defines the buffer zone around green open spaces as a 10-min walking distance, which equates to approximately 800 m for an average person (Hansmann et al., 2007). For the city center of Nanchang, it is defined as Bayi Square, as it is the intersection of major roads in the old city including Ruzi Road, Beijing Road, Zhongshan Road, and Bayi Avenue and the first two metro lines of the city, Line 1 and Line 2. At the same time, the most important commercial district, the oldest block, most tourist attractions, and many government agencies in the city are gathered around it. So, Bayi Square is regarded as the city center of Nanchang.

These results in an evaluation index system where the objective environmental elements of green open spaces are the independent variables and resident running status is the dependent variable, as shown in Table 1.

2.2.3 Selection of indicators for residents' running status

For runners, running in green open spaces is influenced not only by the internal physical environment but also by external elements. Running performance and experience result from the interaction between the internal and external physical environments of these spaces. When both environments are suitable for running, the number of runners running in the space may increase, along with extended running times, longer distances, and faster paces. This can also lead to more frequent runs, greater enjoyment, and the formation of a positive exercise atmosphere (Ni et al., 2024). Therefore, based on relevant data from the Keep, running flow, running distance, running time, and running pace were selected as dependent variables representing residents' running status.

2.3 Data collection

In China, Keep is widely used, with 50% of all sports and fitness app users on the platform. According to KEEP's prospectus (https://www1.hkexnews.hk/listedco/listconews/sehk/2023/0630/2023063000203.pdf) and platform of TalkingData (https://mi.talkingdata.com), in 2022, its monthly active users were 36.4 million, the highest among all online sports applications in China, accounting for about 10% of the total Chinese fitness population (people who exercise twice or more per week) of 37.4 million during the same period, and more than half of all sports application users in China. Meanwhile, the users of KEEP software are mainly young people under the age of 30 (76.6%), residents of large and medium-sized cities like Nanchang (54.2%) and those with a bachelor's degree or above (76%), with more females (58%) than males (42%). The running data for this study is sourced from the running exercise data shared by Keep users. When a user presses the start run button on the software in a green open space, he is considered running by KEEP until the end run button is pressed, during which his running data is recorded by KEEP. This study records and summarizes the running check-in data from 87 green open spaces, in 5 years of 2019–2023, resulting in a total of 22,001 data records, which are biasedly distributed in the green open spaces.

In terms of environmental data, the areas of green open spaces, NDVI, and building land areas are vectorized using sky map remote sensing images in QGIS (https://map.tianditu.gov.cn). Synthetic surface tracks within the

Table 1 All the elements used in the study

Type	Variable and number	Unit	Method
Internal environmental elements	Total area *X1*	km^2	GIS spatial measurement and calculation
	NDVI *X2*	-	GIS spatial measurement and calculation
	Blue landscape *X3*	-	With blue landscape, it is marked as 1. Otherwise, it is marked as 0
	LSI *X4*	-	$\frac{2\sqrt{\pi \times Area}}{Perimeter}$
	Road length *X5*	km	GIS spatial measurement and calculation
	Road density *X6*	km/km^2	$\frac{Roadlength}{Area}$
	Road width *X7*	m	GIS spatial measurement and calculation
	Synthetic surface track *X8*	-	With synthetic surface track, it is marked as 1. Otherwise, it is marked as 0
	Number of public toilets *X9*	-	GIS spatial measurement and calculation
External environmental elements	Density of intersections nearby *X10*	$-/km^2$	GIS spatial measurement and calculation
	Road density nearby *X11*	km/ km^2	GIS spatial measurement and calculation
	Residential land ratio nearby *X12*	%	GIS spatial measurement and calculation
	Population density nearby *X13*	$people/ km^2$	Vectorization of the heat map
	Distance to the city center *X14*	km	GIS spatial measurement and calculation
	Average night light *X15*	cd/m^2	GIS spatial measurement and calculation
	Density of bus stops nearby *X16*	$-/ km^2$	GIS spatial measurement and calculation
	Building land ratio nearby *X17*	%	GIS spatial measurement and calculation
	Commercial land ratio nearby *X18*	km^2	GIS spatial measurement and calculation
Running elements	Runing flow *Y1*	person times/year	Showing in Sect. 3.1
	Running distance *Y2*	km	Showing in Sect. 3.1
	Running time *Y3*	min	Showing in Sect. 3.1
	Running pace *Y4*	min/km	Showing in Sect. 3.1

spaces are identified through Baidu Live Map (https://map.baidu.com), as well as relevant images and videos uploaded by Keep users. The number of nearby public toilets and road network data are obtained from Open Street Map's global road network data (https://www.openstreetmap.org). Surrounding residential data is sourced from the third national land survey from Department of Natural Resources of Jiangxi Province, while population quantity and density are sourced from the seventh population census from National Bureau of Statistics of China (https://www.stats.gov.cn/).

3 Research method

3.1 Running data processing

3.1.1 Running flow data

Running flow refers to the number of running records check-in data. This study analyzes the average annual running flow and its calculation formula is as follows (unit: person times/year):$A_i = 365 * \frac{v_i * b_i}{d_i}$

In the equation, Ai represents the average annual running flow of green open spaces, vi denotes the total running flow on the recorded day, bi indicates the proportion of running flow, and di signifies the total number of days.

3.1.2 Running distance data

This study analyzes the average running distance in green open spaces, calculated using the following formula (unit: kilometers):

$$K_i = \frac{\sum x_i}{i}$$

In the equation, Ki represents the average running distance, xi denotes the running distance of a single record, and i indicates the total number of records.

3.1.3 Running time data

This study analyzes the average running time, calculated using the following formula (unit: minutes):$T_i = \frac{\sum x_i}{i}$

In the equation, Ti represents the average running time, xi denotes the running time of a single record, and i signifies the total number of records.

3.1.4 Running pace data

This study uses average running pace for analysis, which involves preprocessing statistical running pace data. Firstly, the average running speed was calculated using the following formula (unit: kilometers per hour):

$$S_i = \frac{\sum x_i}{i}$$

In the equation, *Si* represents the average running pace, *xi* denotes the running pace of a single record, and *i* indicates the total number of records.

After obtaining the average speed, the running speed (distance covered per hour) was converted into running pace (time taken per kilometer) to calculate the mean running pace in green open spaces. The conversion formula is as follows (unit: minutes/kilometer):

$$s_i' = \frac{60}{s_i}$$

In the equation, s_i' represents the average running pace, and s_i is denotes the average speed from the previous equation. For example, if the result s_i of the previous formula is 10 km per hour, then the pace s_i' will be 60 divided by 10 to obtain 6 min per kilometer.

3.2 Environmental data processing

3.2.1 Pearson correlation

Based on the indicator system developed earlier, this study conducted correlation analysis to examine the relationships between the dependent variables (running flow, running distance, running time, running pace) and the indicators. The Pearson correlation coefficient was used to quantify the strength of these relationships, calculated using the following formula:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

Considering the relationships between variables, this study uses a matrix to assess the correlations among independent variables.

3.2.2 Normalization

Due to the units of variables vary considerably, normalization of each variable is necessary before conducting multilinearity test and ridge regression. This involves converting data from different scales to a common scale between 0 to 1 to avoid the impact of different variable magnitudes on the results. The method used is Min Max Normalization, whose formula is as follows:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

In the equation, x represents variable value before normalization, x_{max} represents the maximum value of the variable, and x_{min} represents the minimum value of the variable.

3.2.3 Multiple linear regression

Multiple linear regression is an extended linear regression method used to analyze the impact of multiple independent variables on a dependent variable. The basic principle is to find the best fitting linear equation by minimizing the sum of squared residuals (i.e. the sum of squared differences between actual and predicted values). This equation can be expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon$$

In the equation, Y represents the dependent variable, $X_1, X_2, \cdots, X_n$ stands for the independent variable, β_0 represents the intercept, $\beta_1, \beta_2, \cdots, \beta_n$ denotes the regression coefficient, and ε represents the error term.

Multiple linear regression is suitable for the following environment: There is a linear relationship between the data; There is no multicollinearity between independent variables (i.e., there is no high correlation between independent variables); The residual term follows a normal distribution and has a constant variance. For example, previous studies used linear regression analysis to test the associations between the percentage of green land cover in the neighborhood and minutes of American elderly people's neighborhood walking per day (Besser et al.,2021).

3.2.4 Variance inflation factor

Multicollinearity between the corresponding independent variables is assessed using the variance inflation factor (VIF) for each variable. In multiple regression analysis, the relationship is defined as follows:

$$VIF = \frac{1}{1-R_i^2}$$

In the equation, R represents the negative correlation coefficient between one independent variable and the others in regression analysis.

A higher variance inflation factor (VIF) indicates an increased potential for collinearity among independent variables. Typically, severe multicollinearity is considered when the VIF exceeds 10.

3.2.5 Ridge regression

Ridge regression is a regression technique used to handle multicollinearity problems. Multicollinearity can lead to unstable estimation of regression coefficients, which in turn affects the predictive ability of the model. Ridge regression stabilizes the estimation of regression coefficients by adding a regularization term (penalty term) to the least squares method. The basic equation of ridge regression is:

$$\beta = (X^T X + \lambda I)^{-1} X^T \bar{y}$$

In the equation, β represents estimation of ridge regression coefficients, X represents independent variable matrix, $\bar{y}$ represents dependent variable vector, λ represents ridge parameters that control the degree of regularization, I represents identity matrix.

Introducing regularization terms, ridge regression can reduce the complexity of the model and minimize the risk of overfitting. In addition, the ridge parameters of ridge regression can be adjusted according to actual situations, making the model more flexible and applicable to different datasets and problems.

Ridge regression is applicable in the following cases: the data exhibits multicollinearity; Hope to reduce the variance of the model and improve its predictive performance by introducing bias; It is necessary to impose certain constraints on the regression coefficients to avoid overfitting.

To address collinearity issues, this article uses a ridge regression model for practical verification. Ridge regression, an improved least squares estimation method, mitigates the impact of collinearity and enhances the accuracy of model parameters estimation. In the field of sports related analysis, there is currently limited research applying ridge regression. Wang X. used ridge regression to analyze the relationship between physical fitness indicators and sports performance of students in Henan Province from 2000 to 2014, and found a positive correlation (Wang X.,2019). However, ridge regression has many applications in other fields, such as previous studies establishing a relationship between remote sensing images and the average breast height diameter and biomass of trees in Maoershan, China using ridge regression when it encountered collinearity problems as in this study (Wang Q. et al., 2011). Therefore, due to the collinearity issue between independent variables, this study utilized ridge regression to analysis and enriched its applications in this field.

4 Impact of environmental elements in green open spaces on running

4.1 Spatial and temporal characteristics of running activities in green open spaces

4.1.1 Temporal characteristics of running activities in green open spaces

Statistical analysis was conducted on the start and end times of 22,201 runner data points collected, examining the running flow within each hour. The results are shown in Fig. 2.

Running in green open spaces in Nanchang City shows distinct peaks in the morning and evening, with morning running flow surpassing evening flow by more than 1.3 times (Liu et al., 2022). The morning peak occurs primarily from 5:00 to 8:00, while the evening peak spans from 19:00 to 21:00. Based on meteorological data, Nanchang residents prefer to run in the morning, particularly when humidity is higher and perceived temperatures are lower.

From a gender perspective, as shown in Fig. 3, females show a higher inclination to run during the morning peak compared to males but are less inclined during the evening peak. Additionally, while the start and end times of both males and females during the morning peak align, males tend to start their evening runs nearly an hour later than females. This difference may be related to women's

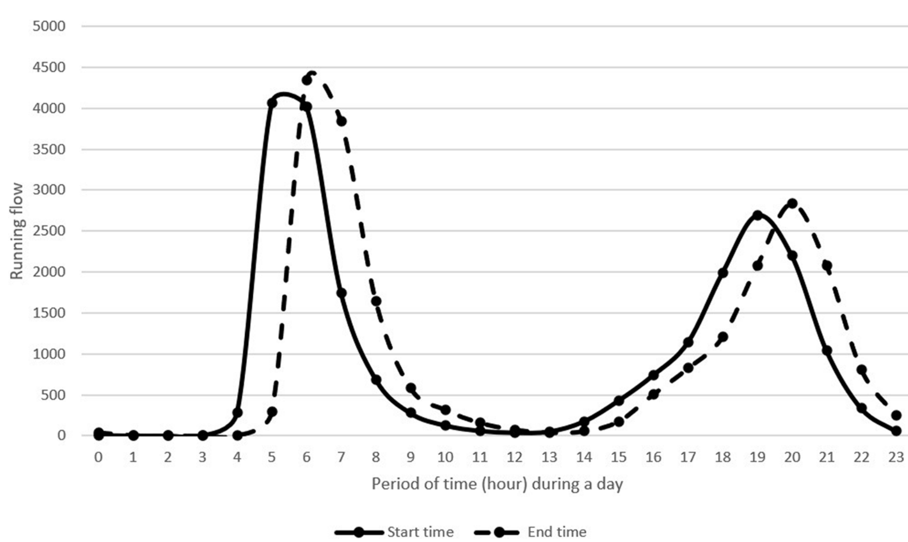

Fig. 2 Time variation of running flow in green open spaces

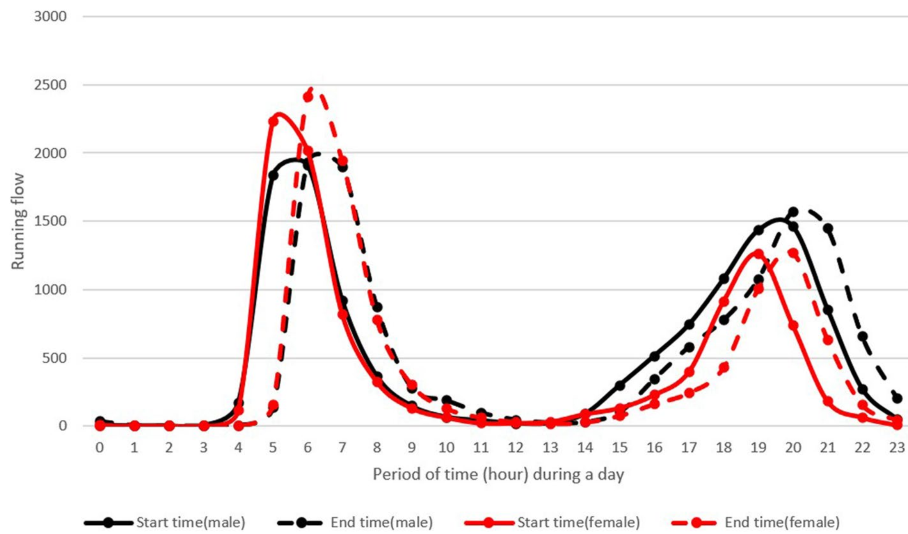

Fig. 3 Start and end times of running in green open spaces by gender

heightened concern about safety factors in the evening in the city.

Looking at different days of the week, Saturday sees the highest running flow, followed by Sunday, while Friday records the lowest running flow, followed by Monday. According to Fig. 4, during weekday mornings, the peak times for starting and ending runs are earlier compared to weekends, with lower running flows. Conversely, on weekends, the morning rush hour for running starts and ends an hour later, accompanied by higher running flows. However, evening running flows on weekends are lower than on weekdays, possibly because individuals with lower running frequencies may prefer weekend mornings for running activities. Nevertheless, the average duration of evening runs on weekends tends to be longer than on weekdays.

From a seasonal perspective, summer shows the highest running flow, followed by spring, autumn, and winter. In spring, summer, and autumn, the peak time for starting a run in the morning is from 5:00 to 6:00, with runs typically ending between 6:00 and 7:00. In winter, these times shift approximately an hour later, likely due to colder early morning temperatures affecting running experiences. During the evening peak, spring sees running peaks between 18:00 and 19:00, while in summer and autumn, this peak occurs about an hour later. This

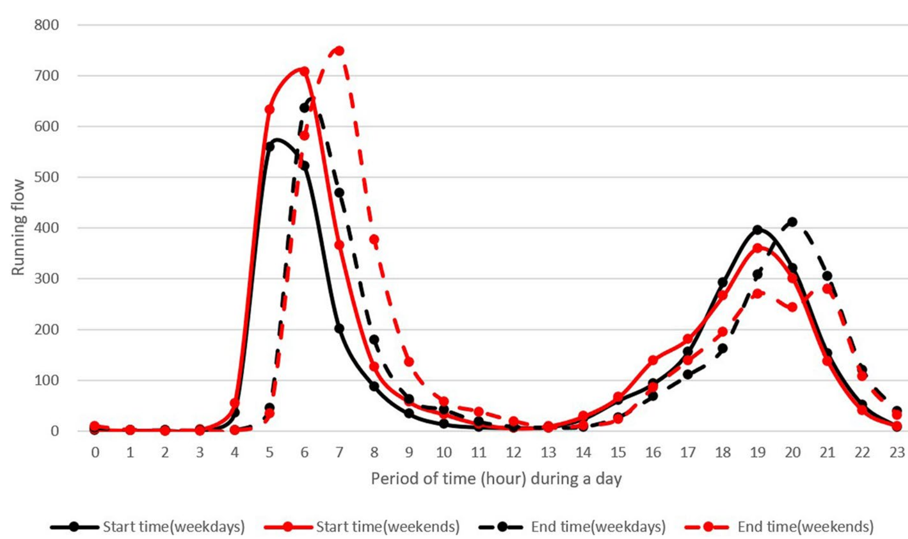

Fig. 4 Start and end times of running in green open spaces on weekdays and weekends

shift may also be influenced by temperature variations impacting the running experience.

4.1.2 Spatial and running status characteristics of running activities in green open spaces

Based on Fig. 5, the annual running activity in green open spaces varies significantly. Taohuahe Park has the lowest annual running flow with 138 records, while Aixihu Park records the highest annual running flow of 31,346 records. This spatial distribution reveals a pattern of higher activity in the east and lower activity in the west.

Regarding running distance, the average distance per run in green open spaces is 8.43 km. Nineteen green open spaces exceed this average, mostly large areas with linear waterfronts and blue landscapes, such as Aixihu Park and Qingshanhu Scenic Area. This indicates that the blue landscape may enhance the running experience for citizens.

Concerning running pace, the average pace for runners in green open spaces is 6 min and 50 s per kilometer. Most green open spaces have an average pace ranging from 6 to 8 min per kilometer. Conversely, smaller green open spaces tend to have slower average paces. This could be due to poor road conditions or because they attractive to low pace runner. Specifically, in Nanchang City, larger green open spaces often have well-developed, wide, and flat greenway systems with high safety, which are specifically designed for running sports, such as Aixi Lake Park, Ganjiang Citizen Park, Yaohu Park, Xianghu Park etc. (Cai et al.,2022; Kuang, 2019; Wenjing et al.,2015; Ye, 2016). Therefore, these spaces have better road conditions and are more attractive to frequent runners. However, smaller green open spaces such as Bayi Park, People's Park, Ruziting Park, Xianshihu Park, etc. do not have dedicated running lanes and have many steps on the way, resulting in poor road conditions. At the same time, the facilities in these spaces are more attractive to elderly people with slower running paces (Li et al., 2023; Moore et al., 2019; Zhang et al., 2020).

In terms of running time, the average duration per run in green open spaces is 55 min. There are 24 green open spaces where the average running time exceeds 55 min, mainly larger parks and linear waterfront green spaces.

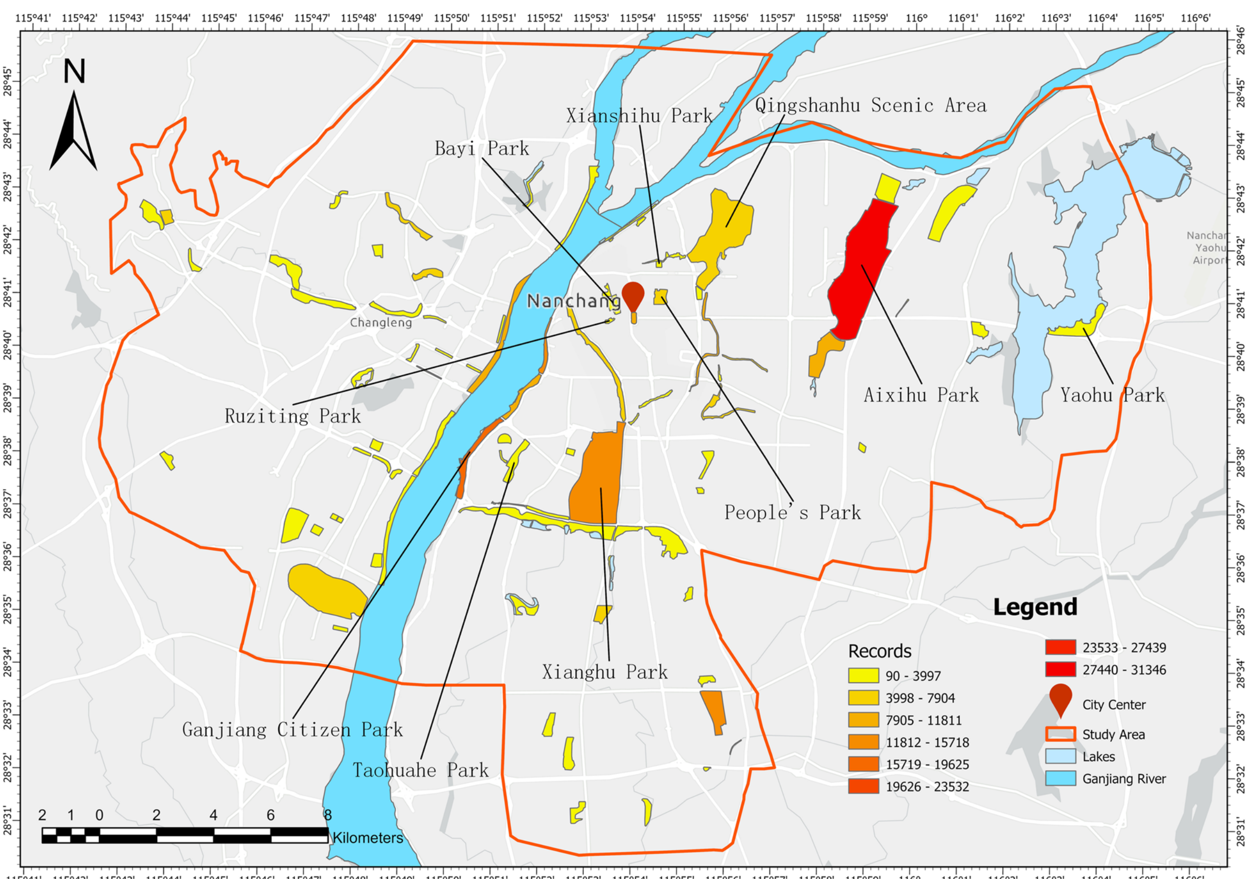

Fig. 5 Green open space running flow chart. Note: The labeled green open spaces will be discussed in the following part

This indicates that users of larger green open spaces tend to spend more time running.

4.2 Correlation analysis between green open space environment and running

Using the indicator system outlined in Sect. 2, correlation analysis is conducted to examine the correlation between the dependent variables (running flow, running distance, running time, running pace) and the indicator system. The Pearson correlation coefficient is used to quantify the strength of the relationship between the environmental characteristics of green open spaces and running performance.

The calculation of Pearson correlation coefficient is based on the covariance and standard deviation of two variables. Covariance measures how two variables change together, while standard deviation measures the degree of dispersion of each variable. The Pearson correlation coefficient is normalized by covariance to a value between -1 and 1, making it easier to interpret. For its result r, $r=1$ indicates a complete positive correlation, meaning that as one variable increases, the other variable also increases proportionally. $r=-1$ indicates complete negative correlation, meaning that one variable increases while the other variable decreases proportionally. $r=0$ indicates no linear correlation, meaning there is no linear relationship between the two variables.

Table 2 presents the correlation coefficients for each variable's independent effect on the dependent variables.

According to Table 2, It can be found that total area, road length, and number of public toilets have the greatest positive impact on running flow, while NDVI has a significant negative impact. For running distance, road length and total area have a significant positive impact, while bus station density nearby, road density, and commercial land ratio nearby have a significant negative effect. As for running time, the factors that have the greatest positive effect on it are road length, number of public toilets, and total area, while the density of nearby bus stops has a negative impact. Meanwhile, LSI, Average night light, and Building land ratio nearby cannot have a significant impact on any dependent variable, therefore they will no longer be considered in subsequent analysis.

For running space, only road width, nearby residential land ratio and commercial land ratio show correlations with running pace. Typically, wider roads facilitate faster running paces. However, in this study, road width is negatively correlated with running pace, possibly because areas with wider roads have higher pedestrian traffic, causing runners to slow down and avoid other people. Running pace is likely more related to the individual abilities of runners rather than the surrounding environment. Therefore, when exploring the elements influencing running pace, it is necessary to focus on individual characteristics and

Table 2 Correlation analysis of internal and external objective elements

Type	Independent	Running flow *Y1*	Running distance *Y2*	Running time *Y3*	Running pace *Y4*
Internal	Total area *X1*	0.665[c]	0.471[c]	0.436[c]	-0.144
Internal	NDVI *X2*	-0.25[b]	-0.225[b]	-0.203[a]	0.168
Internal	Blue landscape *X3*	0.199[a]	0.177	0.197[a]	-0.037
Internal	LSI *X4*	-0.106	-0.077	-0.129	-0.059
Internal	Road length *X5*	0.664[c]	0.484[c]	0.461c	-0.129
Internal	Road density *X6*	-0.09	-0.277[c]	-0.252[b]	0.135
Internal	Road width *X7*	0.261[c]	0.184[a]	0.131	-0.213[a]
Internal	Synthetic surface track *X8*	0.224[a]	0.236[b]	0.31c	0.165
Internal	Number of public toilets *X9*	0.585[c]	0.450[c]	0.445[c]	-0.056
External	Density of intersections nearby *X10*	0.064	-0.201	-0.217[b]	0.043
External	Road density nearby *X11*	0.04	-0.208[a]	0.352[c]	0.172
External	Residential land ratio nearby *X12*	0.13	-0.09	0.334[c]	0.228[a]
External	Population density nearby *X13*	0.23a	0.073	0.302[c]	0.092
External	Distance to the city center *X14*	-0.227[a]	0.103	0.069	-0.06
External	Average night light *X15*	0.127	-0.168	-0.118	0.202
External	Density of bus stops nearby *X16*	-0.223[a]	-0.339[b]	-0.307[c]	0.139
External	Building land ratio nearby *X17*	0.118	-0.053	-0.054	-0.053
External	Commercial land ratio nearby *X18*	-0.073	-0.257[c]	-0.22[b]	0.176[b]

[a], [b], [c] represent significance levels of 1%, 5%, and 10%, respectively

abilities of runners, rather than solely on external environmental elements.

4.3 Multilinearity test

In the following analysis, in order to make the results comparable, all independent variables are first normalized.

After conducting a correlation analysis, it was found that some indicators did not show a significant correlation with the dependent variables. Specifically, running pace was not significantly correlated with objective environmental elements. Therefore, in multilinearity test and ridge regression, running pace will be excluded, and only the dependent variables of running flow, running distance, and running time and the normalized independent variables with the significant level better than 10% will be retained. Before proceeding with the regression analysis, it is necessary to verify whether multicollinearity exists between the different variables.

4.3.1 Results of multiple linear regression model for running flow

A multiple linear regression model is established using the internal and external environmental elements of green open spaces as the independent variables, with running flow as the dependent variable. The results are shown in Table 3. The R^2 value of the model is 0.518, indicating that the variables in the table can explain 51.8% of the changes in running flow, suggesting a good model fit. The formula of the model is as follows:

$$y = 0.101 + 0.363 * X1 - 0.075 * X2 - 0.003 * X3 + 0.158 * X5 + 0.075 * X7 + 0.031 * X8 + 0.027 * X9 + 0.035 * X13 - 0.103 * X14 - 0.03 * X16$$

In the equation, y represents normalized running flow. $X1$, $X2$,..., $X16$ represent normalized independent variables. The specific correspondence between the code and independent variables is shown in Table 1.

The elements that have a positive impact on running flow, in descending order of impact, are total area, road length, road width, population density nearby, synthetic surface track and number of public toilets, while the elements with negative effects are distance to the city center, NDVI, density of bus stops nearby and blue landscape.

4.3.2 Results of multiple linear regression model for running distance

Based on the multiple linear regression model for running flow, the same environmental element indicators were maintained, running distance was used as the dependent variable, and the results are presented in Table 4. The synthetic surface track and total nearby population are positively correlated with running distance, while road density, NDVI, and nearby building land area are negatively correlated. The model formula is as follows:

$$y = 0.49 + 0.026 * X1 - 0.093 * X2 + 0.302 * X5 - 0.145 * X6 + 0.024 * X7 + 0.03 * X8 + 0.223 * X9 - 0.143 * X11 - 0.166 * X16 - 0.081 * X18$$

In the equation, y represents normalized running distance. $X1$, $X2$,..., $X18$ represent normalized independent variables. The specific correspondence between the code and independent variables is shown in Table 1.

Table 3 Results of the multiple linear regression model for running flow

	Non-standardized coefficient		Standardized coefficient	t	P	VIF	R^2	Adjust R^2	F
	B	SD	Beta						
Constant	0.101	0.076	-	1.316	0.192	-	0.518	0.455	F = 8.276 P = 0.000***
Total area *X1*	0.363	0.204	0.378	1.778	0.079*	7.228			
NDVI *X2*	-0.075	0.070	-0.100	-1.072	0.287	1.400			
Blue landscape *X3*	-0.003	0.036	-0.008	-0.081	0.936	1.547			
Road length *X5*	0.158	0.211	0.184	0.747	0.457	9.691			
Road width *X7*	0.075	0.067	0.097	1.115	0.268	1.217			
Synthetic surface track *X8*	0.031	0.027	0.102	1.158	0.250	1.236			
Number of public toilets *X9*	0.027	0.123	0.031	0.222	0.825	3.168			
Population density nearby *X13*	0.035	0.078	0.042	0.454	0.651	1.381			
Distance to the city center *X14*	-0.103	0.053	-0.186	-1.934	0.057*	1.475			
Density of bus stops nearby *X16*	-0.030	0.092	-0.028	-0.332	0.741	1.150			
Dependent variable: Running flow									

***, * represent significance levels of 1%, 5%, and 10%, respectively

Table 4 Linear regression model results for running distance

	Non-standardized coefficient		Standardized coefficient	t	*P*	VIF	R^2	Adjust R^2	F
	B	SD	Beta						
Constant	0.490	0.085	-	5.788	0.000***	-	0.325	0.237	*F* = 3.702 *P* = 0.000 ***
Total area *X1*	0.026	0.388	0.018	0.066	0.947	8.376			
NDVI *X2*	-0.093	0.112	-0.084	-0.828	0.410	1.168			
Road length *X5*	0.302	0.385	0.236	0.784	0.435	10.323			
Road density *X6*	-0.145	0.245	-0.079	-0.591	0.556	2.040			
Road width *X7*	0.024	0.116	0.021	0.204	0.839	1.161			
Synthetic surface track *X8*	0.030	0.046	0.067	0.662	0.510	1.186			
Number of public toilets *X9*	0.223	0.217	0.171	1.026	0.308	3.157			
Road density nearby *X11*	-0.143	0.160	-0.103	-0.895	0.373	1.511			
Density of bus stops nearby *X16*	-0.166	0.181	-0.103	-0.920	0.361	1.432			
Commercial land ratio nearby *X18*	-0.081	0.118	-0.075	-0.687	0.494	1.350			
Dependent variable: Running distance									

*** represent significance levels of 1%, 5%, and 10%, respectively

The elements that have a positive impact on running distance are road length, number of public toilets, synthetic surface track, total area and road width, sequentially, while the elements with negative effects are density of bus stops nearby, road density, road density nearby, NDVI, and commercial land ratio nearby.

4.3.3 *Multiple linear regression model results for running time*

Based on the multiple linear regression model analyzing running flow, while keeping the environmental indicators unchanged, running time was selected as the dependent variable, with results detailed in Table 5. The model formula is as follows:

$$y = 0.462 - 0.101 * X1 - 0.067 * X2 + 0.004 * X3 + 0.295 * X5 - 0.148 * X6 + 0.065 * X8 + 0.309 * X9 - 0.266 * X10 + 0.021 * X11 + 0.013 * X12 + 0.071 * X13 - 0.112 * X16 - 0.038 * X18$$

Table 5 Linear regression model results for running time

	Non-standardized coefficient		Standardized coefficient	t	P	VIF	R^2	Adjust R^2	F
	B	SD	Beta						
Constant	0.462	0.113	-	4.089	0.000***	-	0.329	0.211	*F* = 2.794 *P* = 0.003***
Total area *X1*	-0.101	0.394	-0.072	-0.257	0.798	8.592			
NDVI *X2*	-0.067	0.124	-0.061	-0.538	0.592	1.420			
Blue landscape *X3*	0.004	0.063	0.007	0.060	0.952	1.567			
Road length *X5*	0.295	0.395	0.234	0.746	0.458	10.824			
Road density *X6*	-0.148	0.249	-0.082	-0.596	0.553	2.092			
Synthetic surface track *X8*	0.065	0.047	0.147	1.376	0.173	1.260			
Number of public toilets *X9*	0.309	0.221	0.240	1.401	0.165	3.247			
Density of intersections nearby *X10*	-0.266	0.149	-0.205	-1.783	0.079*	1.460			
Road density nearby *X11*	0.021	0.180	0.015	0.117	0.907	1.904			
Residential land ratio nearby *X12*	0.013	0.105	0.014	0.127	0.899	1.313			
Population density nearby *X13*	0.071	0.141	0.058	0.507	0.614	1.452			
Density of bus stops nearby *X16*	-0.112	0.185	-0.071	-0.604	0.547	1.502			
Commercial land ratio nearby *X18*	-0.038	0.129	-0.035	-0.292	0.771	1.605			
Dependent variable: Running time									

***, * represent significance levels of 1%, 5%, and 10%, respectively

In the equation, y represents normalized running time. $X1$, $X2$,..., $X18$ represent normalized independent variables. The specific correspondence between the code and independent variables is shown in Table 1.

These elements have a positive impact on running time: number of public toilets, road length, population density nearby, synthetic surface track, road density nearby, residential land ratio nearby and blue landscape, sequentially, while these elements have negative effects: density of intersections nearby, road density, density of bus stops nearby, total area, NDVI, and commercial land ratio nearby.

The complete linear regression model results of standardized coefficients are shown in Table 6. In Tables 3–5, there are cases where the VIF approaches or exceeds 10, indicating a multicollinearity between the independent variables. So, ridge regression analysis is needed to address these issues.

4.4 Ridge regression model results and analysis of various elements

4.4.1 Ridge regression model analysis results of running flow

According to Ridge trace, the regression coefficient stabilizes when the K value is 0.193. At this point, the R^2 value is 0.513, the F value is 8.096, and the P value is 0.000***, indicating an average model fit. According to Table 7, The formula for the non-standardized ridge regression model is:

$$y = 0.085 + 0.261 * X1 - 0.065 * X2 + 0.002 * X3 + 0.177 * X5 + 0.069 * X7 + 0.026 * X8 + 0.074 * X9 + 0.043 * X13 - 0.08 * X14 - 0.033 * X16$$

In the equation, y represents normalized running flow. $X1$, $X2$,..., $X16$ represent normalized independent variables. The specific correspondence between the code and independent variables is shown in Table 1.

According to the standardized regression coefficients in Table 7, the environmental elements that have a positive impact on running flow are as follows: total area (0.272), road length (0.207), road width (0.09), number of public toilets (0.085), synthetic surface track (0.084), population density nearby (0.051). The following have negative impacts: distance to the city center (-0.145), NDVI (-0.087), density of bus stops nearby (-0.031). And blue landscape has almost no impact on running flow.

4.4.2 Ridge regression model analysis results of running distance

According to Ridge trace, the regression coefficient stabilizes when the K value is 0.244. At this point, the R^2 value is 0.322, the F value is 3.652, and the P value is 0.001***. According to Table 8, The formula for the non-standardized ridge regression model is:

$$y = 0.477 + 0.142 * X1 - 0.078 * X2 + 0.191 * X5 - 0.12 * X6 + 0.031 * X7 + 0.028 * X8 + 0.194 * X9 - 0.123 * X11 - 0.158 * X16 - 0.079 * X18$$

In the equation, y represents normalized running distance. $X1$, $X2$,..., $X18$ represent normalized independent variables. The specific correspondence between the code and independent variables is shown in Table 1.

Table 6 Comprehensive linear regression model results of standardized coefficients

Type	Independent	Running flow *Y1*	Running distance *Y2*	Running time *Y3*
Internal	Total area *X1*	0.378*	0.018	-0.072
Internal	NDVI *X2*	-0.100	-0.084	-0.061
Internal	Blue landscape *X3*	-0.008	-	0.007
Internal	LSI *X4*	-	-	-
Internal	Road length *X5*	0.184	0.236	0.234
Internal	Road density *X6*	-	-0.079	-0.082
Internal	Road width *X7*	0.097	0.021	-
Internal	Synthetic surface track *X8*	0.102	0.067	0.147
Internal	Number of public toilets *X9*	0.031	0.171	0.240
External	Density of intersections nearby *X10*	-	-	-0.205*
External	Road density nearby *X11*	-	-0.103	0.015
External	Residential land ratio nearby *X12*	-	-	0.014
External	Population density nearby *X13*	0.042	-	0.058
External	Distance to the city center *X14*	-0.186*	-	-
External	Average night light *X15*	-	-	-
External	Density of bus stops nearby *X16*	-0.028	-0.103	-0.071
External	Building land ratio nearby *X17*	-	-	-
External	Commercial land ratio nearby *X18*	-	-0.075	-0.035

* represent significance levels of 1%, 5%, and 10%, respectively

Table 7 Ridge regression model results for running flow

K = 0.193	Non-standardized coefficient		Standardized coefficient:	t	*P*	R^2	Adjust R^2	F
	B	SD	Beta					
Constant	0.085	0.053	-	1.601	0.113	0.513	0.449	8.096 *P* = 0.000***
Total area X1	0.261	0.070	0.272	3.713	0.000***			
NDVI X2	-0.065	0.052	-0.087	-1.243	0.217			
Blue landscape X3	0.002	0.026	0.006	0.086	0.932			
Road length X5	0.177	0.060	0.207	2.941	0.004***			
Road width X7	0.069	0.054	0.090	1.295	0.199			
Synthetic surface track X8	0.026	0.021	0.084	1.219	0.226			
Number of public toilets X9	0.074	0.069	0.085	1.072	0.287			
Population density nearby X13	0.043	0.059	0.051	0.720	0.473			
Distance to the city center X14	-0.080	0.040	-0.145	-2.031	0.045**			
Density of bus stops nearby X16	-0.033	0.075	-0.031	-0.445	0.657			
Dependent variable: running flow								

***, ** represent significance levels of 1%, 5%, and 10%, respectively

Table 8 Ridge regression model results for running distance

K = 0.244	Non-standardized coefficient		Standardized coefficient:	t	*P*	R^2	Adjust R^2	F
	B	SD	Beta					
Constant	0.477	0.069	-	6.921	0.000***	0.322	0.234	3.652 *P* = 0.001***
Total area *X1*	0.142	0.107	0.099	1.331	0.187			
NDVI *X2*	-0.078	0.085	-0.071	-0.921	0.360			
Road length *X5*	0.191	0.091	0.149	2.103	0.038**			
Road density *X6*	-0.120	0.147	-0.065	-0.817	0.416			
Road width *X7*	0.031	0.089	0.027	0.345	0.731			
Synthetic surface track *X8*	0.028	0.035	0.063	0.814	0.418			
Number of public toilets *X9*	0.194	0.110	0.149	1.771	0.080*			
Road density nearby *X11*	-0.123	0.110	-0.089	-1.123	0.265			
Density of bus stops nearby *X16*	-0.158	0.129	-0.098	-1.230	0.222			
Commercial land ratio nearby *X18*	-0.079	0.086	-0.072	-0.910	0.365			
Dependent variable: running distance								

***, **, * represent significance levels of 1%, 5%, and 10%, respectively

According to the standardized regression coefficients in Table 8, the environmental elements that have a positive impact on running flow are as follows: road length (0.149), number of public toilets (0.149), total area (0.099), synthetic surface track (0.063), road width (0.027). The following have negative impacts: density of bus stops nearby (-0.098), road density nearby (-0.089), commercial land ratio nearby (-0.072), NDVI (-0.071), road density (-0.065).

4.4.3 Ridge regression model analysis results of running time

According to Ridge trace, the regression coefficient stabilizes when the K value is 0.260. At this point, the R^2 value is 0.322, the F value is 2.703, and the *P* value is 0.004***. According to Table 8, The formula for the non-standardized ridge regression model is:

$$y = 0.454 + 0.092 * X1 - 0.046 * X2 + 0.011 * X3 + 0.168 * X5 - 0.092 * X6 + 0.051 * X8 + 0.229 * X9 - 0.197 * X10 - 0.02 * X11 + 0.013 * X12 + 0.067 * X13 - 0.119 * X16 - 0.054 * X18$$

In the equation, y represents normalized running time. $X1$, $X2$,..., $X18$ represent normalized independent variables. The specific correspondence between the code and independent variables is shown in Table 1.

According to the standardized regression coefficients in Table 9, the environmental elements that have a positive

Table 9 Ridge regression model results for running time

K = 0.260	Non-standardized coefficient		Standardized coefficient:	t	*P*	R^2	Adjust R^2	F
	B	SD	Beta					
Constant	0.454	0.082	-	5.535	0.000***	0.322	0.203	2.703 *P* = 0.004***
Total area *X1*	0.092	0.103	0.065	0.891	0.375			
NDVI *X2*	-0.046	0.086	-0.042	-0.539	0.591			
Blue landscape *X3*	0.011	0.042	0.021	0.263	0.793			
Road length *X5*	0.168	0.087	0.133	1.931	0.057*			
Road density *X6*	-0.092	0.146	-0.051	-0.633	0.528			
Synthetic surface track *X8*	0.051	0.035	0.114	1.461	0.148			
Number of public toilets *X9*	0.229	0.107	0.178	2.144	0.035**			
Density of intersections nearby *X10*	-0.197	0.104	-0.152	-1.900	0.061*			
Road density nearby *X11*	-0.020	0.111	-0.014	-0.177	0.860			
Residential land ratio nearby *X12*	0.013	0.075	0.014	0.177	0.860			
Population density nearby *X13*	0.067	0.097	0.054	0.688	0.493			
Density of bus stops nearby *X16*	-0.119	0.127	-0.075	-0.939	0.350			
Commercial land ratio nearby *X18*	-0.054	0.087	-0.050	-0.619	0.538			
Dependent variable: running time								

***, **, * represent significance levels of 1%, 5%, and 10%, respectively

impact on running flow are as follows: number of public toilets (0.178), road length (0.133), synthetic surface track (0.114), total area (0.065), population density nearby (0.054), blue landscape (0.021), residential land ratio nearby (0.014). The following have negative impacts: density of intersections nearby (-0.152), density of bus stops nearby (-0.075), road density (-0.051), commercial land ratio nearby (-0.050), NDVI (-0.042), road density nearby (-0.014).

4.4.4 Ridge regression model analysis results summary

The complete ridge regression model results of standardized coefficients are shown in Table 10.

Table 10 Comprehensive ridge regression model results of standardized coefficients

Type	Independent	Running flow *Y1*	Running distance *Y2*	Running time *Y3*
Internal	Total area *X1*	0.272[c]	0.099	0.065
Internal	NDVI *X2*	-0.087	-0.071	-0.042
Internal	Blue landscape *X3*	-0.006	-	0.021
Internal	LSI *X4*	-	-	-
Internal	Road length *X5*	0.207[c]	0.149[b]	0.133[a]
Internal	Road density *X6*	-	-0.065	-0.051
Internal	Road width *X7*	0.090	0.027	-
Internal	Synthetic surface track *X8*	0.084	0.063	0.114
Internal	Number of public toilets *X9*	0.085	0.149[a]	0.178[b]
External	Density of intersections nearby *X10*	-	-	-0.152[a]
External	Road density nearby *X11*	-	-0.089	-0.014
External	Residential land ratio nearby *X12*	-	-	0.014
External	Population density nearby *X13*	0.051	-	0.054
External	Distance to the city center *X14*	-0.145[c]	-	-
External	Average night light *X15*	-	-	-
External	Density of bus stops nearby *X16*	-0.031	-0.098	-0.075
External	Building land ratio nearby *X17*	-	-	-
External	Commercial land ratio nearby *X18*	-	-0.072	-0.050

[a, b, c] represent significance levels of 1%, 5%, and 10%, respectively

From Table 10, each independent variable has a consistent impact on the three running performance elements. That is to say, the impact of a single independent variable on the three elements is similar. Total area has a positive effect on all three running performance elements, with running flow being the most affected. As mentioned in Sect. 4.1.2, in Nanchang city, large green open spaces often have more complete running infrastructure. Larger green open spaces also have bigger space and more running route to create conditions for longer running times and distances. However, NDVI has a moderate negative effect on the three running performance elements. It may be due to vegetation-covered areas being unsuitable for running activities, thus reducing the space's capacity for runners. Additionally, excessive vegetation might attract mosquitoes, posing a threat to runners. Higher population density, contribute to increased runner presence in green open spaces. Meanwhile, the blue landscape has little effect on the three elements except for running time, as it may delight runners and prolong their running time. The length of the road is like the total area, has a significant positive impact on running performance. Road density is not conducive to running performance, possibly due to the high density of roads disrupting the complete running experience, and high road density in Nanchang City's green open spaces mainly reflects more small paths, but most runners prefer the main roads. Next, road width, synthetic running tracks, and the number of public toilets are all beneficial for running performance, indicating that improved public infrastructure is necessary to enhance residents' running performance. Especially, the number of public toilets plays the most important role among various environmental elements in increasing running distance and running time as going to the toilet is necessary for runners during long-time or long-distance running.

For the external elements, the density of surrounding roads and intersections nearby is negatively correlated with running distance and time, with a close relationship between intersection density and running time. This may be because they bring segmentation and fragmentation and disrupt the continuity of running, as runners might be distracted or interrupted by traffic lights, causing them to stop or end their run. On the contrary, the proportion of residential land and population density in the surrounding area are positively correlated with running performance. As the distance from the city center increases, the running flow tends to rapidly decrease. That is because green open spaces closer to the city center are generally surrounded by developed earlier and more crowded areas, which usually means higher population density and housing prices, as well as better infrastructure such as roads. They bring more runners, while high housing prices mean that residents living in these areas are more affluent to afford housing prices, so they have more time to exercise (Park et al.,2017). Unlike existing research on walking, according to the results of this study, a high proportion of commercial areas and a high density of bus stop have a negative impact on running distance and time (Sugiyama et al., 2019). The crowded environment in these areas may have weakened residents' running performance.

The results of the ridge regression show a high degree of consistency with the results of the multivariate linear test in Sect. 4.3, and the significance has been improved. However, there are also differences. Firstly, the ridge regression results show more conservative coefficient estimates due to the presence of regularization terms, which means that the coefficients are generally smaller. Regarding the coefficient changes of each independent variable, in ridge regression, the positive impact of total area on running distance is amplified, making it the third most significant positive factor affecting running distance, and it has a positive effect on running time instead of a negative effect, which is more in line with the actual situation. The impact of blue landscape on running time and the impact of the number of public toilets on running flow have also been amplified by more than double, while the impact of road density nearby on running time has changed from positive to negative, and the other standardization coefficients have decreased or changed little.

5 Discussions and conclusion

This study primarily focuses on analyzing the characteristics of running in green open spaces in Nanchang City, highlighting the impact of both internal and external environmental elements on running performance. By using analysis tools such as SPSS and GIS, the study investigates the relationship between internal and external environments of green open spaces and running flow, running distance, and running time. A regression model is developed to objectively evaluate how green open spaces influence running. The key conclusions are as follows:

In terms of temporal characteristics, running flow peaks in the morning exceed those in the evening. During the morning peak, female participation surpasses that of males, whereas the trend reverses during the evening. On weekends, the morning peak starts an hour later than on weekdays, with overall weekend running flow surpassing that of weekdays. Summer sees the highest running flow, followed by spring, autumn, and winter. In terms of spatial characteristics, running activity shows an eastern bias with lower participation in the west. High running flow and performance are mainly concentrated in large parks and linear waterfront areas, while few green open spaces exhibit these characteristics.

Among the internal and external environmental elements in green open spaces affecting the running behavior of Nanchang residents, significant correlations are as follows: Total area (0.272) and road length (0.207) positively correlated with the running flow, while distance to the city center (-0.145) and NDVI (-0.087) negatively correlated with the running flow. Running distance shows positive correlations with the number of public toilets (0.149) and road length (0.149), while it is negatively correlated with density of bus stops nearby (-0.098) and road density nearby (-0.089). Running time is positively correlated with number of public toilets (0.178) and road length (0.133), and negatively correlated with density of intersections nearby (-0.152) and density of bus stops nearby (-0.075). Running pace, however, appears unaffected by the internal and external environmental elements of green open spaces.

In summary, for the improvement of the three running performance indicators of running flow, running time, and running distance in green open spaces, factors of total area, road length, population density nearby, number of public toilets, synthetic surface tracks, road width, and residential land ratio nearby play a positive role, while road density, NDVI, road density nearby, intersection density nearby, distance from the city center, bus stop density nearby, and the commercial land ratio nearby have a negative effect. Therefore, in order to attract more citizens to run in green open spaces and improve their running performance to ensure their health, in addition to building and maintaining more green open spaces, the government and relevant departments should also pay attention to those. For the existing green open spaces, they should expand their area as much as possible, provide facilities that are conducive to running, such as building plastic tracks and more public toilets, widening existing roads, convert some green spaces into spaces for running and repel mosquitoes, and add water landscapes to green open spaces to attract runners. For new green open spaces, it is advisable to choose areas with a high population density, a high proportion of residential areas and a low proportion of commercial areas, and they'd better get close to the city center and have few bus stops. At the same time, efforts should be made to expand their area and plan for comprehensive running infrastructure. However, efforts to improve running speed in improving green open spaces will have little effect, and should start with the runner's own conditions. Finally, in response to the low evening running traffic among females, measures such as increasing lighting should be taken to enhance the safety of green open spaces at night to address females' concerns.

Acknowledgements
Not applicable.

Authors' contributions
Tianwei Fang wrote and revised, the initial manuscript, performed the visualization, and participated in experiments and data analysis. Linguo Zhou designed the experimental method, collected data, organized the data, conducted experiments, and analyzed the data. Zhenrao Cai performed ridge regression analysis and rewrite relevant sections during revising. Zhijin Tan participated in writing the initial manuscript. Chao Chen participated in the revision of the initial manuscript. Jiali Zheng added references, revised wording, and checked for grammar and spelling errors during the manuscript revision process. Chaoyang Fang conceptualized the experimental approach.

Funding
Graduate innovation fund project of Jiangxi Provincial Department of Education (No. YJS2023013).

Data availability
The data that support the findings of this study are available on request from the corresponding author, Chaoyang Fang, upon reasonable request.

Declarations

Consent for publication
This paper does not contain any individual person's data in any form.

Competing interests
Funding: Graduate innovation fund project of Jiangxi Provincial Department of Education (No. YJS2023013)

Received: 5 July 2024 Revised: 30 September 2024 Accepted: 20 October 2024
Published online: 04 November 2024

References

Abdelkarim S.B., Ahmad A.M., Ferwati S., & Naji K.(2023). Urban Facility Management Improving Livability through Smart Public Spaces in Smart Sustainable Cities. *Sustainability*, 15(23):16257. https://doi.org/10.3390/su152316257

Akpinar A., & Cankurt M.(2017). How are characteristics of urban green space related to levels of physical activity: Examining the links. *Indoor and Built Environment*, 26(8):1091–1101. https://doi.org/10.1177/1420326x16663289

Bankoski A., Harris T.B., McClain J.J., Brychta R.J., Caserotti P., Chen K.Y., et al. (2011). Sedentary Activity Associated With Metabolic Syndrome Independent of Physical Activity. *Diabetes Care*, 34(2):497–503. https://doi.org/10.2337/dc10-0987

Besser, L. M., & Mitsova, D. P. (2021). Neighborhood Green Land Cover and Neighborhood-Based Walking in U.S. Older Adults. American Journal of Preventive Medicine, 61(1), e13–e20. https://doi.org/10.1016/j.amepre.2021.01.013

Bodin M., & Hartig T.(2003). Does the outdoor environment matter for psychological restoration gained through running? *Psychology of Sport and Exercise*, 4(2):141–153. https://doi.org/10.1016/S1469-0292(01)00038-3

Borgers J., Vanreusel B., Vos S., Forsberg P., & Scheerder J.(2016). Do light sport facilities foster sports participation? A case study on the use of bark running tracks. *International Journal of Sport Policy and Politics*, 8(2):287–304. https://doi.org/10.1080/19406940.2015.1116458

Christiansen, L. B., Cerin, E., Badland, H., Kerr, J., Davey, R., Troelsen, J., van Dyck, D., Mitáš, J., Schofield, G., Sugiyama, T., Salvo, D., Sarmiento, O. L., Reis, R., Adams, M., Frank, L., & Sallis, J. F. (2016). International comparisons of the associations between objective measures of the built environment and transport-related walking and cycling: IPEN adult study. *Journal of Transport & Health, 3*(4), 467–478. https://doi.org/10.1016/j.jth.2016.02.010

Coombes E., Jones A.P., & Hillsdon M.(2010). The relationship of physical activity and overweight to objectively measured green space accessibility and use. *Social Science & Medicine*, 70(6):816–822. https://doi.org/10.1016/j.socscimed.2009.11.020

De Jong K., Albin M., Skärbäck E., Grahn P., & Björk J.(2012). Perceived green qualities were associated with neighborhood satisfaction, physical activity, and general health: Results from a cross-sectional study in suburban and rural Scania, southern Sweden. *Health & Place*, 18(6):1374–1380. https://doi.org/10.1016/j.healthplace.2012.07.001

Deelen I., Janssen M., Vos S., Kamphuis C.B.M., & Ettema D.(2019). Attractive running environments for all? A cross-sectional study on physical environmental characteristics and runners' motives and attitudes, in relation to the experience of the running environment. *BMC Public Health*, 19(1):366. https://doi.org/10.1186/s12889-019-6676-6

Dinnie, E., Brown, K. M., & Morris, S. (2013). Reprint of "Community, cooperation and conflict: Negotiating the social well-being benefits of urban greenspace experiences." *Landscape and Urban Planning, 118*, 103–111. https://doi.org/10.1016/j.landurbplan.2013.07.011

Ettema D.(2016). Runnable Cities:How Does the Running Environment Influence Perceived Attractiveness, Restorativeness, and Running Frequency? *Environment and Behavior*, 48(9):1127–1147. https://doi.org/10.1177/0013916515596364

Gao, B., Huang, Q., He, C., Sun, Z., & Zhang, D. (2016). How does sprawl differ across cities in China? A multi-scale investigation using nighttime light and census data. *Landscape and Urban Planning, 148*, 89–98. https://doi.org/10.1016/j.landurbplan.2015.12.006

Gao W., Sanna M., Chen Y.-H., Tsai M.-K., & Wen C.-P.(2024). Occupational Sitting Time, Leisure Physical Activity, and All-Cause and Cardiovascular Disease Mortality. *JAMA Network Open*, 7(1):e2350680-e2350680. https://doi.org/10.1001/jamanetworkopen.2023.50680

Hansmann R., Hug S.-M., & Seeland K.(2007). Restoration and stress relief through physical activities in forests and parks. *Urban Forestry & Urban Greening*, 6(4):213–225. https://doi.org/10.1016/j.ufug.2007.08.004

Harden S.R., Schuurman N., Larson H., & Walker B.B.(2024). The utility of street view imagery in environmental audits for runnability. *Applied Geography*, 162:103167. https://doi.org/10.1016/j.apgeog.2023.103167

Huang D., Jiang B., & Yuan L.(2022). Analyzing the effects of nature exposure on perceived satisfaction with running routes: An activity path-based measure approach. *Urban Forestry & Urban Greening*, 68:127480. https://doi.org/10.1016/j.ufug.2022.127480

Jiang H., Dong L., & Qiu B.(2022). How Are Macro-Scale and Micro-Scale Built Environments Associated with Running Activity? The Application of Strava Data and Deep Learning in Inner London. *ISPRS International Journal of Geo-Information*, 11(10):504. https://doi.org/10.3390/ijgi11100504

Jiao, Y., Yu, H., Wang, Z., Wei, Q., & Yu, Y. (2017). Influence of individual factors on thermal satisfaction of the elderly in free running environments. *Building and Environment, 116*, 218–227. https://doi.org/10.1016/j.buildenv.2017.02.018

Kaczynski, A. T., Koohsari, M. J., Stanis, S. A. W., Bergstrom, R., & Sugiyama, T. (2014). Association of Street Connectivity and Road Traffic Speed with Park Usage and Park-Based Physical Activity. *American Journal of Health Promotion, 28*(3), 197–203. https://doi.org/10.4278/ajhp.120711-QUAN-339

Kuang, Y. (2019). Landscape Planning and Design of Urban Forest Park Trails: A Case Study of Yaohu Country Forest Park in Nanchang City. Journal of Landscape Research, 11(6), 17–20. https://doi.org/10.16785/j.issn1943-989x.2019.6.005

Lee D.-c., Pate R.R., Lavie C.J., Sui X., Church T.S., & Blair S.N.(2014). Leisure-Time Running Reduces All-Cause and Cardiovascular Mortality Risk. *Journal of the American College of Cardiology*, 64(5):472–481. https://doi.org/10.1016/j.jacc.2014.04.058

Lee I.M., Shiroma E.J., Lobelo F., Puska P., Blair S.N., & Katzmarzyk P.T.(2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet*, 380(9838):219–229. https://doi.org/10.1016/S0140-6736(12)61031-9

Li, B., Wen, Y., Peng, B., & Tan, H. (2023). Landscape Upgrading and Transformation of Old Urban Parks from the Perspective of Place Attachment; a Case Study of Nanchang Bayi Park. Meteorological and Environmental Research, 14(2), 28–41. https://doi.org/10.19547/j.issn2152-3940.2023.02.006

Liu O.Y., & Russo A.(2021). Assessing the contribution of urban green spaces in green infrastructure strategy planning for urban ecosystem conditions and services. *Sustainable Cities and Society*, 68:102772. https://doi.org/10.1016/j.scs.2021.102772

Liu Y., Hu J., Yang W., & Luo C.(2022). Effects of urban park environment on recreational jogging activity based on trajectory data: A case of Chongqing, China. *Urban Forestry & Urban Greening*, 67:127443. https://doi.org/10.1016/j.ufug.2021.127443

Moore, J. B., Schuller, K., Cook, A., Lu, Y., Yuan, Z., & Maddock, J. E. (2019). An Observational Assessment of Park-based Physical Activity in Older Adults in Nanchang. *China. American Journal of Health Behavior, 43*(6), 1119–1128. https://doi.org/10.5993/AJHB.43.6.9

Murtagh E., & McKee D.(2012). Contribution of primary school physical education class to daily moderate-vigorous physical activity. *Journal of Science and Medicine in Sport*, 15:S91. https://doi.org/10.1016/j.jsams.2012.11.219

Ni L., Zhang D., Yang Y., & Huang J.(2024). Exploring the influence of relative attractiveness in green spaces on urban movements: A potential to kinetic energy framework. *Journal of Cleaner Production*, 434:139850. https://doi.org/10.1016/j.jclepro.2023.139850

Park, M. H., & Hwang, E. H. (2017). Effects of family affluence on the health behaviors of Korean adolescents. *Japan Journal of Nursing Science, 14*(3), 173–184. https://doi.org/10.1111/jjns.12146

Petrunoff N.A., Yi N.X., Dickens B., Sia A., Koo J., Cook A.R., et al.(2021). Associations of park access, park use and physical activity in parks with wellbeing in an Asian urban environment: a cross-sectional study. *International Journal of Behavioral Nutrition and Physical Activity*, 18(1):87. https://doi.org/10.1186/s12966-021-01147-2

Piercy K.L., Troiano R.P., Ballard R.M., Carlson S.A., Fulton J.E., Galuska D.A., et al.(2018). The Physical Activity Guidelines for Americans. *JAMA*, 320(19):2020–2028. https://doi.org/10.1001/jama.2018.14854

Schuurman N., Rosenkrantz L., & Lear S.A.(2021). Environmental Preferences and Concerns of Recreational Road Runners. *International Journal of Environmental Research and Public Health*, 18(12):6268. https://doi.org/10.3390/ijerph18126268

Sugiyama, T., Rachele, J. N., Gunn, L. D., Burton, N. W., Brown, W. J., & Turrell, G. (2019). Land use proportion and walking: Application of isometric substitution analysis. *Health & Place, 57*, 352–357. https://doi.org/10.1016/j.healthplace.2018.12.004

Theodoratou M., Dritsas I., Saltou M., Dimas V., Spyropoulos A., Nikolopoulou E., et al.(2016). Physical exercise and students' mental health. *European Psychiatry*, 33(S1):s219-s219. https://doi.org/10.1016/j.eurpsy.2016.01.533

Tian Z., Yang W., Zhang T., Ai T., & Wang Y.(2022). Characterizing the activity patterns of outdoor jogging using massive multi-aspect trajectory data. *Computers, Environment and Urban Systems*, 95:101804. https://doi.org/10.1016/j.compenvurbsys.2022.101804

Wang, Q., Hu, H., & Zhang, Y. (2011). Research on Maoershan Stand Factors and 3s Information Based on Ridge Regression Model. 2011 3RD INTERNATIONAL CONFERENCE ON ENVIRONMENTAL SCIENCE AND INFORMATION APPLICATION TECHNOLOGY ESIAT 2011, VOL 10, PT C, 10(C), 1998–2004. https://doi.org/10.1016/j.proenv.2011.09.313

Wang X. (2019). Analysis of the Physical Fitness Status of Students in Henan Province and the Biological Characteristics Influencing Sports Related Physical Fitness (Master's Thesis, Zhengzhou University). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201902&filename=1019098282.nh

Wenjing, L., & Qi, L. (2015). Ecological Landscape Design of Urban Wetland Park: A Case Study of the Aixi Lake Wetland Park in Nanchang City. Journal of Landscape Research, 7(3), 57-.

Cai Xiao,&Su Zhenyu (2022). Perceived Evaluation of Pedestrian Environment on Binjiang Greenway in Nanchang City: A Case Study of Ganjiang Citizen Park Urban Architecture, 19 (17), 190–194. https://doi.org/10.19892/j.cnki.csjz.2022.17.42

Xue, F., Gou, Z., & Lau, S. S. Y. (2017). Green open space in high-dense Asian cities: Site configurations, microclimates and users' perceptions. *Sustainable Cities and Society, 34*, 114–125. https://doi.org/10.1016/j.scs.2017.06.014

Ye, K. (2016). Exploration on Design of Waterfront Landscape and Building in Xianghu Lake Park of Nanchang City. Journal of Landscape Research, 8(6), 5-. https://doi.org/10.16785/jissn1943-989x.2016.6.002

Zhang, L., Yin, Z., Chen, C., Hu, H., & Liu, H. (2020). Accessibility Evaluation and Layout Optimization of Urban Parks in Xihu District, Nanchang City Based on Walking of the Elderly People. Journal of Landscape Research, 12(4), 53–57. https://doi.org/10.16785/j.issn1943-989x.2020.4.013

Zhang R., Liu S., Li M., He X., & Zhou C.(2021). The Effect of High-Density Built Environments on Elderly Individuals' Physical Health: A Cross-Sectional Study in Guangzhou, China. *International Journal of Environmental Research and Public Health*, 18(19):10250. https://doi.org/10.3390/ijerph181910250

Zhang, S., Liu, N., Ma, B., & Yan, S. (2024). The effects of street environment features on road running: An analysis using crowdsourced fitness tracker data and machine learning. *Environment and Planning. B, Urban Analytics and City Science*, *51*(2), 529–545. https://doi.org/10.1177/23998083231185589

Zhong Q., Li B., & Chen Y.(2022). How Do Different Urban Footpath Environments Affect the Jogging Preferences of Residents of Different Genders? Empirical Research Based on Trajectory Data. *International Journal of Environmental Research and Public Health*, 19(21):14372. https://doi.org/10.3390/ijerph192114372

Zhu L., Zhu K., & Zeng X.(2023). Evolution of landscape pattern and response of ecosystem service value in international wetland cities: A case study of Nanchang City. *Ecological Indicators*, 155:110987. https://doi.org/10.1016/j.ecolind.2023.110987

Zhu Y., Ding J., Zhu Q., Cheng Y., Ma Q., & Ji X.(2017). The Impact of Green Open Space on Community Attachment—A Case Study of Three Communities in Beijing. *Sustainability*, 9(4):560. https://doi.org/10.3390/su9040560

Publisher's Note