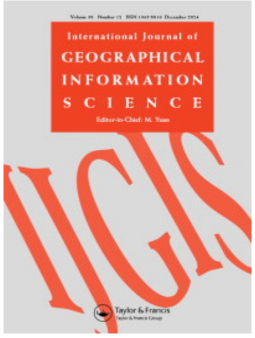

International Journal of Geographical Information Science

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tgis20

A local indicator of stratified power

Jiao Hu, Yongze Song & Tingbin Zhang

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
https://doi.org/10.1080/13658816.2024.2437811

RESEARCH ARTICLE

A local indicator of stratified power

Jiao Hu[a,b], Yongze Song[b] and Tingbin Zhang[a]

[a]College of Earth and Planet Science, Chengdu University of Technology, Chengdu, China; [b]School of Design and the Built Environment, Curtin University, Perth, Australia

ABSTRACT

Spatial stratified heterogeneity measures spatial association by power of determinant (PD) that compares variations within strata and across space. Models based on stratified heterogeneity have been extensively used across various fields to analyze PD from a spatial perspective. However, stratified heterogeneity in the local regions has not been investigated, although it can significantly influence the overall PD measurements in large-scale studies. This study proposes a local indicator of stratified power (LISP) to analyze local spatial stratified association and demonstrate how spatial stratified association changes spatially and in local regions. The LISP model was implemented to examine the local potential determinants of the thickness variations in lake-terminating glaciers of the Greater Himalayas. The results indicate that LISP can reveal spatial association at various local positions, effectively mitigating the underestimation or overestimation of PD values in local regions that are probably missed in global spatial stratified association models. The developed LISP model provides a reliable methodological framework for exploring local spatial associations and identifying local determinants in broad fields.

ARTICLE HISTORY
Received 7 August 2024
Accepted 1 December 2024

KEYWORDS
Spatial stratified heterogeneity; local power; spatial association; local determinants; glacier

1. Introduction

Spatial association is one of the most widely studied issues in spatial data analysis (Páez *et al.* 2002). Accurately quantifying the spatial association between geographic variables can promote the advancement of spatial statistical inference (Wang *et al.* 2020b, Song 2022, Zhang *et al.* 2024a), and has been widely utilized in the analysis of spatiotemporal issues across various domains such as ecology, environmental science, geology, public health, economics, and architecture (Ben-Moshe and Itzkovitz 2019, Song *et al.* 2020, Harvey and O'Neale 2024, Qian *et al.* 2024).

Spatial association is generally examined using three categories of spatial approaches, including local, scale-based, and stratified association models. The local-based spatial association primarily investigates localized clusters of regions with similar geographic attributes, local autocorrelation, and local heterogeneity. Techniques, such as the local indicators of spatial association (LISA) and Getis-Ord Gi are effective for assessing whether geographic attributes exhibit spatial clustering (Anselin 1995, Ord

CONTACT Yongze Song yongze.song@curtin.edu.au

and Getis 1995, Chen *et al*. 2023, Seoane *et al*. 2023). Local heterogeneity emphasizes fine-scale differences within specific areas, while other types (such as stratified, scale, and generalized heterogeneity) focus on broader variations across different groups, scales, or the entire study area (Luo *et al*. 2023, Zhang *et al*. 2024b). For example, geographically weighted regression (GWR) and its enhancements are used to measure the localized geographical heterogeneity effects (Huang *et al*. 2010, Fotheringham *et al*. 2017, Ge *et al*. 2017, Majumder *et al*. 2023). The spatial scaling association primarily aims to characterize the complex scaling structures of geographic attributes (Jiang 2015). For instance, the ht-index is used to quantify spatial scaling association (Jiang and Yin 2014), often demonstrating a scaling law where smaller geographic features are more numerous than larger ones. Lastly, spatial stratified association, such as spatial stratified heterogeneity, explores associations between geographic variables by comparing intra-stratum variance with inter-stratum variance (Wang *et al*. 2016).

Spatial stratification is the process of dividing geographic space into strata, or distinct regions, using geographical variables such as environmental and socio-economic conditions (Wang *et al*. 2016). The stratification process can group similar spatial data into homogeneous strata and distinguish them from different regions to better understand variable associations through spatial patterns. Spatial stratified heterogeneity provides effective models for examining spatial stratified association without mathematical assumptions about the geographic variables, ensuring the reliability and applicability of these models in practice (Song *et al*. 2020). The geographic detector (GD), a widely used spatial stratified heterogeneity-based model, measures stratified association using statistical variance and quantifies the association between explanatory and response variables through the power of determinant (PD) value (Wang *et al*. 2016). A series of enhanced models have been developed based on the GD framework over the past decade to improve the performance and effectiveness in analyzing stratified heterogeneity. For example, the optimal parameter-based geographic detector (OPGD) optimizes the algorithms and parameters for the spatial discretization of explanatory variables (Song *et al*. 2020), the interactive detector of spatial association (IDSA) enhances the spatial interactions of variables through fuzzy overlay approaches for more effectively quantifying spatial interaction association (Song and Wu 2021), and the geographically optimal zones-based heterogeneity (GOZH) model identifies the multiple variables-determined optimal regions for improving the adaptiveness of GD models in dealing with multiple explanatory variables (Luo *et al*. 2022). GD and its enhanced models have been widely applied across various fields, including studies on urbanization, ecological and environmental changes, climate change, and cryospheric variations (Feng *et al*. 2021, Li *et al*. 2021, Chen *et al*. 2022, Dasgupta *et al*. 2022, Zhang *et al*. 2024b).

Local analysis of geographic features is a critical component of geospatial analysis (Huang *et al*. 2023). When significant spatial association is present, global values may not be applicable across the entire study area (Fotheringham and Brunsdon 1999). Local analysis methods address this issue by defining sub-regions to compute multiple local values (Boots and Okabe 2007), playing a key role in identifying hotspots and their extents (Longley and Batty 1997), and demonstrating how relationships among geographic variables change spatially (Fotheringham *et al*. 2009). Since the initial proposal of LISA, local analysis has garnered considerable interest, with studies focusing on various topics such as categorical data (Boots 2006), points on networks (Yamada and Thill 2007), optimal spatial weights (Getis and Aldstadt 2004, Aldstadt and Getis 2006), and space-time and income mobility (Rey 2016).

However, existing research on spatial stratified heterogeneity-based models only reveals the overall level of spatial association but ignores differences in association among different regions. GD and its enhanced models primarily calculate global PD values by assessing the explanatory power of explanatory variables on the response variable. Nevertheless, the explanatory power of variables may significantly vary across different geographical locations, especially in large-scale studies. Therefore, research on spatial association based on local scales is critically important. Analysis at the local scale enables a finer-grained examination of the spatial association at each location and allows for assessing the explanatory power of explanatory variables on response variables that vary spatially.

This study develops a local indicator of stratified power (LISP) to analyze the spatial stratified association at local regions and demonstrate how locally stratified association varies across space. LISP was employed in identifying locally stratified determinants of glacier thickness in the Greater Himalayas. The remainder of this article is structured as follows: Section 2 outlines the steps in developing the LISP, Section 3 presents a case study utilizing LISP, Section 4 showcases the results of the study, Section 5 discusses the strengths and weaknesses of LISP, and finally, Section 6 provides the conclusion.

2. Local indicator of stratified power

This study proposes a local indicator of stratified power (LISP) to analyze the local variation of spatial stratified association. Figure 1 shows the process of LISP, comprising three sequential steps. First, the optimal extent for local is identified using a spatial variogram to satisfy the criterion of a sufficiently small range and ensure that the data within the extent exhibit adequate heterogeneity and association. The second step is to analyze the PD of individual variables on the response variable through GD modeling with optimal spatial discretization algorithms and parameters at the local scale. The final step is to quantify the local PD of interaction variables, including the interaction of a pair of spatial variables and the interaction of multiple variables, using a tree-based spatial discretization approach and stratified heterogeneity approaches.

2.1. Determining the extent of local variability

Determining the local extent within the LISP model can elucidate the relationship between the size of the extent and the degree of dispersion of the response variable within that extent. Many approaches can be used to identify the extent of local variability, such as replay's K function (Dixon 2001), spatial change detection approach (Zhang *et al*. 2024a), and local coefficient of variation (Rao *et al*. 2022). Semivariogram functions are used to identify the extent of local variability in this study. The assumption of spatial second-order stationarity is required when applying the LISP model, as

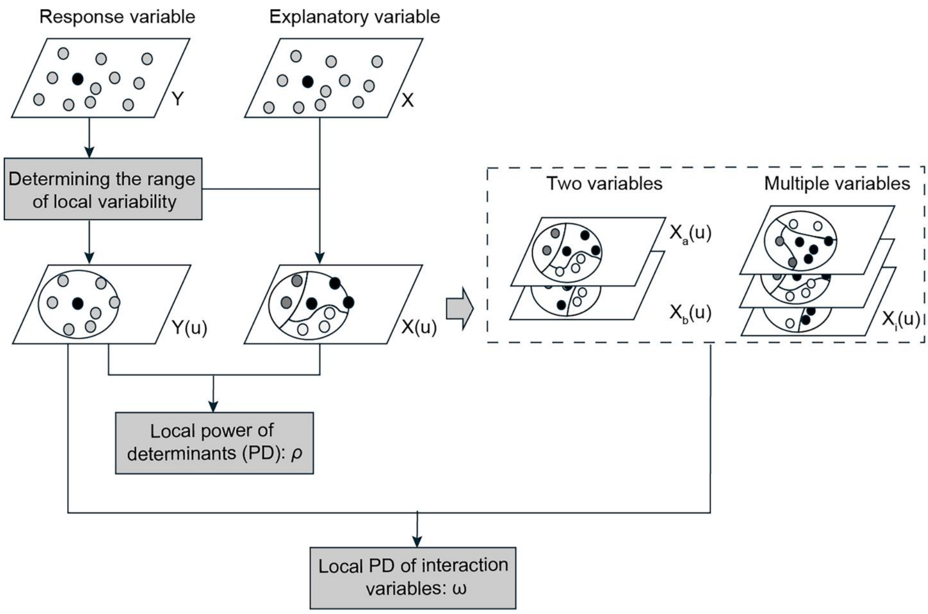

Figure 1. The technical flowchart of the local indicator of stratified power (LISP) for examining the local power of determinants.

it ensures that a reasonable local extent can be derived using semivariogram functions. Thus, this assumption should be tested to verify its validity before proceeding with the analysis. This process is divided into three stages, estimating parameters of empirical semivariograms, fitting semivariogram functions, and determining the local extent.

The semivariogram is a robust geostatistical tool used to analyze the spatial correlation of a variable to both distance and direction. Given two locations separated by a distance of h units, and the difference in the variable values at these two locations, the semivariogram represents one-half of the variance of the differences. The true semivariogram is inherently unknowable (Olea 2006), necessitating the use of its estimated values instead. In this study, we employ Equation (1) to obtain an unbiased estimate of the semivariogram at a distance of h:

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [z(y_i + h) - z(y_i)]^2, \tag{1}$$

where $z(y_i)$ represents the value of a response variable at location y_i, and $n(h)$ denotes the number of locations separated by a distance of h units.

Empirical semivariograms are spatially discrete, making it challenging to accurately and reliably determine the optimal distance. Therefore, it is necessary to fit the calculated empirical semivariogram data using theoretical models. Common theoretical models include the exponential, linear, and power (Cressie 2015). Among them, the most commonly used is the exponential model:

$$\gamma(h) = c_0 + c\left(1 - \exp\left(\frac{-h}{a}\right)\right) \tag{2}$$

where h represents distance. c_0 indicates the semivariance value as the h approaches zero, and c indicates the semivariance value as the distance approaches infinity. a represents the distance of the semivariance first reaches the value of c. Due to the inherent noise in empirical semivariograms, subjective judgment is necessary to determine which theoretical model is better suited for fitting the empirical semivariograms.

After fitting semivariogram functions, the next step is to determine the local extent. We chose twice of a to adequately cover local spatial variability that may contain the locally stratified heterogeneity within the local extent. The a represents the maximum distance where significant spatial correlation exists (Mazzella and Mazzella 2013), and using twice of a ensures a more comprehensive capture of local spatial variability. The extent of local (κ) is calculated by:

$$\kappa = 2 \times a \tag{3}$$

Thus, the local extent κ can effectively presents the local spatial variability or locally stratified heterogeneity and will be used to examine the local power of determinants in subsequent steps.

2.2. Local power of individual determinants

The local PD values of individual variables are calculated using an optimal parameters-based geographical detector (OPGD) (Song *et al.* 2020) at each local extent κ. The factor detector is the core component of OPGD and uses a Q-statistic to determine the relative significance of explanatory variables. The local PD at the uth location ($\rho(u)$) can be calculated as follows (Wang *et al.* 2010):

$$\rho(u) = 1 - \frac{\sum_{z=1}^{M} N_z \sigma_z^2}{N\sigma^2} \tag{4}$$

where N and σ^2 represent the number and standard deviation of observations within the local extent κ of location u. N_z is the number of observations and σ_z^2 refers standard deviation of observations within geographical stratum z ($z = 1, \ldots, M$) of the local extent.

The significance of the $\rho(u)$ is tested using an F-test, which is calculated as follows:

$$F = \frac{N-M}{M-1}\frac{\rho(u)}{1-\rho(u)} \sim F(M-1, N-M; \zeta) \tag{5}$$

where M and N represent the number of regions and observations within the local extent, and ζ is calculated by:

$$\zeta = \left[\sum_{j=1}^{m} \bar{Y}_j^2 - \frac{1}{N}\left(\sum_{j=1}^{M} \bar{Y}_j\sqrt{N_j}\right)^2\right] / \sigma^2 \tag{6}$$

where $\bar{Y}_j$ is the mean of the jth stratum within the local extent.

2.3. Local power of interaction determinants

The local PD of interaction variables, such as the interactions of two variables or multiple variables, is calculated by the GOZH (Luo *et al.* 2022) at each local extent κ.

The local PD of multiple variables within the local extent κ at the location u ($\omega(u)$) can be calculated as follows:

$$\omega(u) = 1 - \frac{\min(SSW_{X,D})}{SST} \tag{7}$$

where SST and SSW are the $N\sigma^2$ and $\sum_{z=1}^{M} N_z\sigma_z^2$ in Equation (4). X are explanatory variables, and D refers to stratified variables. The D is computed using a tree-based spatial discretization approach (Luo *et al.* 2022).

In addition, the contribution of an explanatory variable to an interaction of multiple or all variables can be calculated in two steps. First, a reculculated PD value is obtained by excluding the specific explanatory variable. Then, the contribution of this variable to the interaction is determined by subtracting this recalculated value from the PD value of the variable interaction (Luo *et al.* 2022).

3. Case study: identifying local stratified determinants of Mountain glacier thickness

3.1. Case study background and study area

Mountain glaciers are crucial elements of Asia's water towers, playing an essential role in supporting downstream needs for domestic, agricultural, and industrial purposes (Jouberton *et al.* 2022). Over several decades, High Mountain Asia has experienced varying degrees of glacier mass loss (Rounce *et al.* 2023). Significantly, lake-terminating glaciers have experienced a much higher mass loss compared the land-terminating glaciers (Brun *et al.* 2019, Liu *et al.* 2020). These glaciers are influenced by several factors including temperature, precipitation, elevation, glacial lakes, and the glacier surface albedo (King *et al.* 2019, Zhang *et al.* 2021, Bolibar *et al.* 2022, Rounce *et al.* 2023, Huber *et al.* 2024). LISP is employed to identify local stratified determinants of lake-terminating glacier thickness.

The study area, the Greater Himalayas, encompasses the Hengduan Mountains, the Himalayas, Nyainqentanglha Mountains, Karakoram, and the Hindu Kush (Figure 2). This region harbors the largest concentration of glaciers outside Antarctica and Greenland, temporarily storing significant freshwater resources (Zhao *et al.* 2022). The Greater Himalayas serve as the origin for 11 major rivers in Asia, supporting over 1.3 billion people who depend on these rivers for freshwater (Xu *et al.* 2009). There are more than 5,000 glacial lakes, including 908 proglacial lakes (glacier lakes connected to glaciers) (Zhang *et al.* 2023). These proglacial lakes are rapidly expanding, accelerating the melting of lake-terminating glaciers (Wang *et al.* 2020a).

3.2. Datasets

The thickness change of glaciers from 2000 to 2020 is derived from the global glacier thickness change product (Hugonnet *et al.* 2021), the spatial resolution is 100 meters. The glacier lake inventory data (Zhang *et al.* 2023) and the glacier inventory data (Pfeffer *et al.* 2014) are used to delineate the extent of lake-terminating glaciers.

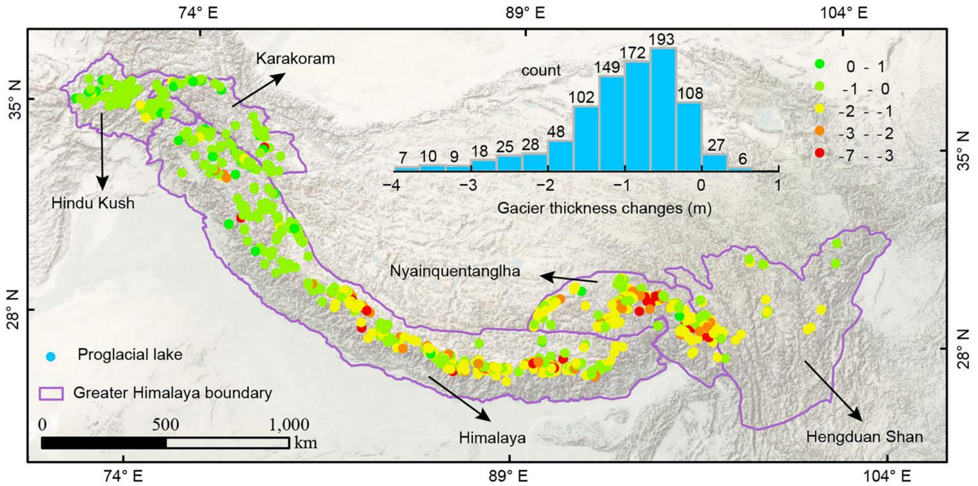

Figure 2. Spatial distribution of the thickness variations of lake-terminating glaciers in the Greater Himalayas. The dots in the chart represent the thickness variations of lake-terminating glaciers. The histogram in the figure shows the frequency of thickness changes at the terminus of glaciers connected to glacier lakes.

Table 1. Explanatory variables for evaluating the spatial distribution of thickness variations in lake-terminating glaciers.

Category	Variable	Data	Period	Spatial resolution
Climate	Summer temperature	ERA5-Land	2000–2020	11,132 m
	Winter temperature	ERA5-Land	2000–2020	11,132 m
	Precipitation	ERA5-Land	2000–2020	11,132 m
Geography	Elevation	NASADEM	2000	30 m
	Slope	NASADEM	2000	30 m
	Aspect	NASADEM	2000	30 m
Environment	Lake area	Glacial lake inventory	2020	30 m
	Surface albedo	MODIS	2000–2020	500 m

A series of potential explanatory variables are collected to explain the change in glacier thickness, categorized into climate, geography, and environment. Table 1 details these variables comprehensively. The climate is characterized by summer temperature, winter temperature, and annual cumulative precipitation, accessed from the ERA5-Land dataset (doi:10.24381/cds.68d2bb30). On the Google Earth Engine (GEE) (Gorelick *et al*. 2017), we conduct linear regression analyses (Montgomery *et al*. 2021) on summer temperature, winter temperature, and annual cumulative precipitation to ascertain their trends from 2000 to 2020. The geography is characterized by elevation, slope, and aspect. Elevation data are accessed from the NASADEM dataset (Crippen *et al*. 2016), and slope and aspect are calculated using this dataset on GEE. The environment is characterized using lake area (by proglacial lakes) and surface albedo. The lake area data are derived from the glacier lake inventory, while surface albedo data are sourced from the MODIS MOD10A1 product (doi: 10.5067/MODIS/MOD10A1.006). Linear regression analysis is conducted to determine the summer surface albedo change trend from 2000 to 2020.

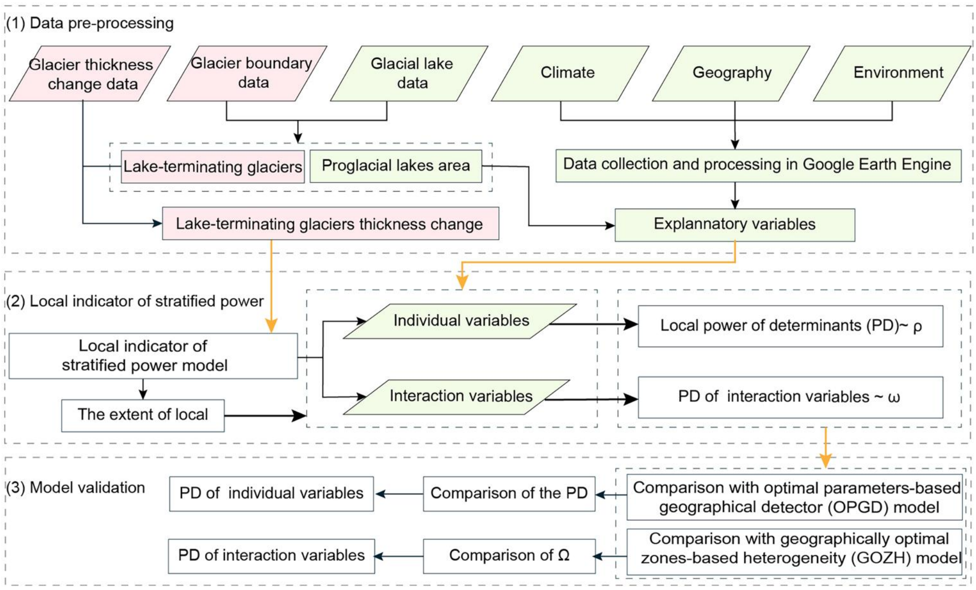

Figure 3. Diagrammatic representation of using LISP to identify key determinants affecting changes in the thickness of lake-terminating glaciers.

3.3. Experiment design

Figure 3 depicts the process flowchart of the local determinants analysis of thickness changes in lake-terminating glaciers in the Greater Himalayas using the LISP model, comprising three main steps. The first step involves data preprocessing of glacier thickness change data and explanatory variable data. The second step entails conducting a local determinants analysis of glacier thickness change using the LISP model. Finally, the model is validated to evaluate its accuracy and performance.

3.3.1. Data pre-processing

Data pre-processing involves three steps. The first step is to extract lake-terminating glaciers and determine lake area based on glacier and glacier lake boundaries. Considering the impact of lakes on glaciers, we extract the area where glaciers overlap with a 1 km buffer around the lakes as the data acquisition extent (DCE) for the response variable and explanatory variables. Next, glacier thickness change data from 2000 to 2020 are extracted as the response variable. Finally, other explanatory variables within the DCE are extracted on the GEE, forming the eight explanatory variables for this case study.

3.3.2. Identifying local stratified determinants of glacier thickness change

This section entails four steps. First, calculating the local extent is the foundation for the LISP. The second step is to analyze the local PD of each variable. The third step is to examine the local PD of variable pairs' interactions through the LISP model. Finally, the assessment of the local PD of all explanatory variables on glacier thickness variation.

3.3.3. Model evaluation

The developed LISP model is evaluated by comparing its derived local effects of PD and the global PD derived from OPGD and GOZH models that calculate overall PD values to assess variable explanatory power across space. The LISP model provides local PD values at each location for both individual variables and interactions of variables. Statistical indicators such as maximum, minimum, and average local PD values are compared to global PD values. The average local PD value reflects the overall explanatory power of a variable, while the maximum and minimum values capture its spatial variation. The effectiveness of the LISP model is examined from the perspective of both individual variables and interactions of variables, as OPGD and GOZH are typically applied to examine the overall explanatory power of these two types of variables. The global PD values in LISP, OPGD, and GOZH models are calculated using the R package "GD" (Song *et al*. 2020).

4. Results

4.1. Local extent exploration

Figure 4 illustrates the semivariogram function between the semivariance of the glacier thickness change and variation distance. The results suggested that the exponential model provides a good fit for the semivariogram function. The exponential semivariogram function-derived range value is 309.15 km, and the local extent for depicting the local variability or locally stratified heterogeneity is 618.30 km, which is used to conduct the local stratified association for the individual variables and interactions of variables in the next few stages.

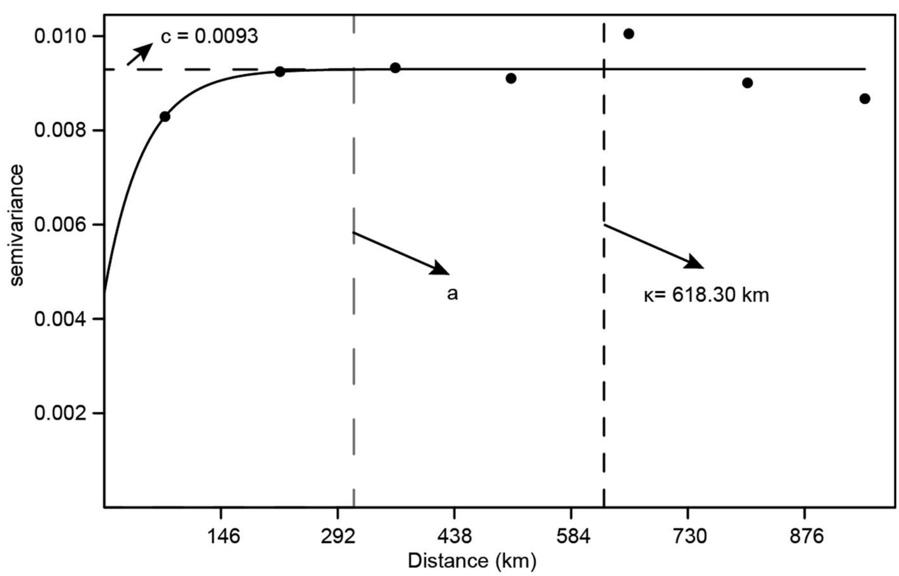

Figure 4. The process of deriving the local extent for the LISP model using the semivariogram function between the semivariance of glacier thickness change and distance.

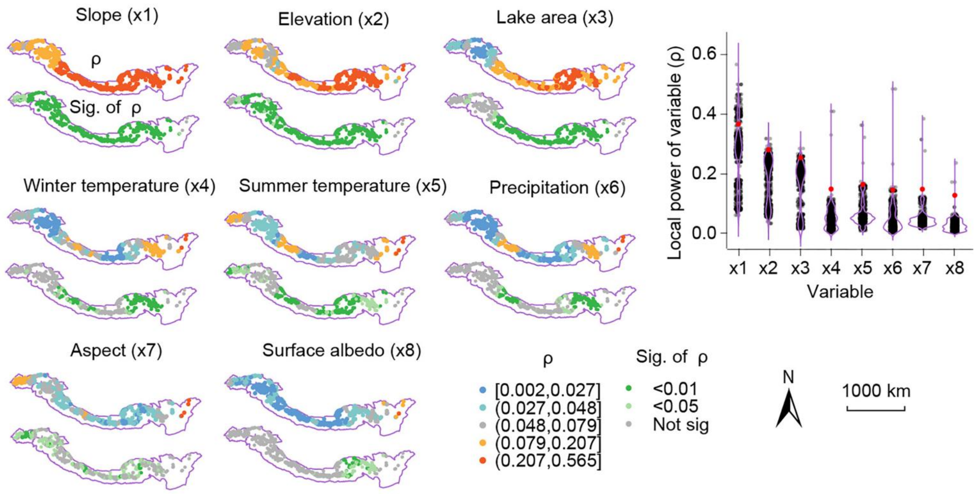

Figure 5. The local power of determinants of individual variables and their significance tests. The statistical plot in the upper right corner is a violin plot for the local power values of each variable, with red dots representing the means.

4.2. Local power of determinants of individual variables

In LISP, ρ values are used to quantify the PD of each variable on the response variable at each specific location. Figure 5 shows the ρ values of individual variables affecting glacier thickness across space. Results show that the LISP model can effectively identify the spatial variation of ρ values for each explanatory variable across the study area with the significance levels. Among these variables, slope exhibits the strongest explanatory power for glacier thickness changes, with higher ρ values in central and eastern regions, and lower ρ values in western regions. Elevation and lake area changes also show high explanatory power for glacier thickness, with similar ρ values in the central and eastern but critical differences in the western of the study area. Surface albedo exhibits the weakest explanatory power for glacier thickness changes, which can be attributed to the lake-terminating glaciers are often covered by debris, resulting in minimal variation in surface albedo changes.

The significance test results for the ρ values of the explanatory variables are presented in Figure 5. Among these variables, slope has the highest proportion of areas with significant locally stratified associations, with over 99.34% of the region showing significance levels below 0.05, except for some regions in the Hengduan Mountains, primarily attributed to the sparse data available. In addition, the major regions for elevation and lake area also show significance levels below 0.05, with only the westernmost and easternmost parts of the study area being non-significant. Notably, the trends in winter temperature variations and annual cumulative precipitation exhibit high significance primarily in the Nyainqentanglha mountains, corresponding to higher PD values. This phenomenon is attributed to the South Asian monsoon climate in the region (Jouberton *et al*. 2022).

4.3. Local power of interaction determinants

The local PD of interaction variables on the response variable is quantified using the ω value. Figure 6(a) displays the spatial distribution of ω values for three sets of interaction variables (slope ∩ lake area, slope ∩ summer temperature, and slope ∩ precipitation) and interaction by all variables. Overall, the ω values computed for the three interaction variable combinations exhibit significant spatial association. Specifically, the slope ∩ lake area shows stronger explanatory power in the central and eastern regions. Conversely, slope ∩ summer temperature and slope ∩ precipitation demonstrate stronger explanatory power in the western part of the central area, with weaker explanatory power in the eastern part. All three combinations exhibit lower PD values in the western regions. The ω values for all variable interactions indicate higher ω values in the central, while lower ω values are observed in the western and eastern, with the maximum ω value reaching 0.707. In this case, since the ω value represents the cumulative effect of two and all variables on the response variable, the significance is high across all locations.

The contribution of each explanatory variable to ω values is shown in Figure 6(b). The contribution of the slope is paramount, albeit exhibiting negative values in the western region (indicating an increase in ω upon excluding the slope). This indicates

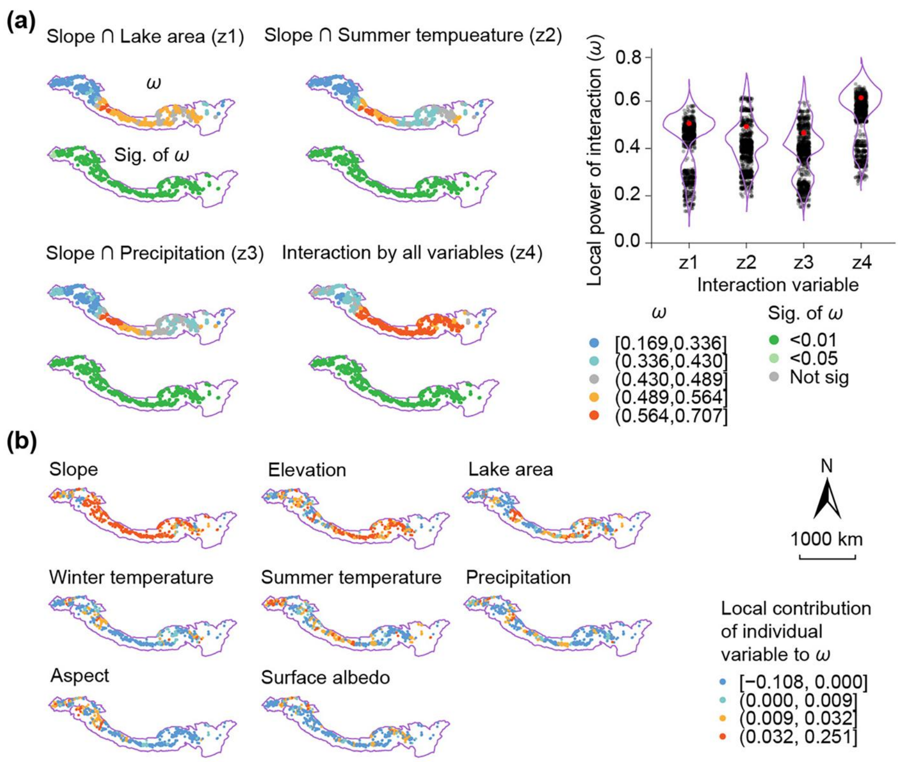

Figure 6. The local power of interaction determinants. (a) The interaction effects with two and all factors and their significance tests. (b) The contribution of each variable to the interaction of multiple variables at a local scale. The statistical plot in the upper right corner is a violin plot for the local power values of each variable.

that, for certain localized regions, having more explanatory variables is not necessarily better. Some variables can influence the contribution of other variables to the ω values. This finding serves as a pivotal criterion for selecting explanatory variables across distinct regions.

4.5. Model evaluation

Individual variables and interaction variables are used to assess the effectiveness of the LISP in analyzing local factors affecting glacier thickness changes on a large scale. This evaluation involved comparing the local PD values derived from LISP models with the global PD values obtained from two widely used enhanced GD models (i.e. OPGD and GOZH). Since OPGD and GOZH models provide PD values on a global scale, we derived the mean, minimum, and maximum local PD values from the LISP model for comparison. Table 2 presents the statistical results based on global and local PD values.

The LISP model demonstrates outstanding performance in exploring individual variables. Specifically, the mean local PD values for most variables exhibit high consistency compared to the global PD values, suggesting that the global GD model captures the average effect of explanatory power across the entire study area. However, there is a slight difference in PD values for elevation calculated using the OPGD and LISP models. Mean local PD value of LISP is 0.181, while the global PD value of OPGD is 0.062 ($p < 0.01$). Results show that the LISP model can compute PD values for all locations, indicating essential spatial variability of PD values, which can't be examined through the OPGD, a global stratified heterogeneity model. As such, the LISP model effectively avoids underestimation or overestimation of PD values for individual variables that can occur with global geographic detector models.

In assessing interaction variables, the LISP and GOZH models were used to calculate the local and global PD values for three combinations: slope ∩ lake area, slope ∩

Table 2. Comparison of the LISP-derived local PD values and global PD values in the existing stratified heterogeneity models.

	Variable	Local PD of LISP mean [min, max]	Percentage of locations with significant LISP ($p < 0.05$)	PD of global geographical detector
Individual variables	Slope	0.266 [0.062, 0.565]	99.34%	0.207**
	Elevation	0.181 [0.033, 0.318]	95.70%	0.062**
	Lake area	0.154 [0.014, 0.285]	75.88%	0.155**
	Winter temperature	0.064 [0.009, 0.363]	73.12%	0.008*
	Summer temperature	0.049 [0.002, 0.409]	46.15%	0.056**
	Precipitation	0.046 [0.002, 0.484]	41.62%	0.066**
	Aspect	0.048 [0.016, 0.384]	51.87%	0.021**
	Surface albedo	0.028 [0.002, 0.236]	25.55%	0.053**
Interaction variables	Slope ∩ Lake area	0.441 [0.169, 0.615]	98.67%	0.361**
	Slope ∩ Summer temperature	0.427 [0.230, 0.652]	100%	0.289**
	Slope ∩ Precipitation	0.401 [0.189, 0.632]	100%	0.372**
	Interaction of all variables	0.549 [0.284, 0.707]	100%	0.547**

Significance levels: **$p < 0.01$; *$p < 0.05$.

precipitation, and slope ∩ summer temperature. The results indicate that average local PD values for all three combinations are higher than their corresponding global PD values. Notably, although the average local and global PD values for the slope ∩ lake area combination are higher than those for the other two combinations, the maximum local PD values for the slope ∩ precipitation and slope ∩ summer temperature combinations are significantly higher than those for the slope ∩ lake area combination. This suggests that the glacier thickness changes in certain areas are primarily controlled by temperature and precipitation, consistent with attribution studies (Hugonnet *et al.* 2021, Huber *et al.* 2024) of most glacier thickness changes (including non-lake-terminating glaciers). This further underscores the reliability of the LISP model in large-scale studies.

In exploring all variables, local PD values and global PD values were calculated using the LISP and GOZH models, respectively. The results indicate that the local PD value with a mean value of 0.549 and ranging from 0.284 to 0.707 and the global PD value (0.547) are generally consistent in terms of the mean value, suggesting that the overall explanatory power of multiple variables derived from the LISP-based local PD aligns well with the existing stratified heterogeneity-based global PD values. However, the local PD can additionally capture spatial variability, demonstrating the influence of variables varies across different regions and enhancing the accuracy and explanatory power of the model.

5. Discussion

With advancements in geographic information science, local analytical methods have gained increasing importance in spatial analysis. Relying solely on global values or sets of global values as model parameters often proves inadequate for many study areas (Boots and Okabe 2007). When geographic data exhibit critical spatial disparity, global metrics become less applicable, failing to account for variations across different regions within the study area and may not even be suitable for any subregions (Fotheringham and Brunsdon 1999). For instance, the local indicators of spatial association (LISA) model, by identifying spatial associations in localized regions (Anselin 1995), has driven the widespread adoption of local analytical techniques. These methods have enhanced spatial models across various domains, including spatial autocorrelation, local heterogeneity, heteroscedasticity, and stratified heterogeneity (Table 3). Based on spatial dependence, the Moran statistic (MORAN 1950) and the Geary c statistic (Geary 1954) are essential metrics for quantifying global spatial autocorrelation. The advancement of local analytical methodologies has significantly enhanced the performance of spatial autocorrelation models and indicators, including the development of LISA (Anselin 1995), the local Geary c statistic (Anselin 2019), and heterogeneous spatial autocorrelation (Zhang *et al.* 2024b). In spatial heterogeneity, geographically weighted regression (GWR), along with its refined models (Brunsdon *et al.* 1998, Fotheringham *et al.* 2017), mitigates the influence of local effects on global linear or nonlinear regression models. To measure spatial heteroscedastic effects, local spatial heteroscedasticity has been employed to test the null hypothesis that local variance is indistinguishable from the variance of the entire study area (Ord and Getis 2012), and

Table 3. Illustrative examples explaining the importance of employing local models in spatial analysis.

Category	Global models and indicators	Local models and indicators
Spatial autocorrelation	Moran statistic (MORAN 1950); Geary c statistic (Geary 1954)	Local Indicator of Spatial Association (Anselin 1995); Local Geary c statistic (Zhang *et al*. 2024b); Heterogeneous spatial autocorrelation (Zhang *et al*. 2024b)
Spatial heterogeneity	Head-tail index (Jiang and Yin 2014)	Geographically weighted regression (Brunsdon *et al*. 1998); Multiscale geographically weighted regression (Fotheringham *et al*. 2017)
Spatial heteroscedasticity	Spatial error model with heteroscedasticity (Toloza *et al*. 2024)	Local spatial heteroscedasticity (Ord and Getis 2012); Spatiotemporal autoregressive conditional heteroscedasticity (Otto 2024)
Spatial stratified heterogeneity	Geographical detectors (Wang *et al*. 2010), optimal parameters-based geographical detectors (Song and Wu 2021)	Local indicator of stratified power (LISP) (This study)

the autoregressive conditional heteroscedasticity model, extended to spatial local contexts, addresses spatial variance dynamics and spill-over effects (Otto 2024).

Despite these advances, existing local models do not address how spatial stratified heterogeneity evolves across different regions. This study addresses this gap by proposing a local indicator of stratified power (LISP) to analyze local spatial stratified association. LISP provides a framework to quantify how spatial stratified association varies across local regions and how these local effects contribute to the overall spatial association. By comparing local and global dynamics of stratified spatial heterogeneity, LISP indicates that spatial patterns, processes, and their determinants vary across space. This study contributes to the existing literature by demonstrating that spatial stratification is not only a global phenomenon but also significantly shaped by localized patterns, which can now be effectively analyzed using the LISP model. This advancement provides new knowledge and tools for identifying regional spatial processes, which are critical for applications such as environmental monitoring, urban planning, and socio-economic analysis.

The LISP model provides several advantages in exploring spatial stratified association at a local scale. First, the LISP can explore the spatial association between response variables and explanatory variables within local regions, thus analyzing local driving mechanisms rather than merely providing a global average relationship. Second, the LISP model can prevent the overestimation or underestimation of PD values that can occur in traditional models studying spatial stratified association. Finally, the LISP model exhibits robust data adaptability, allowing it to optimally determine the representative local extent, whether dealing with gridded or irregularly distributed data. The advantages of the LISP model effectively enhance applicability. Heterogeneity depends on size and varies in both large and small scales (Luo *et al*. 2023, Zhang *et al*. 2024b). The geographical extent of this study spans 1395,632 km^2, which allows for the identification of regional and localized heterogeneity. However, patterns may vary at smaller or larger scales.

However, this research has several limitations and future studies are recommended to address the following aspects. First, anisotropy is a critical aspect of semivariograms, and incorporating anisotropy analysis may add significant value to enhancing the model. Second, this study defines the local extent for examining local variability using samples within a specific distance around the target location. However, in sparsely sampled areas, the samples involved in the PD calculation are limited, reducing the reliability of the results. Utilizing a fixed number of nearest observations to the target location for PD calculation might address this issue, but it may lead to insufficient association of the response variable in densely sampled areas. Finally, the second-order stationarity assumed in the semivariogram, which is not always appropriate for spatial data, should be examined for various cases and practical applications. Thus, future research should adopt more advanced and precise methods to define the local range to enhance the performance of the LISP model.

6. Conclusion

This study proposes a local indicator of stratified power (LISP) model to address critical gaps in analyzing spatial stratified heterogeneity by capturing local variations in determinant effects commonly ignored by global models, thereby advancing geographical information science through enhanced local spatial analysis and a broader understanding of regional variability in spatial associations. By defining optimal extents of local heterogeneity and quantifying local power for individual and interaction variables, LISP effectively mitigates the underestimation or overestimation of the power of determinants, providing more accurate spatial associations in various local regions. This study provides readers with a powerful tool to analyze spatial variability and local determinants in large-scale studies, enhancing the analysis of how spatial associations vary across regions and supporting more accurate and region-specific decision-making. The model has substantial potential for broad applications across various fields in exploring local spatial determinants, assessing spatial variations, and advancing spatial decision-making in complex geospatial issues. Future research is recommended to expand the model by incorporating anisotropy analysis, adaptive extent definitions, flexible variability assumptions, spatiotemporal analyses, and testing its adaptability across diverse geospatial contexts.

Acknowledgment

We extend our sincere gratitude to the editor and the anonymous reviewers for their insightful comments and suggestions, which have significantly enhanced the quality of this manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the China Scholarship Council [Grant Number: 202308510292], and the Second Tibetan Plateau Scientific Expedition and Research Program [Grant Number: 2019QZK0307].

Notes on contributors

Jiao Hu is a PhD candidate at the College of Earth and Planetary Sciences, Chengdu University of Technology, Chengdu, China. He is also a Research Associate at Curtin University, Perth, Australia. His research interests encompass spatial statistics, deep learning, remote sensing, and glaciers. Emails: hujiao@stu.cdut.edu.cn; jiao.hu@curtin.edu.au

Yongze Song is a Senior Lecturer at Curtin University, Australia. His research interests include spatial statistics, geospatial intelligence, sustainable infrastructure, and sustainable development. He serves as an Associate Editor for the journals GIScience & Remote Sensing, and the International Journal of Applied Earth Observation and Geoinformation. Email: Yongze.Song@curtin.edu.au

Tingbin Zhang is a professor at the College of Earth and Planetary Sciences at Chengdu University of Technology and the State Environmental Protection Key Laboratory of Synergetic Control and Joint Remediation for Soil and Water Pollution (Chengdu University of Technology), Chengdu, China. His research interests include environment, ecology, hydrology, and remote sensing. Email: zhangtb@cdut.edu.cn

ORCID

Jiao Hu http://orcid.org/0000-0002-9644-9763
Yongze Song http://orcid.org/0000-0003-3420-9622
Tingbin Zhang http://orcid.org/0000-0002-2480-9774

Data availability statement

Data and codes supporting the findings of this study are available at https://doi.org/10.6084/m9.figshare.26518651.v3

References

Aldstadt, J., and Getis, A., 2006. Using AMOEBA to create a spatial weights matrix and identify spatial clusters. *Geographical Analysis*, 38 (4), 327–343.

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

Anselin, L., 2019. A local indicator of multivariate spatial association: extending Geary's c. *Geographical Analysis*, 51 (2), 133–150.

Ben-Moshe, S., and Itzkovitz, S., 2019. Spatial heterogeneity in the mammalian liver. *Nature Reviews. Gastroenterology & Hepatology*, 16 (7), 395–410.

Bolibar, J., *et al.*, 2022. Nonlinear sensitivity of glacier mass balance to future climate change unveiled by deep learning. *Nature Communications*, 13 (1), 409.

Boots, B., 2006. Local configuration measures for categorical spatial data: binary regular lattices. *Journal of Geographical Systems*, 8 (1), 1–24.

Boots, B., and Okabe, A., 2007. Local statistical spatial analysis: inventory and prospect. *International Journal of Geographical Information Science*, 21 (4), 355–375.

Brun, F., *et al.*, 2019. Heterogeneous influence of glacier morphology on the mass balance variability in High Mountain Asia. *Journal of Geophysical Research: Earth Surface*, 124 (6), 1331–1345.

Brunsdon, C., Fotheringham, S., and Charlton, M., 1998. Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47 (3), 431–443.

Chen, J., *et al.*, 2022. Magnitudes and patterns of large-scale permafrost ground deformation revealed by Sentinel-1 InSAR on the central Qinghai-Tibet Plateau. *Remote Sensing of Environment*, 268, 112778.

Chen, W., Wang, G., and Zeng, J., 2023. Impact of urbanization on ecosystem health in Chinese urban agglomerations. *Environmental Impact Assessment Review*, 98, 106964.

Cressie, N., 2015. *Statistics for spatial data*. Hoboken, NJ: John Wiley & Sons.

Crippen, R., *et al.*, 2016. NASADEM global elevation model: methods and progress. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B4, 125–128.

Dasgupta, S., *et al.*, 2022. Siting priorities for congestion-reducing projects in Dhaka: a spatio-temporal analysis of traffic congestion, travel times, air pollution, and exposure vulnerability. *International Journal of Sustainable Transportation*, 16 (12), 1078–1096.

Dixon, P.M., 2001. Ripley's K function. *In*: A.H. El-Shaarawi and W.W. Piegorsch, eds. *Encyclopedia of environmetrics*, Volume 3. Chichester: Wiley, 1796–1803.

Feng, R., *et al.*, 2021. Urban ecological land and natural-anthropogenic environment interactively drive surface urban heat island: an urban agglomeration-level study in China. *Environment International*, 157, 106857.

Fotheringham, A.S., and Brunsdon, C., 1999. Local forms of spatial analysis. *Geographical Analysis*, 31 (4), 340–358.

Fotheringham, A.S., Brunsdon, C., and Charlton, M.E., 2009. Geographically weighted regression. *The Sage Handbook of Spatial Analysis*, 1, 243–254.

Fotheringham, A.S., Yang, W., and Kang, W., 2017. Multiscale geographically weighted regression (MGWR). *Annals of the American Association of Geographers*, 107 (6), 1247–1265.

Ge, Y., *et al.*, 2017. Geographically weighted regression-based determinants of malaria incidences in northern China. *Transactions in GIS*, 21 (5), 934–953.

Geary, R.C., 1954. The contiguity ratio and statistical mapping. *The Incorporated Statistician*, 5 (3), 115–146.

Getis, A., and Aldstadt, J., 2004. Constructing the spatial weights matrix using a local statistic. *Geographical Analysis*, 36 (2), 90–104.

Gorelick, N., *et al.*, 2017. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27.

Harvey, E.P., and O'Neale, D.R.J., 2024. The importance of spatial heterogeneity in disease transmission. *Nature Physics*, 20 (7), 1053–1054.

Huang, B., Wu, B., and Barry, M., 2010. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science*, 24 (3), 383–401.

Huang, F., *et al.*, 2023. Mapping local climate zones for cities: a large review. *Remote Sensing of Environment*, 292, 113573.

Huber, C.J., *et al.*, 2024. High-altitude glacier archives lost due to climate change-related melting. *Nature Geoscience*, 17 (2), 110–113.

Hugonnet, R., *et al.*, 2021. Accelerated global glacier mass loss in the early twenty-first century. *Nature*, 592 (7856), 726–731.

Jiang, B., 2015. Geospatial analysis requires a different way of thinking: the problem of spatial heterogeneity. *GeoJournal*, 80 (1), 1–13.

Jiang, B., and Yin, J., 2014. Ht-index for quantifying the fractal or scaling structure of geographic features. *Annals of the Association of American Geographers*, 104 (3), 530–540.

Jouberton, A., *et al.*, 2022. Warming-induced monsoon precipitation phase change intensifies glacier mass loss in the southeastern Tibetan Plateau. *Proceedings of the National Academy of Sciences of the United States of America*, 119 (37), e2109796119.

King, O., *et al.*, 2019. Glacial lakes exacerbate Himalayan glacier mass loss. *Scientific Reports*, 9 (1), 18145.

Li, H., Jia, P., and Fei, T., 2021. Associations between taste preferences and chronic diseases: a population-based exploratory study in China. *Public Health Nutrition*, 24 (8), 2021–2032.

Liu, Q., *et al.*, 2020. Interannual flow dynamics driven by frontal retreat of a lake-terminating glacier in the Chinese Central Himalaya. *Earth and Planetary Science Letters*, 546, 116450.
Longley, P.A., and Batty, M., 1997. *Spatial analysis: modelling in a GIS environment*. Chichester, England: John Wiley & Sons.
Luo, P., *et al.*, 2022. Identifying determinants of spatio-temporal disparities in soil moisture of the Northern Hemisphere using a geographically optimal zones-based heterogeneity model. *ISPRS Journal of Photogrammetry and Remote Sensing*, 185, 111–128.
Luo, P., *et al.*, 2023. A generalized heterogeneity model for spatial interpolation. *International Journal of Geographical Information Science*, 37 (3), 634–659.
Majumder, S., *et al.*, 2023. Multiscale GIS based-model to assess urban social vulnerability and associated risk: evidence from 146 urban centers of Eastern India. *Sustainable Cities and Society*, 96, 104692.
Mazzella, A., and Mazzella, A., 2013. The importance of the model choice for experimental semivariogram modeling and its consequence in evaluation process. *Journal of Engineering*, 2013, 1–10.
Montgomery, D.C., Peck, E.A., and Vining, G.G., 2021. *Introduction to linear regression analysis*. Hoboken, NJ: John Wiley & Sons.
MORAN, P.A.P., 1950. Notes on continuous stochastic phenomena. *Biometrika*, 37 (1-2), 17–23.
Olea, R.A., 2006. A six-step practical approach to semivariogram modeling. *Stochastic Environmental Research and Risk Assessment*, 20 (5), 307–318.
Ord, J.K., and Getis, A., 1995. Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, 27 (4), 286–306.
Ord, J.K., and Getis, A., 2012. Local spatial heteroscedasticity (LOSH). *The Annals of Regional Science*, 48 (2), 529–539.
Otto, P., 2024. A multivariate spatial and spatiotemporal ARCH model. *Spatial Statistics*, 60, 100823.
Páez, A., Uchida, T., and Miyamoto, K., 2002. A general framework for estimation and inference of geographically weighted regression models: 2. Spatial association and model specification tests. *Environment and Planning A: Economy and Space*, 34 (5), 883–904.
Pfeffer, W.T., *et al.*, 2014. The Randolph glacier inventory: a globally complete inventory of glaciers. *Journal of Glaciology*, 60 (221), 537–552.
Qian, H., *et al.*, 2024. Effects of climate and environmental heterogeneity on the phylogenetic structure of regional angiosperm floras worldwide. *Nature Communications*, 15 (1), 1079.
Rao, J., *et al.*, 2022. Infrared small target detection based on weighted local coefficient of variation measure. *Sensors*, 22 (9), 3462.
Rey, S.J., 2016. Space–time patterns of rank concordance: local indicators of mobility association with application to spatial income inequality dynamics. *Annals of the American Association of Geographers*, 106 (4), 788–803.
Rounce, D.R., *et al.*, 2023. Global glacier change in the 21st century: every increase in temperature matters. *Science (New York, N.Y.)*, 379 (6627), 78–83.
Seoane, P.A., *et al.*, 2023. Astrophysics with the laser interferometer space antenna. arXiv.org.
Song, Y., 2022. The second dimension of spatial association. *International Journal of Applied Earth Observation and Geoinformation*, 111, 102834.
Song, Y., *et al.*, 2020. An optimal parameters-based geographical detector model enhances geographic characteristics of explanatory variables for spatial heterogeneity analysis: cases with different types of spatial data. *GIScience & Remote Sensing*, 57 (5), 593–610.
Song, Y., and Wu, P., 2021. An interactive detector for spatial associations. *International Journal of Geographical Information Science*, 35 (8), 1676–1701.
Toloza, J.D., Melo, O.O., and Cruz, N.A., 2024. Spatial error models with heteroskedastic normal perturbations and joint modeling of mean and variance.
Wang, J., *et al.*, 2010. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. *International Journal of Geographical Information Science*, 24 (1), 107–127.

Wang, X., *et al.*, 2020a. Glacial lake inventory of high-mountain Asia in 1990 and 2018 derived from Landsat images. *Earth System Science Data*, 12 (3), 2169–2182.

Wang, J., Gao, B., and Stein, A., 2020b. The spatial statistic trinity: a generic framework for spatial sampling and inference. *Environmental Modelling & Software*, 134, 104835.

Wang, J.-F., Zhang, T.-L., and Fu, B.-J., 2016. A measure of spatial stratified heterogeneity. *Ecological Indicators*, 67, 250–256.

Xu, J., *et al.*, 2009. The melting Himalayas: cascading effects of climate change on water, biodiversity, and livelihoods. *Conservation Biology: The Journal of the Society for Conservation Biology*, 23 (3), 520–530.

Yamada, I., and Thill, J.-C., 2007. Local indicators of network-constrained clusters in spatial point patterns. *Geographical Analysis*, 39 (3), 268–292.

Zhang, G., *et al.*, 2023. Underestimated mass loss from lake-terminating glaciers in the greater Himalaya. *Nature Geoscience*, 16 (4), 333–338.

Zhang, Y., *et al.*, 2021. Albedo reduction as an important driver for glacier melting in Tibetan Plateau and its surrounding areas. *Earth-Science Reviews*, 220, 103735.

Zhang, Z., *et al.*, 2024a. Robust interaction detector: a case of road life expectancy analysis. *Spatial Statistics*, 59, 100814.

Zhang, Z., Li, Z., and Song, Y., 2024b. On ignoring the heterogeneity in spatial autocorrelation: consequences and solutions. *International Journal of Geographical Information Science*, 38 (12), 2545–2571.

Zhao, F., *et al.*, 2022. Rapid glacier mass loss in the Southeastern Tibetan Plateau since the year 2000 from satellite observations. *Remote Sensing of Environment*, 270, 112853.