geographical analysis

Geographical Analysis (2022) **0**, 1–14

Applying Local Indicators of Spatial Association to Analyze Longitudinal Data: The Absolute Perspective

Ran Tao, Yuzhou Chen

School of Geosciences, University of South Florida, Tampa, FL USA

Local Indicators of Spatial Association (LISA) are a class of spatial statistical methods that have been widely applied in various scientific fields. When applying LISA to make longitudinal comparisons of spatial data, a common way is to run LISA analysis at each time point, then compare the results to infer the distributional dynamics of spatial processes. Given that LISA hinges on the global mean value that often varies across time, the LISA result generated at time T_i reflects the spatial patterns strictly with respect to T_i. Therefore, the typical comparative cross-sectional analysis with LISA can only characterize the relative distributional dynamics. However, the relative perspective alone is inadequate to comprehend the full picture, as the patterns are not directly associated with the changes of the spatial process's intensity. We argue that it is important to obtain the absolute distribution dynamics to complement the relative perspective, especially for tracking how spatial processes evolve across time at the local level. We develop a solution that modifies the significance test when implementing LISA analysis of longitudinal data to reveal and visualize the absolute distribution dynamics. Experiments were conducted with Mongolian livestock data and Rwanda population data.

Introduction

Local Indicators of Spatial Association (LISA) are a class of spatial statistical methods that have been widely applied in many fields such as geography, regional science, demography, transportation, urban planning, and epidemiology. Proposed by Anselin in 1995, LISA serves two purposes: to be interpreted as indicators of local pockets of first-order nonstationarity (or hot spots), and to assess the influence of individual locations on the magnitude of the global statistic and to identify "outliers" (Anselin 1995). The original LISA family included several statistics, including the local Moran's *I*, the local Gamma, and the local Geary's *C*, with a focus on examining local spatial autocorrelation. A closely related work around the same time was the local G statistic, which focuses on measuring first-order nonstationarity, or hot spot detection (Getis and Ord 2010; Ord and Getis 1995). Since then, the LISA family has expanded with many new statistics that either examine a different spatial property or suit a specific application domain.

Correspondence: Ran Tao, School of Geosciences, University of South Florida, Tampa, FL, USA
e-mail: rtao@usf.edu

Submitted: 6 August 2021; Revised version accepted: 18 February 2022

doi: 10.1111/gean.12323

Noticeable examples include the local spatial heteroscedasticity (LOSH; Ord and Getis 2012); the bivariate spatial association measures (Anselin, Syabri, and Smirnov 2002; Lee 2001); the multivariate local Geary's *C* (Anselin 2019); LISA in a network space (Yamada and Thill 2010); LISA for origin-destination flow data (Berglund and Karlström 1999; Tao and Thill 2020), etc. The LISA statistics are designed to examine spatial patterns only at the time when the dataset was collected. In other words, they only handle cross-sectional spatial data defined as a collection of observations (behavior) for multiple spatial units at a single point in time. On the other hand, longitudinal spatial data are widely available and of great study value. The best example is the spatial panel data, also known as spatial panels, such as census data that records a long list of socioeconomic variables over time on the same set of geographic units. There is a strong need for longitudinal comparisons of the distributional dynamics, namely how the spatial patterns of the same study area change over time.

Thanks to the increasingly available toolkits that can run LISA analysis and output attractive result maps (Bivand and Piras 2015), someone with no rigorous training in spatial statistics can carry out the analysis through the graphical user interface (GUI) of software such as GeoDa and ArcGIS Pro. The result maps are self-explanatory as they depict the distribution of clusters of strong or weak spatial association in the study area that are statistically significant. For a typical spatial panel data, a series of result maps can be generated to illustrate the spatial patterns of the same study area in each year. A common way to infer the distributional dynamics of spatial processes is to make direct comparisons of these result maps on at least three aspects. First, to compare the total number of local patterns, e.g., to see which year has more high-value clusters. Second, to compare the spatial distribution of local patterns to see whether the detected clusters shift their locations over the years. Third, to compare the patterns of the same local region for tracking its changes.

Such comparative cross-sectional analysis with LISA has been found in many published works. For example, Saizen, Maekawa, and Yamamura (2010) performed LISA analysis on a Mongolian livestock population dataset and compared the results of four different years; Kowe, Mutanga, and Odindi (2020) applied LISA statistics on a multi-temporal Zimbabwe Landsat data to study long-term vegetation fragmentation; Lutz (2019) evaluated the European digital single market strategy by running LISA analysis on the longitudinal Eurostat data from 2011 to 2016; Seabrook (2009) compared the LISA results of 2004 and 2008 U.S. presidential elections to study the "Obama Effects"; Wang et al. (2014) made comparison of space-time patterns of Beijing hand-foot-mouth disease from 2008 to 2012 based on LISA maps. Other examples from published articles are too many to list, of which the study topics include, but are not limited to air pollution (Chen, Xu, and Li 2018), regional economic growth (Kulkarni, Haynes, and Stough 2011), tourism flows (Xing-zhu and Qun 2014), tuberculosis (Ge, Zhang, and Wang 2016), and poverty (Katumba, Cheruiyot, and Mushongera 2019).

However, direct comparison of LISA results across time can only characterize the relative distributional dynamics. Arribas-Bel and Tranos (Arribas-Bel and Tranos 2018) illustrate this with the so-called "pseudo space-time LISA", which is to apply LISA to generate cross section results at each hour in a day. The results are comparable with the understanding that they are relative to the hourly mean. It is worth mentioning that there are scholars both inside and outside geospatial science disciplines not fully aware of this limitation, and therefore draw questionable conclusions about the spatiotemporal processes of their interest. This fact hints at a strong need to formally discuss the differences between relative and absolute dynamics. We argue that both perspectives are equally important, and the latter is particularly

useful for tracking how spatial processes evolve across time at the local level. This study aims at providing an easy-to-implement solution to modify LISA analyses to reveal and visualize the absolute distributional dynamics, as an important complement to the existing relative perspective.

The idea behind our solution is basically re-evaluating distribution dynamics from Rey (2001). Recognizing that different null hypotheses can generate very different LISA results (Sauer, Oshan, and Rey 2021), our basic logic is to modify the benchmark of significance test for comparable results. Our solution should not be confused with the research line of developing local indicators of space-time autocorrelation, which integrates the spatial autocorrelation and temporal autocorrelation with a spatiotemporal weight matrix to identify spatiotemporal clusters. For example, the space-time local Moran's *I* (Shen, Li, and Si 2016; Lee and Li 2017), and the space-time local G statistic (Wang and Lam 2020). In contrast, we attempt to isolate comparable spatial patterns by removing the influence of temporal nonstationarity.

Relative and absolute

The original LISA statistics were designed without considering the temporal dimension. For instance, the arguably most widely adopted LISA, namely the local Moran's *I* (Anselin 1995), is shown as equation (1),

$$I_i = \frac{n\left(x_i - \overline{x}\right)\sum_j^n w_{ij}\left(x_j - \overline{x}\right)}{\sum_i^n \left(x_i - \overline{x}\right)^2}, \tag{1}$$

where I_i is the local Moran's *I* statistic of spatial unit i; n is the total number of units; x_i is the attribute value of region i; $\overline{x}$ is the average of x_i; w_{ij} is the spatial weight between unit i and unit j . There is no time-related element of the formula, as the statistic was designed to measure spatial autocorrelation at the time of data collection by default. In other words, LISA results are implicitly anchored at the time of data collection, i.e., the result generated at time T_i, only reflects the spatial patterns with respect to T_i. Comparing the LISA results from two different time points can only infer changes in relative patterns.

To explain why such comparative cross-sectional analysis with LISA can only lead to relative distributional dynamics and how it might be inadequate, we take the conceptual example shown in Fig. 1. In all three subfigures, the curved surface represents the spatial process of interest, e.g., population distribution in a region. The vertical axis denotes the numeric value or the intensity level of the spatial process. There are two local clusters with extreme values, one is the peak-like high-value cluster and the other is the valley-like low-value cluster. The semitransparent surface above the mean-value surface indicates the significance upper bound derived from Monte Carlo simulations, and any part of the spatial process above the upper bound is deemed a statistically significant cluster of high values with strong spatial autocorrelation (commonly referred as high-high (HH) cluster). For example, the red part in Fig. 1a is a HH cluster as it surpasses the pink significance upper-bound surface. Similarly, the semitransparent surface below the mean-value surface indicates the significance lower bound, and any part of the spatial process below it is deemed statistically significant cluster of low values with strong spatial autocorrelation (commonly referred as low-low (LL) cluster). For example, the turquoise part in Fig. 1a is a LL cluster as it is below the yellow significance lower-bound surface.

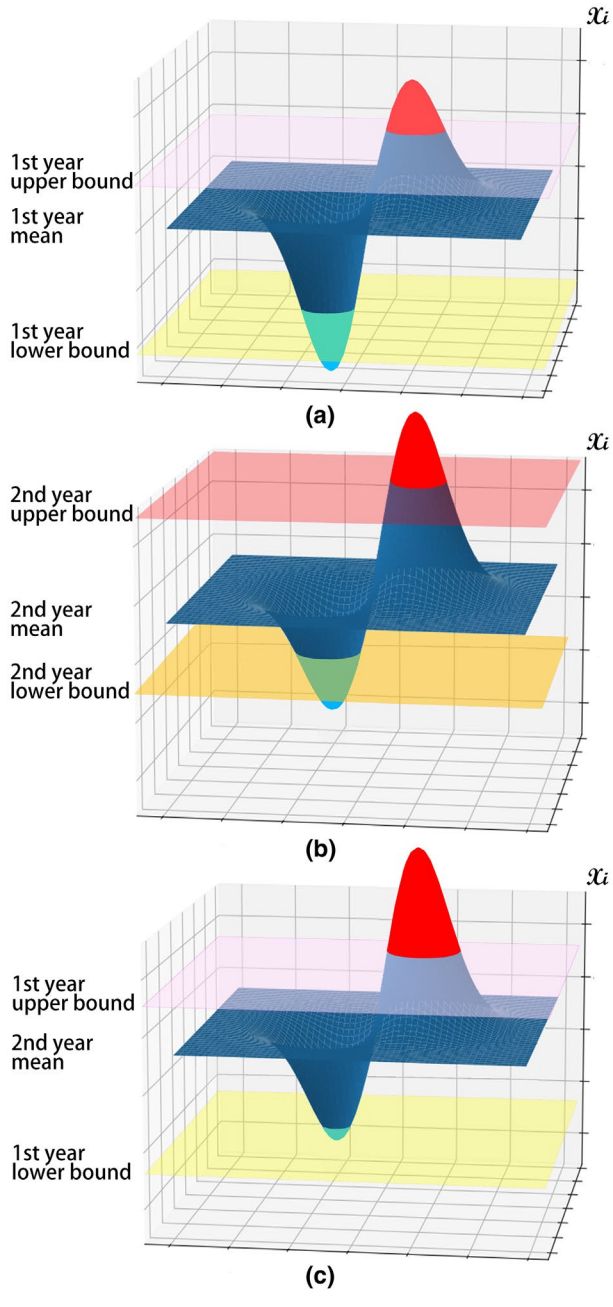

Figure 1. Comparison of relative and absolute distributional dynamics. (a) Year 1 spatial process with year 1 significance envelopes; (b) year 2 spatial process with year 2 significance envelopes; (c) year 2 spatial process with year 1 significance envelopes.

Fig. 1a shows the situation in year 1 and Fig. 1b shows year 2. From year 1 to year 2, there is a global increase of value in the study area: the peak becomes higher; the valley becomes shallower; the mean-value surface is lifted. Due to the global value boost, the significance envelopes in the

second year become higher, too. As a result, the two clusters in Fig. 1b keep approximately the same sizes as year 1, despite the increase of their absolute values. The comparison between Fig. 1a and b is a typical comparative cross-sectional analysis with LISA. The conclusions drawn from such comparison are from the relative perspective. In this case, the relative distributional dynamics are considered stable between the two years. First, the number of each type of local clusters remains the same as there is only one HH cluster and one LL cluster in either year. Second, the spatial distributions of clusters are also similar as their locations remain unmoved in the study area. Third, the size of each cluster, namely the part above the upper bound or below the lower bound, is relatively stable as well.

However, in most real-world scenarios, the dynamics of spatial processes are much more complicated than this example. And LISAs serve as an important set of exploratory spatial data analysis (ESDA) toolkits to help understand complex spatial processes (Anselin 1998). Therefore, a natural logic is to run LISA analyses first, observe and compare their patterns to make inferences about the spatial processes. While knowing the relative distributional dynamics is rather useful, it can be inadequate to comprehend the full picture of how spatial processes evolve over time. The relative distributional pattern is technically irrelevant to the absolute value changes. The spatial process itself can be drastically intensified or weakened to maintain stable relative distributional dynamics, on the condition that the changes happen consistently and synchronously in space. In other words, the changes between two time points do not favor or disfavor particular local areas to create relative advantages or disadvantages. For example, the case shown in Fig. 1a and b maintains a stable relative pattern despite a significant increase in absolute value. On the other hand, little change of absolute value at a local area can still result in a highly dynamic relative pattern, if the global mean value has changed. Therefore, we argue that a more comprehensive manner to grasp distributional dynamics of spatial processes is to combine both relative and absolute perspectives. Since the default way of comparing LISA results infers the relative dynamics, it is the purpose of this paper to bring up a solution to reveal the absolute dynamics.

Our solution has a simple logic: using the same benchmark for significance tests of different years. More specifically, we use the same significance envelopes to decide the LISA results of different years. Fig. 1c illustrates the basic idea of our strategy: we adopt the significance envelopes of year 1 to determine LISA results of year 2. In this way, we observe that the HH cluster becomes larger as it includes more neighboring spatial units, while the LL cluster has shrunk as only the very bottom portion is deemed significant. The observation is consistent with the absolute value changes between these two years. The solution is particularly useful to track the dynamics of spatial process at the local level. For example, the intensity of spatial process becomes even more extreme at the HH cluster's location but moderated at the LL cluster's location. Such patterns are hidden by the stable relative pattern derived from the typical comparative cross-sectional analysis with LISA.

Implementation

To reveal the absolute distributional dynamics from time point T_p to time point T_q, we use the same benchmark, e.g., the one at time T_p, to perform significant tests and to derive comparable LISA results. The general steps of implementation are summarized as follows, taking the local Moran's I as the representative LISA statistic.

a) Calculate the local Moran's I of T_p using equation (2). Equation (2) is the same as the original formula in equation (1). The only difference is the newly added footnote T_p that specifies the time of calculation. x_{i,T_p} is the attribute value of spatial unit i at time T_p; $\overline{x}_{T_p}$ is the average value at time T_p; the spatial weight matrix w_{ij} is assumed not varying with time.

$$I_{i,T_p} = \frac{n\left(x_{i,T_p} - \overline{x}_{T_p}\right)\sum_j^n w_{ij}\left(x_{j,T_p} - \overline{x}_{T_p}\right)}{\sum_i^n \left(x_{i,T_p} - \overline{x}_{T_p}\right)^2} \quad (2)$$

b) To evaluate the statistical significance of I_{i,T_p}, Monte Carlo simulation is a common method. Each time, randomly distribute all x_{i,T_p} values in the study area and calculate the local Moran's I as step a). Repeat the simulation for N times (e.g., $N = 999$), and sort all the calculated I_{i,T_p} results for each unit. The thresholds of the top and bottom 5% (if setting the significance level at 0.05) are the upper-bound and lower-bound envelopes of significance, respectively. Compare the original I_{i,T_p} values with the two envelopes to determine whether the local patterns are significant.

c) Calculate the local Moran's I of $T_q$using equation (3). $I_{i,T_q/T_p}$is the local Moran's I value of spatial unit $i$for comparing $T_q$with T_p. $x_{i,T_q}$is the attribute value of spatial unit $i$at time T_q. Unlike the original formula, the average value $\overline{x}_{T_p}$is taken from time $T_p$instead of T_q, so does the standard deviation in the denominator,

$$I_{i,T_q/T_p} = \frac{n\left(x_{i,T_q} - \overline{x}_{T_p}\right)\sum_j^n w_{ij}\left(x_{j,T_q} - \overline{x}_{T_p}\right)}{\sum_i^n \left(x_{i,T_p} - \overline{x}_{T_p}\right)^2} \quad (3)$$

d) Run Monte Carlo simulations to test the significance of $I_{i,T_q/T_p}$. Because our strategy is to use the benchmark of T_p, the simulation is done by randomizing all x_{i,T_p} values in the study area. The two envelopes will stay the same as the ones in step b). Comparing the $I_{i,T_q/T_p}$ values with the envelopes, we can determine the significance of local patterns.

e) The local patterns at step b) and step d) are now directly comparable to infer absolute distributional dynamics of the spatial process from T_p to T_q.

It is worth mentioning that while we primarily use the local Moran's I as an example, the solution is applicable to other LISA or LISA-like statistics such as the local Geary's C, the local G statistic, and LOSH statistic, with slight modifications according to their formulas.

Experiments

Case 1: Mongolia stock

We did our first experiment with the livestock population data at the sub-provincial level in Mongolia. This dataset is publicly available from the web portal of the Mongolia National Statistical Office. More importantly, this is the same dataset used in Saizen, Maekawa, and Yamamura (2010), in which we found issues in their LISA result interpretations. We focus on analyzing two types of livestock, namely goat and sheep, as both experienced dramatic changes in terms of population and spatial distribution. The maps in the first row of Fig. 2 illustrate the goat density in four different years selected for discussion in Saizen, Maekawa, and Yamamura (2010). The total population of goats was 5.6 million in 1992 and almost doubled to 11 million by 1999. It then slightly decreased to 9.1 million by 2002 but gained significantly to 15.4 million by 2006. Goats in Mongolia traditionally concentrated in both the western and central regions and sprawled to other regions as time went by.

The maps in the second row of Fig. 2 show the original local Moran's *I* result corresponding to each year, of which Saizen, Maekawa, and Yamamura (2010) made following observations:

> "Comparing the LISA cluster maps for the period 1992–2006, the HH cluster was observed in western Mongolia in 1992, and its distribution had shifted to the east in 1999. In 2002, the HH cluster had disappeared and was once again distributed in western Mongolia in 2006… Instead, it was shown that the HH goat cluster was shifting, and its distribution was different from those of other animals' HH clusters."

Interestingly, the authors found peculiarity of their observations by saying that "these changes do not seem to be associated with the change in population density." Nevertheless, the authors could not properly explain this finding because they were not fully aware of the limitations of direct comparisons of LISA results from different years. In fact, the second row of Fig. 2 depicts only moderate changes of LISA results from 1992 to 2006, which indicate a generally stable pattern of Mongolian goat's distributional dynamics from the relative perspective. Considering that the goat population almost tripled during that period, it is safe to conclude that the increase was rather even across space. For example, most regions in the central and western side of country remain nonsignificant in all four maps in row 2 of Fig. 2, as they received similar degrees of goat population increase throughout years. However, the relative perspective alone is insufficient, especially when trying to explain local patterns that contradict with the underlying trend of population changes. For example, the LL cluster located at the southeastern corner of Mongolia became larger, despite noticeable increase of goat population in that region; several HH clusters detected in 1992 and 1999 turned into nonsignificant LISA patterns in 2002 and 2006, in contrast with the overall growing trend of goat population.

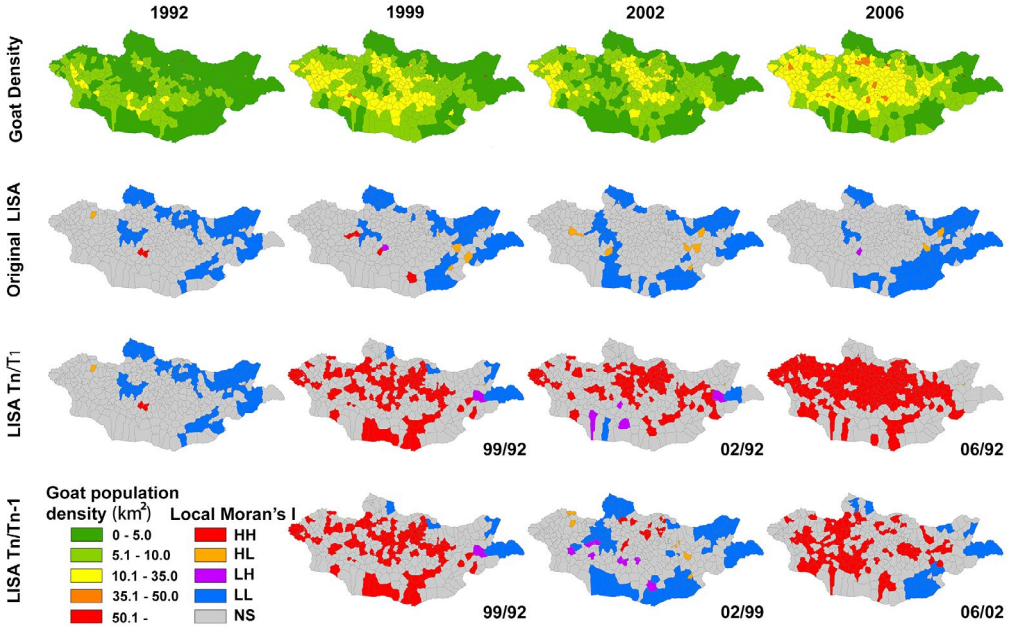

Figure 2. Mongolian goat density and the local Moran's *I* results.

We applied our solution to reveal the absolute distributional dynamics that go along with goat population trend at both global and local scales. We generated two batches of LISA results with different parameter settings. The maps in the third row of Fig. 2 are the results using the first year of observation (T_1 = 1992) as the benchmark. Therefore, all four maps in the third row are directly comparable. For example, in contrast with 1992, there are many more HH clusters in 2006, while the LL clusters have disappeared. In 2002, the number of HH clusters is higher than in 1992, but lower than in 1999. The maps in the fourth row of Fig. 2 are the results using the previous year of observation ($T_{n\,1}$) as the benchmark. In this way, the absolute dynamics can be revealed in a more sensitive way. For instance, the "02/99" map shows many LL clusters but only a few HH clusters. It accurately picks up the setback of goat farming industry from the decade-long booming trend. The "06/02" map depicts a strong comeback and proves that the previous setback was temporary.

Combing both relative and absolute perspectives can lead us to a more comprehensive view. Taking the southeastern corner as an example, the "06/92" map shows the region as nonsignificant LISA pattern from absolute perspective, indicating the increase of goat density was not significant enough to form a HH cluster there. Meanwhile, there are many HH clusters detected in the rest of Mongolia. The degree of goat density increase in the southeastern corner does not match the average level in the country, which results in a relative disadvantage for that region. In consequence, we observe an expanded LL cluster in that region in row 2 of Fig. 2, showing a negative trend of relative dynamics. The disappearance of several HH clusters in row 2 of Fig. 2 is another example of a contradiction between relative and absolute dynamics. The cause was not any decrease in goat density in those local areas. Instead, many regions have received a drastic increase in goat density, which made high-density areas rather evenly distributed in the central and west parts of the country. Therefore, the previous HH clusters lost their relative advantage as the only high-density regions. Of course, there are examples that the relative and absolute dynamics that correspond to each other. The region near the north border exhibits a positive trend from both relative and absolute perspectives, as it turned from an LL cluster to nonsignificant pattern in the second row and evolved as a new HH cluster in the third row.

Sheep is the other livestock selected for analysis. The sheep population in Mongolia was 14.6 million in 1992, 15.2 million in 1999, 10.6 million in 2002, and 14.8 million in 2006. It maintained a stable size in most years except for 2002, in which there was a nearly 50% drop. The original LISA results in the second row of Fig. 3 are similar to each other, indicating a stable relative pattern in general. An interesting example is at the northeastern corner of Mongolia. The LL clusters detected in 1999 almost completely disappeared in 2002, despite drastic decrease of sheep population in that period. A jointly analysis with the absolute dynamics can help explain this. By applying our solution with the first year of observation (T_1 = 1992) as the benchmark of significance test, the "02/92" map has the least number of HH clusters and the greatest number of LL clusters among the four maps in the third row. Looking closely at the transition from "99/92" map to "02/92" map, we found the pattern in the northeastern corner remained stable, while many other regions exhibited a negative trend: most of the HH clusters disappeared and many new LL clusters emerged. In other words, the northeastern corner did not experience a huge sheep density drop from 1999 to 2002 as the rest of Mongolia, which attributes to the improvement of its relative status reflected by the disappearance of LL clusters in the second row of Fig. 3.

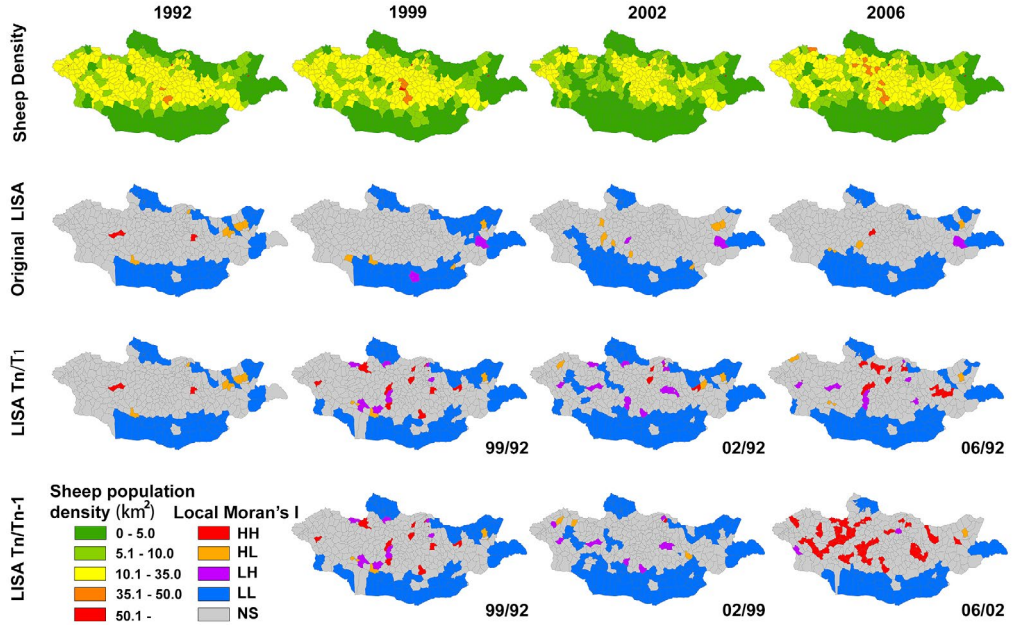

Figure 3. Mongolia sheep density and the local Moran's *I* results.

Using the previous year of observation as the benchmark, we can observe even more obvious patterns. In the fourth row of Fig. 3, the "02/99" map reflects the decrease of sheep population from 1999 to 2002: there is no HH cluster but widely distributed LL clusters. On the other hand, the "06/02" map reflects a steep increase in sheep population from 2002 to 2006, as many HH clusters have emerged while some LL clusters have disappeared. The spatial distribution of local patterns is also consistent with reality. For instance, according to the density maps in the first row, sheep were more concentrated in 2006 than in 1992 despite similar population sizes. Reflected in the "06/92" map, we observe several newly emerged HH clusters in central Mongolia.

Case 2: Rwanda population

Local Moran's I

Our second experiment was conducted with a Rwanda population dataset. Rwanda's population has been increasing quickly since the genocide in 1994. The high spatial and temporal nonstationarity makes it a perfect candidate for testing our solution. In this case study, we analyzed the 1km-resolution population raster data of 2004, 2007, 2010, and 2013, downloaded from the public source WorldPop (https://www.worldpop.org/). The maps in the first row of Fig. 4 depict the population data in the four selected years. Rwanda's population had grown from 8.7 million in 2004 to 10.8 million in 2013, with a relatively steady annual growth rate of around 2.5%. The traditional populous areas, including the capital city Kigali near the geographic center, and the northwestern part bordering the Demographic Republic of the Congo, had received a large proportion of the increased population. Meanwhile, there emerged some new clusters of residents, such as the northeastern part bordering Uganda.

Like the previous experiments, we generated three batches of local Moran's *I* result: one with the original formula, one with the modified version using the first year of observation

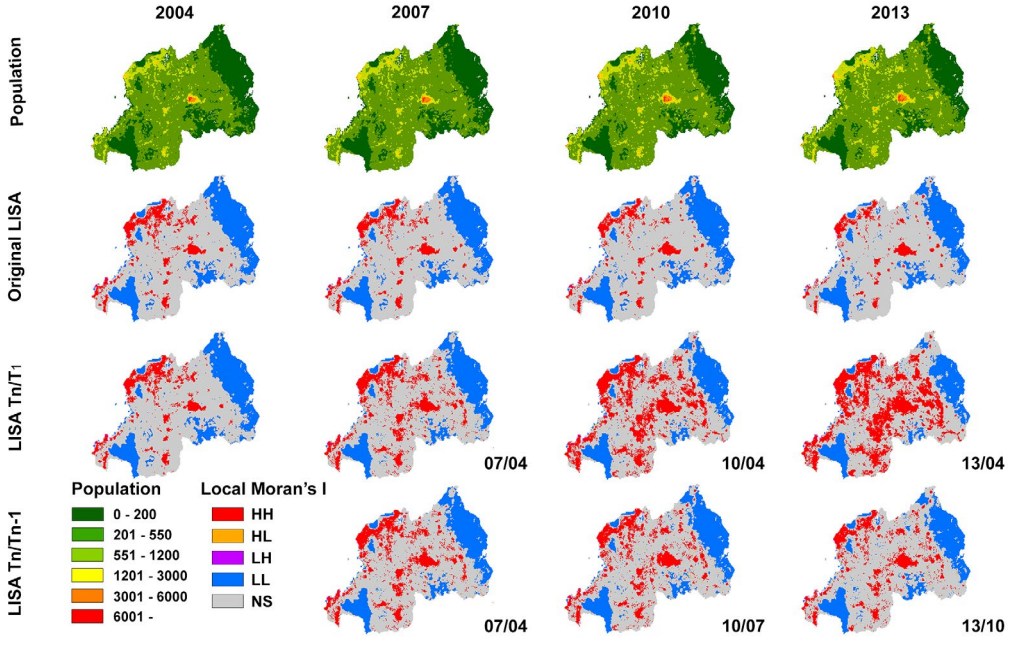

Figure 4. Rwanda population and the local Moran's *I* results.

(T_1 = 2004) as the benchmark, and one with the modified version using the previous year of observation ($T_{n\,1}$) as the benchmark. The second row of Fig. 4 lists the maps of the original LISA results that illustrate relative distributional dynamics of Rwanda population. While similar patterns have been kept for many regions, there are some noticeable differences, including the reduction of HH cluster in the northwestern region, the newly formed HH clusters in the eastern part, and the reduction of LL clusters in the northeastern corner.

In the third row of Fig. 4, the four selected years were put on the same scale for comparison as they all use year 2004 as the benchmark of significance test. From left to right, we observe an increasing trend of HH clusters and a decreasing trend of LL clusters. Both trends are consistent with the trend of population growth. Overall, the HH clusters that represent populous areas had been growing and spreading out in space. The original HH clusters in 2004 were preserved and enlarged over the years. New HH clusters emerged in several locations, especially in the southern and eastern parts of the country. Regarding the LL clusters that indicate homogenous rural or desert areas, their total size had been shrinking. In particular, the original LL clusters in the northeastern corner turned into either nonsignificant or even HH clusters because of population influx. Combining the findings of row 2 and row 3, we found that the population growth in the eastern part of Rwanda outpaced the national average. It results in newly formed HH clusters and reduction of LL clusters from both relative and absolute perspectives. On the other hand, the growth in the northwestern region was slower than the national average, which results in more HH clusters from absolute perspective but fewer HH clusters from relative perspective.

The fourth row of Fig. 4 lists the result maps using the previous year of observation as the benchmark. However, the observable difference is moderate compared with their counterparts in

the third row. The fact that the three maps in the fourth row look alike indicates a steady population growth rate in Rwanda, unlike the previous case study in which the Mongolian livestock population experienced drastic fluctuation in some years.

The local G statistic

We also tested our approach with the local G statistic. To be specific, we adopted the G_i^* version in our experiment, as shown in equation (4),

$$G_i^*(d) = \frac{\sum_{j=1}^{n} w_{ij}(d)\, x_j - \overline{X} \sum_{j=1}^{n} w_{ij}(d)}{S\sqrt{\frac{\left[n\sum_{j=1}^{n} w_{ij}(d)^2 - \left(\sum_{j=1}^{n} w_{ij}(d)\right)^2\right]}{n-1}}} \quad (4)$$

where G_i^* is the local G statistic that includes the value of x at spatial unit i, or the situation when $j = i$; n is the total number of features; x_j is the attribute value at unit j; where w_{ij} is a binary spatial weight matrix with ones for any links that are within distance d of a given i; $\overline{X}$ is the mean value of the observations; S is the standard deviation.

Despite the difference in equations, our solution applied to the local G statistic generates highly comparable results as to the local Moran's I. Fig. 5 displays three batches of G_i^* result: one with the original formula, one with the modified version using the first year of observation (T_1 = 2004) as the benchmark, and one with the modified version using the previous year of observation ($T_{n\,1}$) as the benchmark. Unlike the local Moran's I, the local G statistic generates only two types of local patterns that are statistically significant, namely the cluster of high values (or known as the "hot spots") and the cluster of low values (or known as the "cold spots"). The two types of clusters are similar as the local Moran's I HH clusters and LL clusters, respectively.

The overall patterns observed in Fig. 5 are not far from what we saw in Fig. 4. The four maps in the second row showing the original G_i^* results are barely distinguishable, despite the huge population growth over the study period. When using the same benchmark for comparing G_i^* results of different years, drastic differences can be observed among the maps in the third row: there are significantly more high-value clusters and fewer low-value clusters in later years. The maps in the last row express a mild but noticeable difference, in comparison with the corresponding original G_i^* map. It reflects the cumulative population growth in the three-year gap.

Discussion and conclusion

When applying LISA statistics to make longitudinal comparisons of spatial data, a common approach is to generate cross section result for each time point and then compare them. It has been adopted by scholars both inside and outside geospatial science disciplines. However, such comparative cross-sectional analysis can only lead to relative distributional dynamics, which may be inadequate to fully understand the underlying spatial process. Moreover, if users who take this approach are not aware of this limitation, the result interpretation can be insufficient or even incorrect. Take the Mongolian livestock study (Saizen, Maekawa, and Yamamura 2010) as an example, the authors observed that the relative patterns are not associated with livestock population changes, but they could not properly explain it.

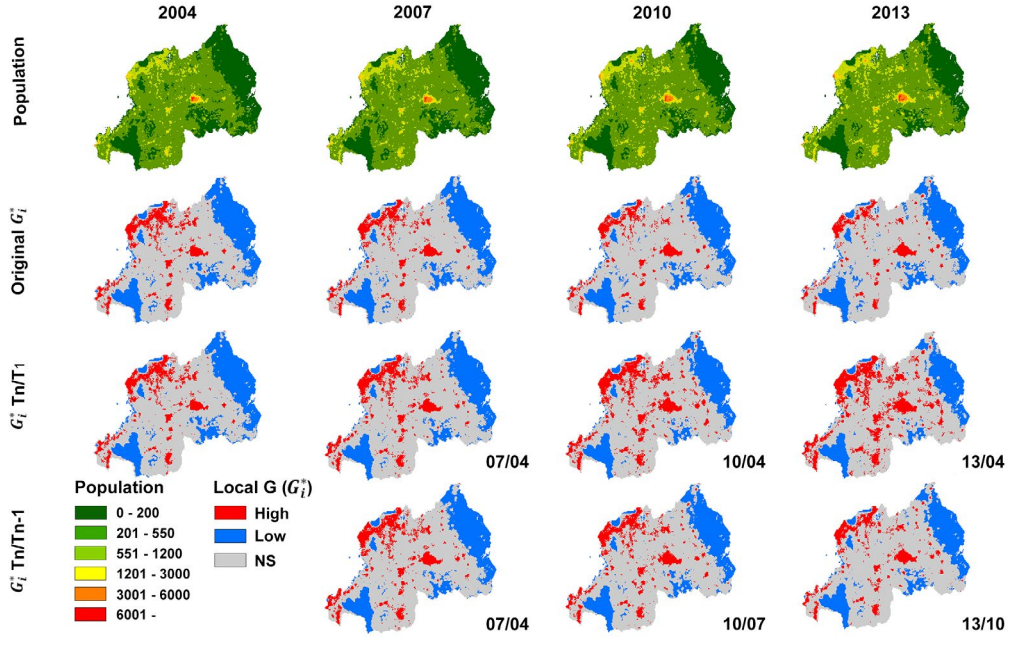

Figure 5. Rwanda population and the local G (G_i^*) results.

We argue that it is equally important to examine absolute distributional dynamics of spatial processes. In this study, we propose a solution by modifying the step of the significance test. We force LISA analysis of different times to use the same benchmark of statistical significance to derive results of the local patterns. The implementation is to use the data mean, standard deviation, and significance envelopes from the benchmark time to calculate LISA results. In our experiments, we tested two versions: using the first year of observation (T_1) as the benchmark for all years and using the previous year of observation ($T_{n\,1}$) as the benchmark for T_n. The former allows for comparing all years using the same standard, and the latter allows for comparing LISA results of two consecutive years.

We purposely selected two case studies for testing our solution, one with Mongolian livestock data and the other with Rwanda population data. By applying our solution, the modified LISA results using the first year of observation (T_1) as the benchmark are consistent with the actual data changes. For example, there are more HH clusters but fewer LL clusters when the population is steadily increasing, such as the Mongolian goat case and the Rwanda population case. Using the previous year of observation ($T_{n\,1}$) as the benchmark can reveal changes in a more sensitive way. For instance, the "02/99" map and the "06/02" map in the Mongolian sheep case reflect the drastic drop and bounce of sheep population, respectively. The maps in the last row of Fig. 4 show steady and smooth population growth in Rwanda. In either case study, we observed in some local areas the relative patterns contradict with their absolute counterparts, or vice versa. For instance, we identified negative relative dynamics of goat population in the southeastern corner of Mongolia despite a growing trend of absolute value. Neither relative nor absolute perspective alone can reveal the panorama of the distributional dynamics of spatial processes. The best practice to combine them for a joint analysis.

The experiment with Rwanda population data demonstrated that our solution can be applied to methods other than the local Moran's *I*. Despite differences in formula and result interpretations, the local G statistic can be modified in the same logic. In fact, the detected hot spots and cold spots are highly comparable with the HH clusters and LL clusters from the local Moran's *I*. The number of hot spots and cold spots and their spatial distributions become consistent with the population growth and relocation in Rwanda. Presumably, our solution can be applied to other LISA statistics and their variants, given the common adoption of Monte Carlo simulation in statistical inference. A future direction is to look beyond the LISA family, i.e., to test whether such a solution can work for a broader range of spatial analytical methods. Again, our solution should not be confused with the family of space-time LISA statistics (e.g., Shen, Li, and Si 2016; Lee and Li 2017; Wang and Lam 2020), which identify clusters of spatiotemporal autocorrelations. However, it is recommended that multiple perspectives to be considered to obtain a panoramic view of the spatiotemporal dynamics.

References

Anselin, L. (1995). "Local Indicators of Spatial Association—LISA." *Geographical Analysis* 27(2), 93–115.

Anselin, L. (1998). "Exploratory Spatial Data Analysis in a Geocomputational Environment." Geocomputation: A Primer. Chichester, NY: Wiley.

Anselin, L. (2019). "A Local Indicator of Multivariate Spatial Association: Extending Geary's C." *Geographical Analysis* 51(2), 133–50.

Anselin, L., I. Syabri, and O. Smirnov (2002). "Visualizing Multivariate Spatial Correlation with Dynamically Linked Windows." In Proceedings, CSISS Workshop on New Tools for Spatial Data Analysis. Santa Barbara, CA.

Arribas-Bel, D., and E. Tranos (2018). "Characterizing the Spatial Structure (s) of Cities "on the Fly": The Space-Time Calendar." *Geographical Analysis* 50(2), 162–81.

Berglund, S., and A. Karlström (1999). "Identifying Local Spatial Association in Flow Data." *Journal of Geographical Systems* 1(3), 219–36.

Bivand, R., and G. Piras (2015). Comparing Implementations of Estimation Methods for spatial Econometrics. American Statistical Association.

Chen, J., C. Xu, K. Li, and M. Song (2018). "A Gravity Model and Exploratory Spatial Data Analysis of Prefecture-Scale Pollutant and CO_2 Emissions in China." *Ecological Indicators* 90, 554–63.

Ge, E., X. Zhang, X. Wang, and X. Wei (2016). "Spatial and Temporal Analysis of Tuberculosis in Zhejiang Province, China, 2009–2012." *Infectious Diseases of Poverty* 5(1), 1–10.

Getis, A., and J. K. Ord (2010). "The analysis of Spatial Association by use of Distance Statistics." Perspectives on Spatial Data Analysis, 127–45. Berlin, Heidelberg: Springer.

Katumba, S., K. Cheruiyot, and D. Mushongera (2019). "Spatial Change in the Concentration of Multidimensional Poverty in Gauteng, South Africa: Evidence from Quality of Life Survey Data." *Social Indicators Research* 145(1), 95–115.

Kowe, P., O. Mutanga, J. Odindi, and T. Dube (2020). "A Quantitative Framework for Analysing Long Term Spatial Clustering and Vegetation Fragmentation in an Urban Landscape Using Multi-Temporal Landsat Data." *International Journal of Applied Earth Observation and Geoinformation* 88, 102057.

Kulkarni, R., K. Haynes, R. Stough, and J. Riggle (2011). "Light Based Growth Indicator (LBGI): Exploratory Analysis of Developing a Proxy for Local Economic Growth Based on Night Lights." *Regional Science Policy & Practice* 3(2), 101–13.

Lee, J., and S. Li (2017). "Extending Moran's Index for Measuring Spatiotemporal Clustering of Geographic Events." *Geographical Analysis* 49(1), 36–57.

Lee, S. I. (2001). "Developing a Bivariate Spatial Association Measure: An Integration of Pearson's R and Moran's I." *Journal of Geographical Systems* 3(4), 369–85.

Lutz, S. U. (2019). "The European Digital Single Market Strategy: Local Indicators of Spatial Association 2011–2016." *Telecommunications Policy* 43(5), 393–410.

Ord, J. K., and A. Getis (1995). "Local Spatial Autocorrelation Statistics: Distributional Issues and an Application." *Geographical Analysis* 27(4), 286–306.

Ord, J. K., and A. Getis (2012). "Local Spatial Heteroscedasticity (LOSH)." *The Annals of Regional Science* 48(2), 529–39.

Rey, S. J. (2001). "Spatial Empirics for Economic Growth and Convergence." *Geographical Analysis* 33(3), 195–214.

Saizen, I., A. Maekawa, and N. Yamamura (2010). "Spatial Analysis of Time-Series Changes in Livestock Distribution by Detection of Local Spatial Associations in Mongolia." *Applied Geography* 30(4), 639–49.

Sauer, J., T. Oshan, S. Rey, and L. J. Wolf (2021). "The Importance of Null Hypotheses: Understanding Differences in Local Moran's under Heteroskedasticity." *Geographical Analysis*. https://doi.org/10.1111/gean.12304.

Seabrook, N. R. (2009). "The Obama Effect: Patterns of Geographic Clustering in the 2004 and 2008 Presidential Elections." *The Forum* 7(2), 1–15. https://doi.org/10.2202/1540-8884.1308.

Shen, C., C. Li, and Y. Si (2016). "Spatio-Temporal Autocorrelation Measures for Nonstationary Series: A New Temporally Detrended Spatio-Temporal Moran's Index." *Physics Letters A* 380(1–2), 106–16.

Tao, R., and J. C. Thill (2020). "BiFlowLISA: Measuring Spatial Association for Bivariate Flow Data." *Computers, Environment and Urban Systems* 83, 101519.

Wang, J., Z. Cao, D. D. Zeng, Q. Wang, X. Wang, and H. Qian (2014). "Epidemiological Analysis, Detection, and Comparison of Space-Time Patterns of Beijing Hand-Foot-Mouth Disease (2008–2012)." *PLoS One* 9(3), e92745.

Wang, Z., and N. S. Lam (2020). "Extending Getis-Ord Statistics to Account for Local Space–Time Autocorrelation in Spatial Panel Data." *The Professional Geographer* 72(3), 411–20.

Xing-zhu, Y., and W. Qun (2014). "Exploratory Space–Time Analysis of Inbound Tourism Flows to China Cities." *International Journal of Tourism Research* 16(3), 303–12.

Yamada, I., and J. C. Thill (2010). "Local Indicators of Network-Constrained Clusters in Spatial Patterns Represented by a Link Attribute." *Annals of the Association of American Geographers* 100(2), 269–85.