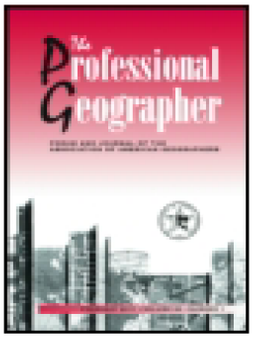

The Professional Geographer

ISSN: 0033-0124 (Print) 1467-9272 (Online) Journal homepage: https://www.tandfonline.com/loi/rtpg20

Extending Getis–Ord Statistics to Account for Local Space–Time Autocorrelation in Spatial Panel Data

Zheye Wang & Nina S. N. Lam

To cite this article: Zheye Wang & Nina S. N. Lam (2020): Extending Getis–Ord Statistics to Account for Local Space–Time Autocorrelation in Spatial Panel Data, The Professional Geographer

To link to this article: https://doi.org/10.1080/00330124.2019.1709215

Published online: 13 Feb 2020.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=rtpg20

Extending Getis–Ord Statistics to Account for Local Space–Time Autocorrelation in Spatial Panel Data

Zheye Wang and Nina S. N. Lam
Louisiana State University

Space and time are both crucial characteristic dimensions of geographic events and phenomena. Although exploratory spatial data analysis (ESDA) can be used to visualize and summarize complex spatial patterns, it has limitations in capturing the temporal dynamics of geographic features. Efforts have been made to incorporate the time dimension into ESDA techniques to detect space–time clustering or trends. Localized space–time statistics that could help in exploratory space–time data analysis (ESTDA), however, are still lacking. Focusing on spatial panel data, our work extended Getis–Ord G_i and G_i^* statistics using a space–time contemporaneous weight matrix and a space–time lagged weight matrix to account for local space–time autocorrelation. Two applications in this article show that the newly developed method can be used to summarize space–time patterns from spatial panel data, identify changes of landscape more consistently, and lend the results readily to visualization. **Key Words: exploratory space–time data analysis (ESTDA), local space–time autocorrelation, space–time G_i and G_i^*, space–time contemporaneous weight matrix, space–time lagged weight matrix.**

空间和时间均为地理事件和现象的关键特征维度。尽管可以使用空间数据探索分析（ESDA）对复杂的空间式进行可视化和概述，但它在捕获地理特征的时空动态方面存在局限性。为检测时空聚类或趋势，科学家设法在 ESDA 技术中纳入了时间维度，但仍然缺乏可用于时空数据探索分析（ESTDA）的局部时空统计数据。我们的研究侧重于空间面板数据，使用时空同期权重矩阵和时空滞后权重矩阵来扩展 Getis—Ord G_i and G_i^* 统计，以解释局部时空自相关性。本文中的两个应用说明，新开发的方法可用于总结空间面板数据中的时空模式，以便更一致地识别地貌变化，更轻松地提供可视化的结果。 **关键词：时空数据探索分析（ESTDA），局部时空自相关，时空 G_i and G_i^*，时空同期权重矩阵，时空滞后权重矩阵。**

El espacio y el tiempo son dimensiones características cruciales de los eventos y fenãmenos geográficos. Aunque el análisis de datos espaciales con carácter exploratorio (ESDA) se puede usar para visualizar y resumir patrones espaciales complejos, tiene limitaciones para captar la dinámica temporal de rasgos geográficos. Se han hecho esfuerzos para incorporar la dimensiãn temporal dentro de las técnicas ESDA para detectar agrupamiento o tendencias del espacio–tiempo. Sin embargo, todavía faltan las estadísticas localizadas del espacio–tiempo que pudieren ayudar en el análisis exploratorio de datos del espacio–tiempo (ESTDA). Enfocándonos en el panel de datos espaciales, nuestro trabajo extendiã las estadísticas Getis–Ord G_i y G_i^* usando una matriz de peso contemporánea del espacio–tiempo y una matriz de peso rezagada del espacio–tiempo para dar cuenta de la autocorrelaciãn del espacio–tiempo local. Dos aplicaciones consideradas en este artículo muestran que el método desarrollado recientemente puede usarse para resumir patrones del espacio–tiempo desde el panel de datos espaciales, identificar de manera más consistente los cambios del paisaje y prestar los resultados de inmediato para visualizaciãn. **Palabras clave: análisis exploratorio de datos del espacio–tiempo (ESTDA), autocorrelaciãn del espacio–tiempo local, espacio–tiempo G_i y G_i^*, matriz de peso contemporánea del espacio–tiempo, matriz de peso rezagada del espacio–tiempo.**

With the development of the Global Positioning System, remote sensing, and geographic information systems (GIS), the past several decades have witnessed a remarkable increase in the production of geographically referenced data and the development of a large variety of spatial analytical techniques (Anselin 1995, 2012; Goodchild et al. 2000; Goodchild and Haining 2004). In this context, exploratory spatial data analysis (ESDA) has become prevalent due to its strength in revealing complex spatial phenomena. As defined by Anselin (1999), ESDA is a set of techniques to "describe and visualize spatial distributions, identify atypical locations (spatial outliers), describe patterns of spatial association (spatial clusters), and suggest different spatial regimes and other forms of spatial instability or spatial non-stationarity" (258). Fundamental to ESDA is the concept of spatial autocorrelation that follows Tobler's first law of geography: Everything is related to everything else, but near things are more related than distant things (Tobler 1970; Anselin 1999). A fair amount of global and local indexes of spatial autocorrelation, such as Moran's *I*, local indicators of spatial association (LISA), and Getis–Ord statistics, have been developed and applied in describing spatial patterns, detecting spatial clusters, and suggesting spatial regimes.

Geographic events and phenomena "have both spatial and temporal dimensions that cannot be meaningfully separated" (Miller 2004, 648). Although ESDA can be used to analyze and visualize complex spatial patterns, it has limitations in capturing the temporal dynamics of geographic features (Ye 2010). In this research, we focus on spatial panel data (also known as spatial panels), which are composed of repeated observations over time on the same set of cross-sectional (geographic) units (e.g., countries, states, counties, census tracts, and ZIP codes).

The Professional Geographer, 0(0) 2020, pages 1–10 © 2020 by American Association of Geographers.
Initial submission, August 2019; revised submissions, October and November 2019; final acceptance, November 2019.
Color versions of one or more figures in the article can be found online at www.tandfonline.com/rtpg
Published by Taylor & Francis Group, LLC.

Compared to cross-sectional data, spatial panel data contain more variation and are generally more informative because of the added time dimension (Elhorst 2003). Many socioeconomic indicators (e.g., annual gross domestic product by state in the United States) that are important for geography, regional science, public health, and other disciplines are actually spatial panels. Although spatial panel data contain richer space–time information than cross-sectional data, ESDA does not use the temporal dimension, and it only focuses on a single cross section. In this regard, it calls for the advancement of exploratory space–time data analysis (ESTDA) to go beyond the static data and better reveal the space–time dynamics of various geographic events and phenomena.

Efforts have been made to incorporate the temporal dimension into ESDA. Here, we identify two important threads of research effort. The first line of research deals with the temporal dynamics of ESDA results (Rey 2001; Rey, Janikas, and Smirnov 2005). For example, Rey (2001) developed a spatial Markov transition matrix by integrating the classification of local Moran statistics into Markov chain modeling. In this way, Rey (2001) compared the regional income distributions (gained from local Moran's *I*) at different time points. Additionally, Rey, Janikas, and Smirnov (2005) formulated a LISA time–path plot to demonstrate how a region's income evolved over time with its neighbors. For each region in a time series, its LISA time–path plot was obtained by the following steps: (1) calculating its spatial lag (the average of the region's first-order neighbors) at each time point; (2) plotting the region and its spatial lag as an individual point in an x–y axis; and (3) connecting all of the points over time to obtain the LISA time path for this region. The pairwise movement revealed by the LISA time path can help in the identification of stability levels of local economic growth.

Different from the first thread, the second line of research consists of multiple space–time clustering methods developed by modifying the spatial algorithms to incorporate the time dimension. Some examples include space–time scan statistics (Kulldorff 2001; Kulldorff et al. 2005), space–time kernel density estimation (Nakaya and Yana 2010; Lee, Gong, and Li 2017), the space–time Ripley's K function (Bailey and Gatrell 1995), and the space–time nearest neighbor test by Jacquez and collaborators (Jacquez 1996; Jacquez, Goovaerts, and Rogerson 2005). Although autocorrelation-based spatial statistics are well-established and widely used, the measurement of space–time autocorrelation has not received sufficient attention in the statistical analysis of space–time clustering or trends. The concept of space–time autocorrelation can be dated back to 1975 with the pioneering work of Cliff and Ord (1975). Since then, the measurement of global spatiotemporal autocorrelation has been sporadically documented in the literature (Cliff and Ord 1981; Griffith 1981; Bertazzon 2003; Dubé and Legros 2013). In particular, local space–time autocorrelation statistics have not been formulated until recently, with most of the work focusing on extending Moran's *I* (Shen, Li, and Si 2016; Lee and Li 2017). Some important alternatives, such as Getis–Ord statistics, have rarely been explored (Griffith 2018). Introduced in Getis and Ord (1992) and Ord and Getis (1995), G_i and G_i^* can identify local patterns in spatial data and distinguish between hot spots (locations of high spatial associations) and cold spots (locations of low spatial associations). As stated by Songchitruksa and Zeng (2010), "While the Moran's *I* either confirms the site of interest as part of its surrounding sites (in cluster) or distinguishes the site from the cluster, it cannot discriminate between patterns that are high-value dominant or low-value dominant" (43). Therefore, extending G_i and G_i^* into the space–time domain, which is the objective of this article, would seem to be promising.

The Emerging Hot Spot Analysis tool in ArcGIS (ESRI 2019) provides a way to extend Getis–Ord statistics to the space–time domain. Users are not well informed of how their space–time adjacency structure is conceptualized, where their specification of space–time weight matrix is derived from, and what intellectual relationship it has with existing efforts, however. Two conceptualizations of space–time dependency structure, contemporaneous and lagged (see "Space–Time Weight Matrix" for more details), have been documented in the literature, and their mathematical expressions have been adopted in the measurement of global space–time autocorrelation (Cliff and Ord 1981; Griffith 1981; Griffith and Paelinck 2018). As recently noted by Griffith (2018), though, these two conceptualizations or specifications have rarely been used in local space–time autocorrelation statistics, because they were developed long before the appearance of local Moran's *I*, Getis–Ord G_i and G_i^*, and other localized spatial statistics. Our study fills this gap by incorporating the two space–time specifications into Getis–Ord G_i and G_i^* to account for local space–time autocorrelation.

Getis–Ord G_i and G_i^*

The null hypothesis of G_i is that the sum of observations at i's ($i = 1, 2, \ldots, n$) neighbors is not larger (or smaller) than one would expect by chance given all of the observations across the entire region (Getis and Ord 1996). Neighborness here is often defined in two ways. One is based on whether two locations (areal units) share a common boundary, whereas the other depends on whether they are within a specified distance from each other. Using either of these two methods, we can get a spatial weight matrix W_s where each entry w_{ij} equals 1 if locations i and j are

geographic neighbors and 0 otherwise. Notably, a location i is not defined as its own neighbor ($w_{ii} = 0$) in G_i. The G_i statistic is given as follows:

$$G_i(W_s) = \frac{\sum_{j,j\neq i} w_{ij}x_j}{\sum_{j,j\neq i} x_j}.$$

Different from the null hypothesis of G_i, the G_i^* statistic does not exclude the observation at i but sums it together with its neighbors. Therefore, a location i is also defined as its own neighbor in G_i^* ($w_{ii} = 1$). The G_i^* statistic is given by the following equation:

$$G_i^*(W_s) = \frac{\sum_j w_{ij}x_j}{\sum_j x_j}.$$

The interpretation of $G_i(W_s)$ and $G_i^*(W_s)$ is straightforward: A positive standardized z value suggests a hot spot (i.e., concentration of high-value locations), whereas a negative z value indicates a cold spot (i.e., concentration of low-value locations). Note that $G_i^*(W_s)$ has been more widely used than $G_i(W_s)$ in practice (Braithwaite and Li 2007).

Space–Time Weight Matrix

Central to the space–time extensions of G_i and G_i^* is the conceptualization and specification of space–time adjacency structure. Griffith (2010, 2012) suggested two conceptualizations of space–time dependency structure that include contemporaneous specification (Figure 1A) and lagged specification (Figure 1B). According to Griffith (2012),

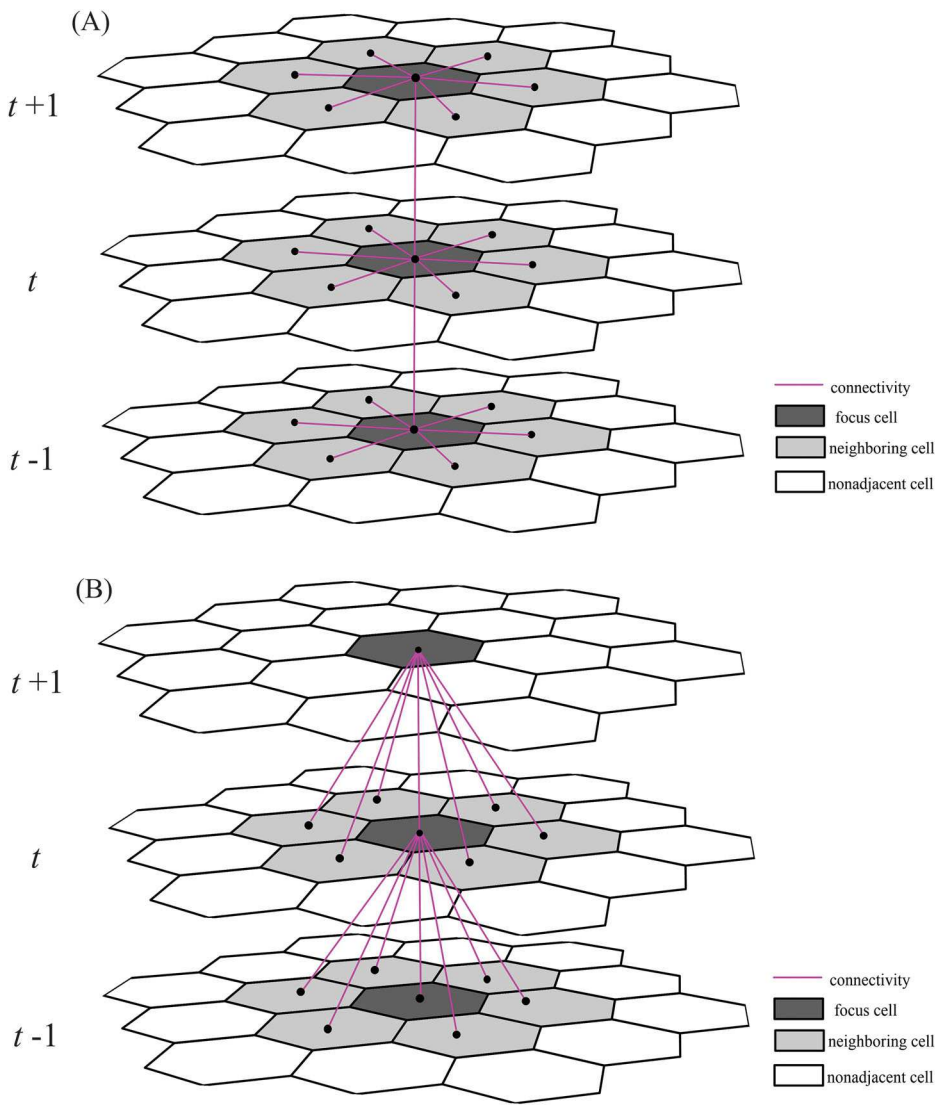

Figure 1 *Two conceptualizations of space–time dependency structure (modified from Griffith 2012): (A) space–time contemporaneous dependency and (B) space–time lagged dependency.*

space–time contemporaneous specification indicates that "a location i links to its preceding *in situ* location as well as the instantaneous neighboring locations," whereas the space–time lagged specification is that "a location i links to its preceding *in situ* location as well as the preceding neighboring locations" (23). Matrix versions of these two conceptualizations have been developed (Griffith 2010, 2012; Griffith and Heuvelink 2012; Griffith and Paelinck 2018) and are given here.

The space–time contemporaneous specification is

$$V_{ST} = I_T \otimes W_S + W_T \otimes I_s.$$

The space–time lagged specification is

$$V_{ST} = W_T \otimes W_S + W_T \otimes I_s,$$

where V_{ST} denotes the space–time connectivity matrix, $\otimes$ is the Kronecker product, I_T is a T by T identity matrix (T is the number of cross-sections in the spatial panel data), W_S is the same n by n spatial weight matrix as in the static G_i and G_i^*, W_T represents the T by T time series connectivity matrix (the upper and lower off-diagonal cells contain 1, and all other cells contain 0), and I_s is the n by n identity matrix. The total number of observations is $N = nT$. Note that the element v_{ij} in V_{ST} equals 1 if i and j are neighbors in space, time, or both and 0 otherwise. Using spatial panel data with 100 cross-sectional units and eight time points (800 observations), for example, the dimensions of these matrices including V_{ST}, I_T, W_S, W_T, and I_s will then be 800 by 800, 8 by 8, 100 by 100, 8 by 8, and 100 by 100, respectively.

An obvious advantage of these two specifications over those in other studies lies in their conciseness and simplicity. They are delivered through matrix operations that demonstrate clearly the extension of a spatial weight matrix (can be acquired with many GIS software programs) to its space–time version, and these matrix operations can be easily performed with many programming tools such as R, Matlab, and Python.

These two specifications have been adopted to calculate the global space–time Moran index (Griffith and Paelinck 2018) and conduct eigenvector space–time filtering (Griffith 2010, 2012). Griffith and Paelinck (2018) suggested that when $T = 1$, space–time contemporaneous specification will reduce to a static (spatial) weight matrix, and the lagged specification will vanish. After assessing the spatial statistical properties of these specifications for the global Moran index, they further stated that the contemporaneous specification is preferred over the lagged one (Griffith and Paelinck 2018). In light of this, we adopted the space–time contemporaneous structure to derive space–time G_i and G_i^* in this research.

Space–Time G_i and G_i^*

Modified from its spatial counterpart, the null hypothesis of space–time G_i can be specified as follows: The sum of observations at i's ($i = 1, 2, \ldots, N$) space–time neighbors (excluding i) is not more extreme (either larger or smaller) than one would expect by chance given all of the observations (excluding i). Replacing the spatial weight w_{ij} with space– time weight v_{ij}, space–time G_i can be modified as

$$G_i(V_{ST}) = \frac{\sum_{j,j\neq i} v_{ij} x_j}{\sum_{j,j\neq i} x_j}.$$

Likewise, the null hypothesis of space–time G_i^* can be stated as follows: The sum of observations at i's ($i = 1, 2, \ldots, N$) space–time neighbors (including i) is not more extreme (either larger or smaller) than one would expect by chance given all of the observations (including i). Substituting the spatial weight w_{ij} with space–time weight v_{ij}, space– time G_i^* is given as follows:

$$G_i^*(V_{ST}) = \frac{\sum_j v_{ij} x_j}{\sum_j x_j}.$$

Similar interpretation can be used to interpret $G_i(V_{ST})$ and $G_i^*(V_{ST})$: A positive standardized z value suggests a space–time hot spot, whereas a negative z value indicates a space–time cold spot.

Inference

Both normal approximation and conditional permutation have been used in the inference of Getis–Ord G_i and G_i^*. With normal approximation, Ord and Getis (1995) derived the expectations and variances of $G_i(W_s)$ and $G_i^*(W_s)$ and redefined each statistic as a standardized z value by subtracting the statistic from its expectation, divided by the standard deviation. Similarly, we can derive a standardized version of $G_i(V_{ST})$ and $G_i^*(V_{ST})$ using the normal approximation. For the derivation of standardized $G_i(V_{ST})$, we begin with the calculation of the mean and variance of x:

$$\bar{x}(i) = \frac{1}{N-1} \sum_{j,j\neq i} x_j$$

$$s_N(i)^2 = \frac{\sum_{j,j\neq i} x_j^2}{N-1} - [\bar{x}(i)]^2.$$

The expected value of $G_i(V_{ST})$ is

$$E(G_i(V_{ST})) = \frac{V_i}{N-1}$$

with a variance of

$$\mathrm{Var}(G_i(V_{ST})) = \frac{V_i(N-1-V_i)}{(N-1)^2(N-2)}\left(\frac{s_N(i)}{\bar{x}(i)}\right)^2,$$

resulting in the following standardized form:

$$G_i(V_{ST}) = \frac{\sum_j v_{ij}x_j - V_i\bar{x}(i)}{s_N(i)\left\{[(N-1)S_{N1i} - V_i^2]/(N-2)\right\}^{1/2}},$$

where $V_i = \sum_{j\neq i} v_{ij}$

$$S_{N1i} = \sum_{j,j\neq i} v_{ij}^2.$$

For the standardized $G_i^*(V_{ST})$, we also begin with the calculation of the mean and variance of x

$$\bar{x}^* = \frac{\sum_j x_j}{N}$$

$$s_N^{*2} = \frac{\sum_j x_j^2}{N} - (\bar{x}^*)^2.$$

The expected value of $G_i^*(V_{ST})$ is

$$E\left(G_i^*(V_{ST})\right) = \frac{V_i^*}{N}$$

with a variance of

$$\mathrm{Var}(G_i(V_{ST})) = \frac{V_i^*(N-V_i^*)}{N^2(N-1)}\left(\frac{s_N^*}{\bar{x}^*}\right)^2,$$

resulting in the following standardized form:

$$G_i^*(V_{ST}) = \frac{\sum_j v_{ij}x_j - V_i^*\bar{x}^*}{s_N^*\left\{[NS_{N1i}^* - V_i^{*2}]/(N-1)\right\}^{1/2}}$$

$$V_i^* = \sum_{j\neq i} v_{ij} + v_{ii} = V_i + v_{ii}$$

$$S_{N1i}^* = \sum_j v_{ij}^2.$$

Because the normal approximation method is less reliable when the number of observations is small ($n < 100$; Anselin 2019), we used conditional permutation as the inference approach for space–time G_i and G_i^*. This consists of fixing the value at i, randomly permuting all remaining values except the one at i for m times, and recalculating the local statistic at i. In this way, we have $m+1$ local statistics (including the observed one) for i and form an empirical reference distribution under the null hypothesis. A standardized z value for each local statistic can then be obtained, and inference can also be made with a pseudo p value of a one-sided test represented by the proportion of local statistics in the reference distribution that are more extreme (either larger or smaller) than the observed one (Anselin 2019).

Just like what local spatial statistics suffer from, the interpretation of pseudo p values in space–time G_i and G_i^* can be affected by multiple comparisons (also known as multiplicity; de Castro and Singer 2006; Anselin 2019). Our space–time local analysis involves multiple inferences and tests from a given spatial panel data set (i.e., each i is given a test) and therefore the resulting pseudo p values might not correctly reflect Type I error (rejection of the true null hypothesis). For example, if $N = 1,000$ null hypotheses are simultaneously tested and all of them are true, given that each test's significance level is set to $\alpha = 0.01$, then ten rejections of true hypotheses would be expected.

There have been several corrections for multiple comparisons. A simple but very conservative approach is the Bonferroni method (Bonferroni 1936). With an overall target p value of α (i.e., overall probability of Type I error) for N simultaneous tests, the Bonferroni bound for each individual test would be α/N. Then, a test is taken as significant when its p value or pseudo p value is no more than α/N. Also controlling the overall probability of Type I error, the corresponding bound for each individual test in Sidak correction is $1-(1-\alpha)^{1/N}$ (Sidak 1967). Both Bonferroni and Sidak methods are conservative multiple comparison procedures that can reduce the possibility of rejecting the true null hypothesis but will also increase the possibility of accepting a false null hypothesis.

In light of this, de Castro and Singer (2006) suggested using an alternative and preferable approach (i.e., false discovery rate [FDR]) proposed by Benjamini and Hochberg (1995). FDR is obtained with a stepwise procedure. First, the pseudo p values (p_i) of the test statistics are sorted in ascending order ($p_1 \leq p_2 \leq p_3 \cdots \leq p_N$). Second, starting from p_N, the first p_i that meets $p_i \leq (i/N)\alpha$ is identified. Third, this p_i is set as the critical p value and all tests are considered significant if their pseudo p values are no more than this value. FDR is less conservative than the Bonferroni and Sidak methods and thus can identify more locations with significant local space–time autocorrelation while controlling for false positives. Hence, we used FDR in this work. We note that different approaches to calculate FDR could affect the number of clusters being detected (Bivand, Müller, and Reder 2009), and future research could be conducted to compare the various methods.

We applied the newly modified space–time G_i^* to analyze two spatial panel data sets, including the U.S. state-level unemployment rates between 2003 and 2017 and county-level Hurricane Sandy tweets (posted between 23 October 2012 and 1 November 2012) in the eastern United States. Of note is that, in both cases, we used queen contiguity to obtain W_S for calculation of V_{ST}.

Application I

The first application examined how Twitter responses to Hurricane Sandy at a community level (here, county) changed across space and over time. Hurricane Sandy was named a tropical storm on 22 October 2012, made its landfall in New Jersey on 29 October 2012, and dissipated on 2 November 2012. Sandy's track between 21 October and 31 October was obtained from the National Oceanic and Atmospheric Administration (n.d.) and mapped in Figure 2. We collected 83,006 geotagged Sandy tweets and aggregated them spatially to each county in fourteen eastern states and temporally to each day from 23 October 2012 to 1 November 2012. Thus, we obtained spatial panel data including 626 cross-sectional units (counties) and ten time points, where each observation represents the number of Hurricane Sandy tweets posted from a given county on a certain day. Note that each observation in this spatial panel was divided by the local population to reduce its impact because larger population tends to yield more social media activity. For more details regarding Sandy tweet collection, please refer to Wang et al. (2019). Setting the overall target p value to 0.01 and using 9,999 permutations, the critical p value obtained via FDR is 0.00203.

Figure 3 shows a clear space–time trend and clustering of Hurricane Sandy tweets. We can easily relate the space–time dynamics of Sandy Twitter activities to the hurricane movements. As Sandy headed northward and approached New Jersey, more and more tweets were posted, which resulted in the decrease in cold spots. Hot spots emerged on 28 October, expanded when Sandy made its landfall, and shrank a day later with the decreasing intensity of the storm. Compared with the hot and cold spot patterns computed from individual cross sections (not included in this article), the hot and cold spot patterns derived from the new space–time index are more distinct and the changes through time are clearly revealed and visualized. Using the space–time indexes, we are able to track the hot and cold spots emerging and disappearing in a more consistent and reliable manner.

Application II

The second application explored the space–time dynamics of unemployment across the conterminous United States (including Washington, DC) during the time period from 2003 to 2017. The annual state-level unemployment rates were obtained from the Bureau of Labor Statistics (n.d.). Setting the

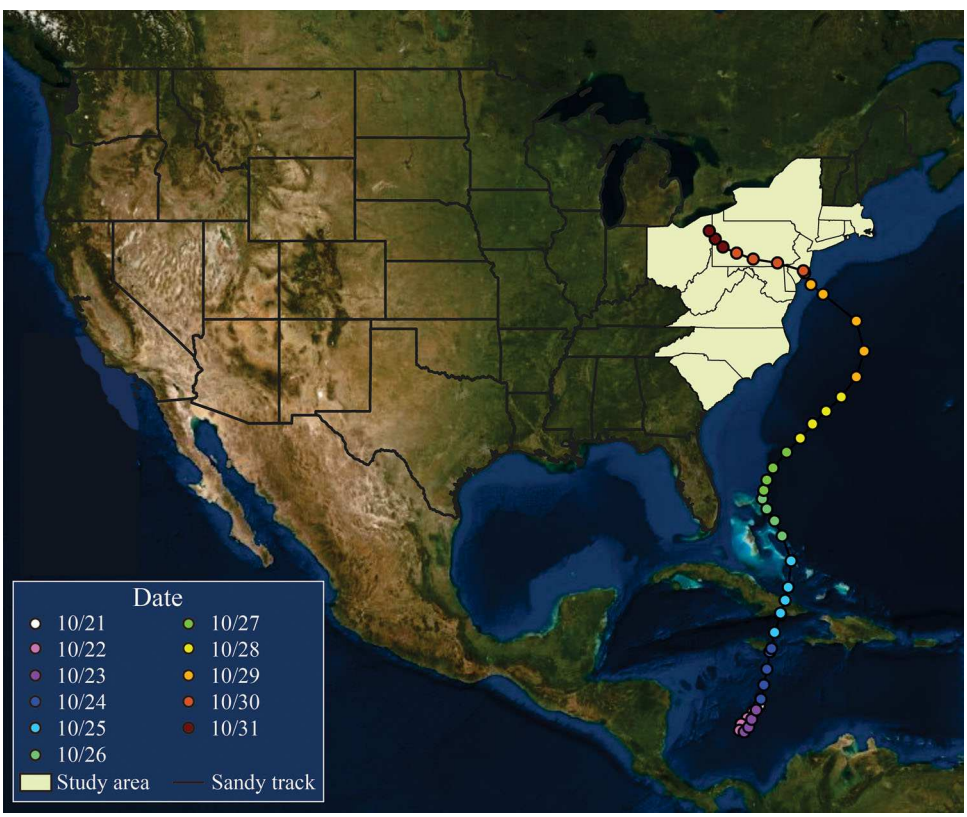

Figure 2 *Study area and Hurricane Sandy track.*

overall target p value to 0.01 and using 9,999 permutations, the critical p value obtained via FDR in these data is 0.0017. Figure 4 shows the temporal dynamics of state-level unemployment rate. As seen from Figure 4, cold spots of unemployment first increased and then decreased between 2003 and 2008, and the year 2008 marks a notable transition of the unemployment landscape. The year 2008 can

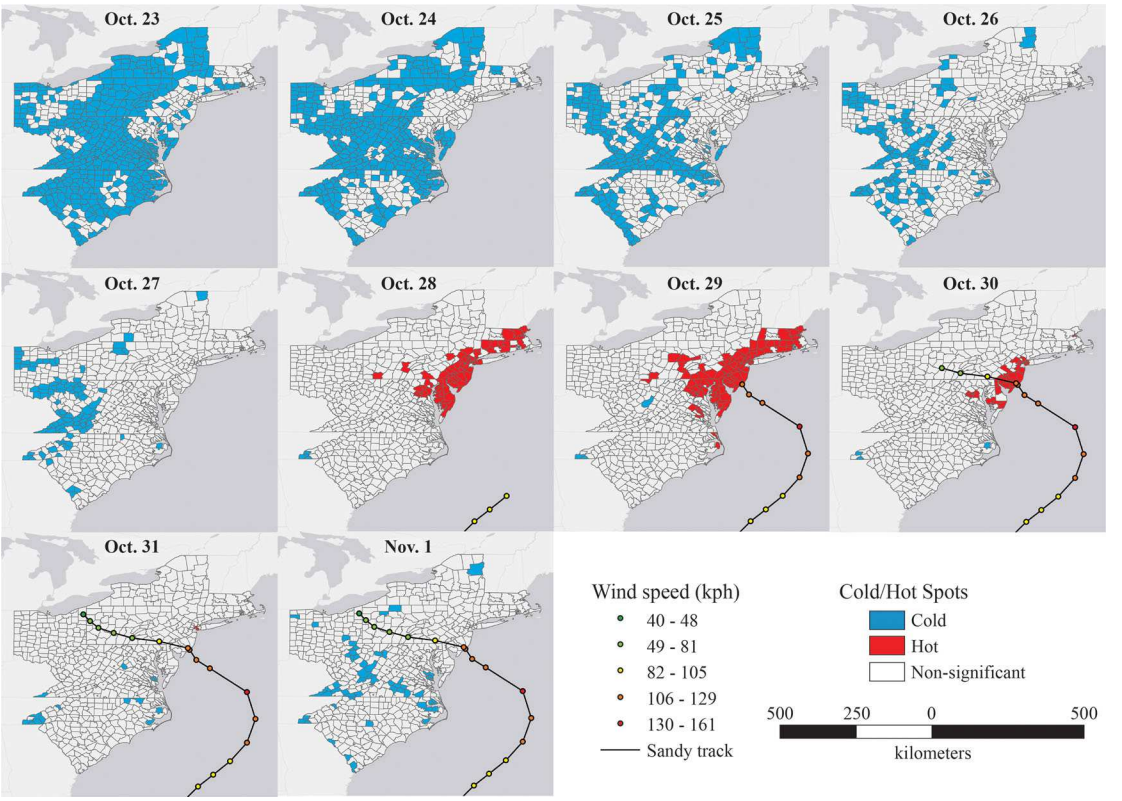

Figure 3 *Space–time hot and cold spots of Hurricane Sandy tweets.*

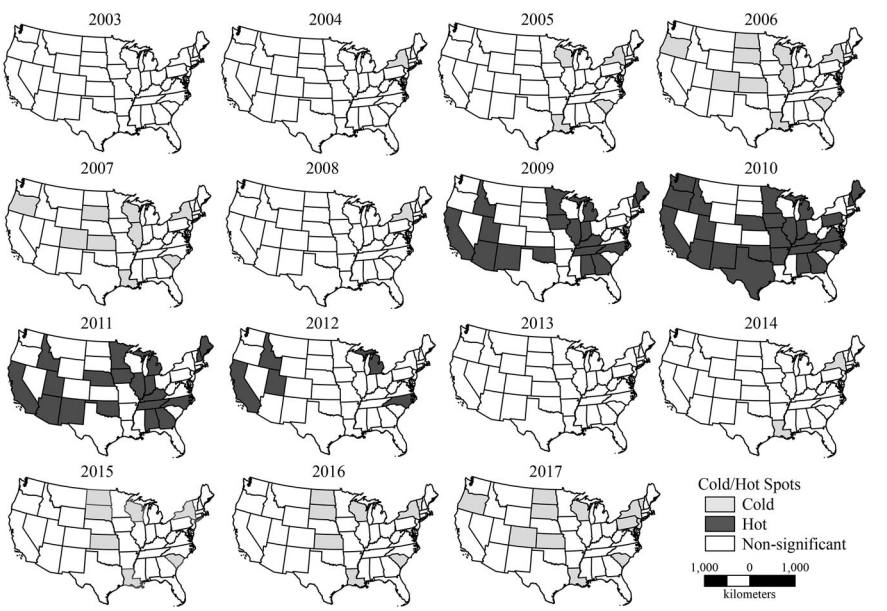

Figure 4 *Space–time hot and cold spots of unemployment rate.*

be viewed as a point of change, because it connected two largely different unemployment landscapes (featured by cold spots in 2007 vs. hot spots in 2009). This pattern can be linked with the 2008 financial crisis, which caused a recession and increased unemployment in the following years. Due to this economic disaster, hot spots of unemployment emerged in 2009, and the unemployment landscape continued to be prominently featured until 2012. The year 2013 marks another transition of unemployment landscapes where hot spots diminished and cold spots were about to return. From 2014 to 2017, economic recoveries reduced the unemployment rate, and cold spots started to show in the unemployment landscape.

Concluding Remarks

We extended the local tests for spatial autocorrelation in Getis and Ord (1992) to a space–time context. The extensions to space–time G_i and G_i^* were accomplished by modifying Getis–Ord statistics to incorporate the space–time adjacency structure originally developed by Griffith (1981). The modification is straightforward and easy to implement. Two applications show that our method can summarize space–time patterns in panel-formatted geographic observations and lend the results readily to visualization. The visualization of space–time clusters in each cross section gives new insights into the temporal transitions of geographic phenomena; the method can thus serve as an effective tool for identifying the point of landscape change. As argued in the Introduction, there is a lack of local tests of space–time clustering in the literature; therefore, our work can be a useful addition to the autocorrelation-based ESTDA techniques.

There is an increasing awareness of the multiplicity problem in local spatial autocorrelation measures. People are realizing that multiple comparisons can result in false spatial clusters and misleading conclusions. In this context, GIS software programs such as GeoDa and ESRI ArcMap have incorporated corrections in their spatial statistical components to address this issue. Multiplicity, however, has received less attention in the local tests of space–time autocorrelation than that of their spatial counterparts. We used the FDR approach in this work to reflect Type I error and identify as many true space–time clusters as possible. Our work is an improvement over existing measurements of local spatiotemporal autocorrelation that overlooked multiple testing.

Simultaneous modeling of local space–time effects can reveal many more details of spatial and temporal processes in geographic phenomena, although the modeling could be subjected to computational obstacles because of the typically large spatial–temporal data sets. Therefore, future efforts could be made to improve the efficiency of algorithms and leverage high-performance computing power to accelerate the computation. Another scalability problem arises when the data set comes with a large number of time points (e.g., 10,000). In this situation, mapping all of the results at every time point will not be feasible. Future work could use time series analysis (e.g., change point detection) to summarize the temporal trends of hot and cold spots. ■

Funding

This article is based on work supported in part by two research grants from the U.S. National Science Foundation: one under the Interdisciplinary Behavioral and Social Science Research (IBSS) Program (Award No. 1620451) and the other under the National Science Foundation Social and Economic Sciences Division (SES) Hurricane Harvey 2017 Program (Award No. 1762600).

ORCID

Zheye Wang http://orcid.org/0000-0002-2037-9850
Nina S. N. Lam http://orcid.org/0000-0002-5344-9368

Literature Cited

Anselin, L. 1995. Local indicators of spatial association—LISA. *Geographical Analysis* 27 (2):93–115. doi: 10.1111/j.1538-4632.1995.tb00338.x.

Anselin, L. 1999. Interactive techniques and exploratory spatial data analysis. *Geographical Information Systems: Principles, Techniques, Management and Applications* 1:251–64.

Anselin, L. 2012. From SpaceStat to CyberGIS: Twenty years of spatial data analysis software. *International Regional Science Review* 35 (2):131–57. doi: 10.1177/0160017612438615.

Anselin, L. 2019. A local indicator of multivariate spatial association: Extending Geary's *C*. *Geographical Analysis* 51 (2):133–50. doi: 10.1111/gean.12164.

Bailey, T. C., and A. C. Gatrell. 1995. *Interactive spatial data analysis*. Vol. 413. Essex, UK: Longman Scientific and Technical.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57 (1):289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x.

Bertazzon, S. 2003. Spatial and temporal autocorrelation in innovation diffusion analysis. In *International Conference on Computational Science and Its Applications*—ICCSA

2003, ed. V. Kumar, M. L. Gavrilova, C. J. K. Tan, and P. L'Ecuyer, 23–32. Berlin: Springer.

Bivand, R., W. G. Müller, and M. Reder. 2009. Power calculations for global and local Moran's *I*. *Computational Statistics & Data Analysis* 53 (8):2859–72. doi: 10.1016/j.csda.2008.07.021.

Bonferroni, C. 1936. Teoria statistica delle classi e calcolo delle probabilita [Statistical theory of classes and calculus of probabilities]. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze* 8:3–62.

Braithwaite, A., and Q. Li. 2007. Transnational terrorism hot spots: Identification and impact evaluation. *Conflict Management and Peace Science* 24 (4):281–96. doi: 10.1080/07388940701643623.

Bureau of Labor Statistics. n.d. Local area unemployment statistics. Accessed September 1, 2019. https://www.bls.gov/lau/

Cliff, A., and J. K. Ord. 1975. Space–time modelling with an application to regional forecasting. *Transactions of the Institute of British Geographers* (64):119–28. doi: 10.2307/621469.

Cliff, A., and J. Ord. 1981. Spatial and temporal analysis: Autocorrelation in space and time. In *Quantitative geography: A British view*, ed. N. Wrigley and R. Bennett, 104–10. London and New York: Routledge and Kegan Paul.

de Castro, M. C., and B. H. Singer. 2006. Controlling the false discovery rate: A new application to account for multiple and dependent tests in local statistics of spatial association. *Geographical Analysis* 38 (2):180–208. doi: 10.1111/j.0016-7363.2006.00682.x.

Dubé, J., and D. Legros. 2013. A spatiotemporal measure of spatial dependence: An example using real estate data. *Papers in Regional Science* 92 (1):19–30. doi: 10.1111/j.1435-5957.2011.00402.x.

Elhorst, J. P. 2003. Specification and estimation of spatial panel data models. *International Regional Science Review* 26 (3):244–68. doi: 10.1177/0160017603253791.

Environmental Systems Research Institute. 2019. How emerging hot spot analysis works. Accessed July 18, 2019. https://pro.arcgis.com/en/pro-app/tool-reference/space-time-pattern-mining/learnmoreemerging.htm.

Getis, A., and J. K. Ord. 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis* 24 (3):189–206. doi: 10.1111/j.1538-4632.1992.tb00261.x.

Getis, A., and J. K. Ord. 1996. Local spatial statistics: An overview. In *Spatial analysis: Modelling in a GIS environment*, ed. P. Longley and M. Batty, 261–77. Cambridge, UK: GeoInformation International.

Goodchild, M. F., L. Anselin, R. P. Appelbaum, and B. H. Harthorn. 2000. Toward spatially integrated social science. *International Regional Science Review* 23 (2):139–59. doi: 10.1177/016001760002300201.

Goodchild, M. F., and R. P. Haining. 2004. GIS and spatial data analysis: Converging perspectives. *Papers in Regional Science* 83 (1):363–85. doi: 10.1007/s10110-003-0190-y.

Griffith, D. A. 1981. Interdependence in space and time: Numerical and interpretative considerations. In *Dynamic spatial models*, ed. D. Griffith and R. MacKinnon, 258–87. Ahhen aan den Riin: Siithoff Noordhoff.

Griffith, D. A. 2010. Modeling spatio-temporal relationships: Retrospect and prospect. *Journal of Geographical Systems* 12 (2):111–23. doi: 10.1007/s10109-010-0120-x.

Griffith, D. A. 2012. Space, time, and space–time eigenvector filter specifications that account for autocorrelation. *Estadística Española* 54 (177):7–34.

Griffith, D. A. 2018. A comment on J. Lee and S. Li's "Extending Moran's Index for measuring spatiotemporal clustering of geographic events" by Daniel A. Griffith. *Geographical Analysis* 50 (4):477–78.

Griffith, D. A., and G. B. Heuvelink. 2012. Deriving space–time variograms from space–time autoregressive (STAR) model specifications. In *Advances in spatial data handling and GIS*, ed. A. Yeh, J. Shi, Y. Leung, and C. Zhou, 3–12. Berlin: Springer.

Griffith, D. A., and J. H. Paelinck. 2018. Space–time autocorrelation. In *Morphisms for quantitative spatial analysis*, ed. D. A. Griffith and J. H. P. Paelinck, 25–34. Cham, Switzerland: Springer.

Jacquez, G. M. 1996. A *k* nearest neighbour test for space–time interaction. *Statistics in Medicine* 15 (18): 1935–49. doi: 10.1002/(SICI)1097-0258(19960930)15:18<1935::AID-SIM406>3.0.CO;2-I.

Jacquez, G. M., P. Goovaerts, and P. Rogerson. 2005. Space–time intelligence systems: Technology, applications and methods. *Journal of Geographical Systems* 7 (1): 1–5. doi: 10.1007/s10109-005-0146-7.

Kulldorff, M. 2001. Prospective time periodic geographical disease surveillance using a scan statistic. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 164 (1):61–72. doi: 10.1111/1467-985X.00186.

Kulldorff, M., R. Heffernan, J. Hartman, R. Assunçao, and F. Mostashari. 2005. A space–time permutation scan statistic for disease outbreak detection. *PLoS Medicine* 2 (3): e59. doi: 10.1371/journal.pmed.0020059.

Lee, J., J. Gong, and S. Li. 2017. Exploring spatiotemporal clusters based on extended kernel estimation methods. *International Journal of Geographical Information Science* 31 (6):1154–77. doi: 10.1080/13658816.2016.1170133.

Lee, J., and S. Li. 2017. Extending Moran's index for measuring spatiotemporal clustering of geographic events. *Geographical Analysis* 49 (1):36–57. doi: 10.1111/gean.12106.

Miller, H. J. 2004. Activities in space and time. In *Handbook of transportation research 5: Transport geography and spatial systems*, ed. P. Stopher, K. Button, K. Haynes, and D. Hensher, 647–60. London: Pergamon/Elsevier Science.

Nakaya, T., and K. Yano. 2010. Visualising crime clusters in a space–time cube: An exploratory data analysis approach using space–time kernel density estimation and scan statistics. *Transactions in GIS* 14 (3):223–39. doi: 10.1111/j.1467-9671.2010.01194.x.

National Oceanic and Atmospheric Administration. n.d. Historical hurricane tracks. Accessed September 1, 2019. https://coast.noaa.gov/hurricanes/

Ord, J. K., and A. Getis. 1995. Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis* 27 (4):286–306. doi: 10.1111/j.1538-4632.1995.tb00912.x.

Rey, S. J. 2001. Spatial empirics for economic growth and convergence. *Geographical Analysis* 33 (3):195–214. doi: 10.1111/j.1538-4632.2001.tb00444.x.

Rey, S. J., M. Janikas, and O. Smirnov. 2005. Exploratory geovisualization of spatial dynamics. Geocomputation

2005 Proceedings, ed. Y. Xie and D. G. Brown. (CD-ROM). Ann Arbor: University of Michigan.

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. doi: 10.1016/j.physleta.2015.09.039.

Sidak, Z. 1967. Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association* 62 (318):626–33. doi: 10.2307/2283989.

Songchitruksa, P., and X. Zeng. 2010. Getis–Ord spatial statistics to identify hot spots by using incident management data. *Transportation Research Record: Journal of the Transportation Research Board* 2165 (1):42–51. doi: 10.3141/2165-05.

Tobler, W. R. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46 (Suppl. 1):234–40. doi: 10.2307/143141.

Wang, Z., N. S. Lam, N. Obradovich, and X. Ye. 2019. Are vulnerable communities digitally left behind in social responses to natural disasters? An evidence from Hurricane Sandy with Twitter data. *Applied Geography* 108:1–8. doi: 10.1016/j.apgeog.2019.05.001.

Ye, X. 2010. Comparative space time dynamics. PhD thesis, University of California, Santa Barbara, and San Diego State University.

ZHEYE WANG is a Postdoctoral Research Associate in the Department of Environmental Sciences, College of the Coast and Environment, Louisiana State University, Baton Rouge, LA 70803. E-mail: zwang3603@gmail.com. His research interests include space–time statistics, location-based social media, disaster vulnerability and resilience, and spatial data science.

NINA S. N. LAM is Abraham Distinguished Professor of Louisiana Environmental Studies in the Department of Environmental Sciences, College of the Coast and Environment, Louisiana State University, Baton Rouge, LA 70803. E-mail: nlam@lsu.edu. Her research interests include geographic information science (GIScience), remote sensing, spatial analysis, environmental health, disaster resilience, and sustainability.