Measuring Population Shift Bias in Tests of Space-Time Interaction

Nicholas Malizia, Elizabeth A. Mack, and Sergio J. Rey

GeoDa Center for Geospatial Analysis and Computation
School of Geographical Sciences and Urban Planning
Arizona State University
975 S. Myrtle Ave. Tempe, AZ 85287-5302, USA
Email: nmalizia@asu.edu

1 Introduction

Tests of space-time interaction detect clustering of events in space and time in excess of "any purely spatial or purely temporal clustering" (Kulldorff, 1998, pg. 58). These tests are widely employed in studies of crime (e.g. Knox, 2002; Grubesic and Mack, 2008) and disease (e.g. Petridou et al., 1996; Rogerson, 2001). By simultaneously considering both the spatial and temporal dimensions of the event patterns, these methods are capable of identifying certain data generating processes and, as a result, are often used to inform etiological work (Ward and Carpenter, 2000). Most of these tests, however, dubiously assume the underlying susceptible population within a study area to be invariant through time and across space. In settings where this assumption does not hold, these tests will detect space-time interaction due to population changes in addition to interaction resulting from the data generating process of interest. The excess interaction observed due to violating this assumption and by failing to account for the changes in the underlying population is referred to as population shift bias (Kulldorff and Hjalmars, 1999). Although recognized, this bias is often not accounted for in practice and its potential impact on results is not fully explored. This paper carries out a simulation to develop a detailed understanding of the impact of population shift bias on three of the most common tests of space-time interaction: the Knox (1964), Mantel (1967), and Jacquez (1996) tests. Additionally, the simulation demonstrates that contrary to prior claims (i.e. Kulldorff and Hjalmars, 1999; Aldstadt, 2007), population shift bias is problematic even in studies with a short temporal extent. To these ends, we simulate events within the dynamic population of a hypothetical metropolitan landscape over the course of one day. We then quantify the amount of population shift bias affecting each of the space-time interaction tests for a number of different population movement scenarios.

2 Interaction Tests

The space-time interaction tests considered in this study are described in further detail below. The methods have been implemented by the authors in Python and are available in the open-source space-time analysis software, PySAL (Rey and Anselin, 2010). Note that in all cases we employ Euclidean distance metrics. Also, events are never considered adjacent to or neighbours of themselves.

To calculate the Knox (1964) test for space-time interaction, critical space and time distance thresholds (δ and τ, respectively) defining adjacency between events are specified by the user. The test statistic is then calculated as the count of event pairs that are adjacent in both time and space. Formally, the test statistic is specified in Equation 1, where n = number of events, a^s = adjacency in space, a^t = adjacency in time, d^s = distance in space, and d^t = distance in time.

$$X = \sum_i^n \sum_j^n a_{ij}^s a_{ij}^t \tag{1}$$

$$a_{ij}^s = \begin{cases} 1, & \text{if } d_{ij}^s < \delta \\ 0, & \text{otherwise} \end{cases}$$

$$a_{ij}^t = \begin{cases} 1, & \text{if } d_{ij}^t < \tau \\ 0, & \text{otherwise} \end{cases}$$

The Mantel test is a modification of the Knox test that considers the space and time distances between all pairs of events, and not just those within critical thresholds (Mantel, 1967). The test statistic is the sum of the products of the spatial and temporal distances between all event pairs in the dataset. The statistic is specified in Equation 2, where, again, d^s and d^t denote distance in space and time, respectively.

$$M = \sum_i^n \sum_j^n d_{ij}^s d_{ij}^t \tag{2}$$

In an effort to address shortcomings of the previous two methods, Jacquez (1996) developed a test using a similar form, based on nearest neighbour distances. The test locates the k nearest neighbours in both space and time for all events and then counts those common to both dimensions for individual events. Formally, the statistic, J_k is defined in Equation 3, where n = number of cases; a^s = adjacency in space; a^t = adjacency in time.

$$J_k = \sum_i^n \sum_j^n a_{ijk}^s a_{ijk}^t \tag{3}$$

$$a_{ijk}^s = \begin{cases} 1, & \text{if event } j \text{ is a } k \text{ nearest neighbour of event } i \text{ in space} \\ 0, & \text{otherwise} \end{cases}$$

$$a_{ijk}^t = \begin{cases} 1, & \text{if event } j \text{ is a } k \text{ nearest neighbour of event } i \text{ in time} \\ 0, & \text{otherwise} \end{cases}$$

To assess the significance of the results for each of these tests, a Monte Carlo approach is traditionally used where in each permutation the temporal coordinates are shuffled and the statistic is recalculated. This generates a distribution of potential values for the statistic (specific to the observed event pattern), which is then used to assess the pseudo-significance of the observed test statistic value. While this approach is appropriate in situations where the susceptible population is static across time, it is inappropriate when the distribution changes heterogeneously through time and space. Using this method in such a context introduces the population shift bias mentioned above.

3 Methods

This study measures the bias introduced by failing to account for shifts in the susceptible population for a hypothetical metropolitan area over the course of one day. To measure the bias, events (i.e. crimes, illnesses) are randomly generated within the population in each of four daily movement scenarios: high movement (where 98% of the individuals change spatial unit for some period during the day); moderate movement (59% change unit); low movement (35% change unit); no movement (all individuals remain within unit). The metropolitan area has a population of 640,000 divided equally among its 40 spatial units (see Figure 1). In each of the dynamic scenarios, the population in the spatial units varies heterogeneously over the course of the day. Some spatial units gain population (employment or shopping locations) at certain points of the day while others lose population (bedroom communities). Additionally, we consider the same scenarios with an additional influx of 400,000 individuals to the metro (visitors or commuters) from the periphery during the day.

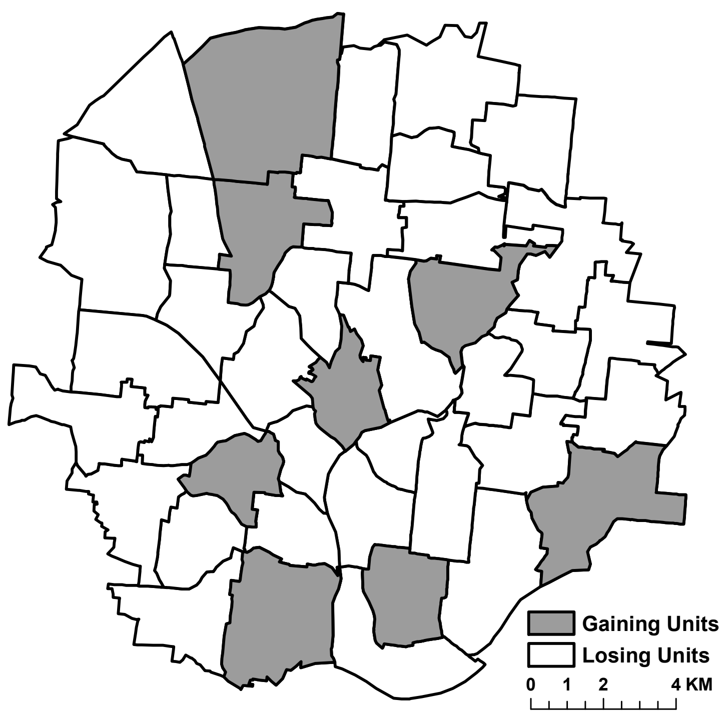

Figure 1: Simulation study area.

To estimate the population shift bias, we follow the methodology of Kulldorff and Hjalmars (1999). For our experiment, this means events are randomly assigned to spatial units in the metro at different hours in the day based on a probability proportional to the population of the spatial unit at each hour of the day. In this example, all individuals were assumed to be susceptible to the events. For each movement scenario, 1000 replications are run where 100 events are randomly simulated. The significance of the test statistics in each replication is assessed using the Monte Carlo approach described above. For each scenario and test combination, the proportion of significant replications (where $\alpha = 0.05$ and 0.01) is recorded. Because there is no population movement in the static scenario, there is no population shift bias; as a result, the proportion of significant replications for this scenario serves as our baseline. The difference between the proportion of significant replications observed for the dynamic population scenarios and that observed for the static population scenario measures the amount of population shift bias present in each of the tests, for each scenario. The parameters used in this study for the Knox and Jacquez tests are outlined in Table 1, no additional parameters were specified for the Mantel test.

4 Results

The results, shown in Table 1, illustrate the sizable impact population shift bias may have on these tests of space-time interaction, even for the short temporal extent considered. Generally speaking, the Knox test was most affected by the population shifts. As the critical distances used by the test increased, observed bias increased as well, in one case up to 95 times the α value. Although this extreme example is partly an artifact of our experimental design, which intended to promote any potential bias by concentrating mobile individuals in the gaining spatial units, the scenarios designed are not implausible and neither, therefore, are the estimates of the bias. Researchers employing this test, especially in an urban context, need to be aware of this susceptibility. Although still affected, the results for the Jacquez test displayed the least amount of bias, likely due to the relative nature of the nearest neighbour distance metric employed by the test. For all tests, any bias observed was increased by the addition of the influx population to the metro area.

The take-home message from this work is that population shift bias must be accounted for when employing tests of space-time interaction regardless of the test employed or the duration of the study. This can be accomplished by using an unbiased form of the test which takes population shift into account. A general template for such unbiased tests is described in Kulldorff and Hjalmars (1999). Future research should concentrate on specific implementations of this form.

5 Acknowledgements

This work was supported by the U.S. National Science Foundation through a Graduate Research Fellowship to Nicholas Malizia.

		Without Influx Population						With Influx Population					
		Low Movement		Moderate Movement		High Movement		Low Movement		Moderate Movement		High Movement	
Test	Parameters	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
Jacquez	$k = 1$	0.009	0.006	0.032	0.016	0.108	0.058	0.030	0.005	0.033	0.020	0.103	0.053
	$k = 2$	0.008	-0.004	0.035	0.011	0.139	0.073	0.020	0.009	0.044	0.028	0.167	0.116
	$k = 3$	0.007	0.001	0.050	0.022	0.257	0.140	0.045	0.017	0.074	0.037	0.241	0.149
	$k = 4$	0.003	-0.005	0.039	0.022	0.302	0.176	0.038	0.012	0.096	0.040	0.299	0.188
	$k = 5$	0.007	-0.003	0.049	0.019	0.390	0.228	0.065	0.019	0.121	0.047	0.376	0.244
Knox	$\delta = 0.5, \tau = 0.25$	0.030	0.021	0.024	0.002	0.062	0.045	0.022	0.017	0.028	0.017	0.082	0.100
	$\delta = 1.0, \tau = 0.25$	0.016	0.009	0.038	0.016	0.125	0.062	0.054	0.030	0.068	0.031	0.158	0.118
	$\delta = 2.0, \tau = 0.25$	0.022	0.023	0.035	0.033	0.170	0.084	0.062	0.031	0.092	0.049	0.216	0.153
	$\delta = 5.0, \tau = 0.25$	0.024	0.009	0.071	0.028	0.327	0.171	0.094	0.031	0.127	0.054	0.383	0.223
	$\delta = 0.5, \tau = 0.50$	0.011	0.012	0.042	0.023	0.160	0.079	0.045	0.032	0.068	0.042	0.200	0.147
	$\delta = 1.0, \tau = 0.50$	0.017	0.011	0.056	0.026	0.266	0.146	0.064	0.034	0.141	0.086	0.319	0.220
	$\delta = 2.0, \tau = 0.50$	0.031	0.016	0.088	0.048	0.411	0.261	0.107	0.052	0.237	0.116	0.473	0.348
	$\delta = 5.0, \tau = 0.50$	0.038	0.016	0.154	0.067	0.661	0.499	0.181	0.075	0.349	0.179	0.766	0.599
	$\delta = 0.5, \tau = 1.00$	0.025	0.004	0.079	0.040	0.332	0.207	0.097	0.043	0.180	0.086	0.416	0.303
	$\delta = 1.0, \tau = 1.00$	0.043	0.016	0.116	0.050	0.528	0.376	0.144	0.067	0.276	0.149	0.587	0.461
	$\delta = 2.0, \tau = 1.00$	0.052	0.019	0.134	0.076	0.709	0.575	0.224	0.104	0.410	0.230	0.777	0.686
	$\delta = 5.0, \tau = 1.00$	0.079	0.039	0.249	0.127	0.896	0.847	0.337	0.176	0.603	0.421	0.926	0.905
	$\delta = 0.5, \tau = 2.00$	0.017	0.009	0.036	0.022	0.433	0.250	0.115	0.050	0.185	0.089	0.501	0.364
	$\delta = 1.0, \tau = 2.00$	0.028	0.011	0.094	0.042	0.649	0.502	0.173	0.072	0.309	0.170	0.737	0.604
	$\delta = 2.0, \tau = 2.00$	0.044	0.016	0.150	0.081	0.790	0.704	0.261	0.123	0.451	0.277	0.852	0.821
	$\delta = 5.0, \tau = 2.00$	0.071	0.042	0.249	0.150	0.908	0.898	0.362	0.203	0.638	0.487	0.929	0.950
Mantel		0.075	0.026	0.139	0.056	0.592	0.407	0.181	0.086	0.216	0.093	0.569	0.384

Table 1: Population shift bias for all combinations of tests and population movement scenarios.

6 References

Aldstadt, J. (2007). An incremental Knox test for the determination of the serial interval between successive cases of an infectious disease. *Stochastic Environmental Research and Risk Assessment*, 21(5):487–500.

Grubesic, T. and Mack, E. (2008). Spatio-temporal interaction of urban crime. *Journal of Quantitative Criminology*, 24(3):285–306.

Jacquez, G. (1996). A k nearest neighbour test for space-time interaction. *Statistics in Medicine*, 15(18):1935–1949.

Knox, E. (1964). The detection of space-time interactions. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 13(1):25–30.

Knox, E. (2002). An epidemic pattern of murder. *Journal of Public Health*, 24(1):34–37.

Kulldorff, M. (1998). Statistical methods for spatial epidemiology: tests for randomness. In Gatrell, A. and Löytönen, M., editors, *GIS and Health*, pages 49–62. Taylor & Francis, London.

Kulldorff, M. and Hjalmars, U. (1999). The Knox method and other tests for space-time interaction. *Biometrics*, 55(2):544–552.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27(2):209–220.

Petridou, E., Revinthi, K., Alexander, F., Haidas, S., Koliouskas, D., Kosmidis, H., Piperopoulou, F., Tzortzatou, F., and Trichopoulos, D. (1996). Space-time clustering of childhood leukaemia in Greece: evidence supporting a viral aetiology. *British Journal of Cancer*, 73(10):1278–1283.

Rey, S. and Anselin, L. (2010). PySAL: A Python library of spatial analytical methods. In Fischer, M. M. and Getis, A., editors, *Handbook of Applied Spatial Analysis*, pages 175–193. Springer.

Rogerson, P. (2001). Monitoring point patterns for the development of space–time clusters. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 164(1):87–96.

Ward, M. and Carpenter, T. (2000). Analysis of time–space clustering in veterinary epidemiology. *Preventive Veterinary Medicine*, 43(4):225–237.