VGI for mapping change in bike ridership

D. Boss[1], T.A. Nelson*[2] and M. Winters[3]

[1]Unviersity of Victoria, Victoria, Canada
[2]Arizona State University, Arizona, USA
[3]Simon Frasier University, Vancouver, Canada
*Email: Trisalyn.nelson@asu.edu

Abstract

Cities are investing in cycling infrastructure to increase the travel by bicycle. Most evaluation of changes to infrastructure are aspatial, simply measuring the number of riders on a street segment. Though a lack of spatial data on cyclists has limited to map-based assessment of ridership, fitness Apps (i.e., Strava) are generating spatially and temporally continuous VGI data cycling. Given concerns about representiveness of samples and completeness of data, we need to evaluate if fitness App data are a suitable tool for monitoring city-wide change in ridership. Our goal is to evaluate if we can conduct a spatially explicit evaluation of change in ridership patterns using VGI from Strava. We conducted our study in Ottawa, Canada using data from May 2015 and 2016. We quantified change in normalized ridership and utilized network-constrained local measure of spatial autocorrelation to identify change in riding patterns. Several spatial weights definitions were evaluated. We contextualized locations of change in ridership pattern with data on infrastructure change. Through preliminary analyses, we detected clusters of increased and decreased ridership that are associated with new instillations of infrastructure and a closed tunnel. While we will continue to require program that monitor all riders, fitness Apps provide a promising source of spatially explicit data that may help urban planners and transportation researchers better understand how people are moving through a city.

Keywords: VGI, Crowdsource, Cycling, Spatial Pattern

1. Context

Cycling is a sustainable mode of transportation with numerous health, environmental and social benefits, yet ridership levels remain low in North America. To encourage increased ridership, cities are making significant investments in cycling infrastructure, a number of cities investing in cycling infrastructure networks (Pucher & Buehler, 2008). It is essential that cities monitor and report on the impact of infrastructure projects on ridership to be accountable to the public and to encourage political will for future investments in cycling infrastructure (Handy et al., 2014). Ridership is typically monitored aspatially, by counting the number of cyclists on a road segment before and after infrastructure changes. Aspatial assessments of ridership do not account for changes in patterns and ignore the potential for cyclist change routes in repose to upgrades in cycling infrastructure. While mapped based monitoring of ridership can better illuminate how patterns of cycling are changing in a city, spatially explicit ridership data has been difficult to obtain. Fitness App data are changing the availability of ridership data and providing a spatially explicit data on where people ride (Jestico et al., 2016). However, fitness App data, like other VGI, are a biased sample of users, as they tend to underrepresent women and young and old riders (Ferster et al. 2017). As well, the number of people using fitness Apps can change through time making the sample size difficult to ascertain.

Our goal is to evaluate if we can utilize bike ridership VGI to monitor change in patterns of cycling across a city in response to cycling infrastructure investment. The data used to represent ridership data generated by users of Strava.com, arguably the most popular commercial cycling App for Ottawa, Canada in May 2015 and 2016.

2.0 Study area and data

Ottawa is Canada's Capital Region and contains over 600 km of bicycle paths. The region has invested significant financial resources in bicycle and multi-use infrastructure over the past several years to promote active transportation. The VGI dataset is from Strava, a mobile health and fitness application and we are utilizing data through a partnership between Strava and the City of Ottawa. The data includes activity counts (bicycle trips) per segment of transportation infrastructure in the Ottawa region, aggregated for weekdays in May 2015 and May 2016. There were a total of 4.49 million activity counts from 52,123 bike trips across 71,205 network segments.

3.0 Methods

To map change in ridership we summed the total count of all activities across the study area for each time period and calculated the proportion of trips that occurred on each segment. We subtracted normalized ridership in 2015 from 2016 on a segment-by-segment basis and created a map of the absolute difference. We visualized the resulting map in an attempt to identify change in the spatial variation of bicycle trips. The normalized change in ridership was used as input for a network based local Moran's I_i to map clustering in ridership change (Yamada and Thill 2007). As all spatial methods are sensitive to definitions of what is nearby, we used three definitions of the spatial weights matrix or neighborhood: first order contiguity with equal weighting; second order contiguity with equal weighting of all neighbors; and second order contiguity with weight varying based on the degree of contiguity (Nelson and Boots 2008). Local Moran's I_i enables evaluation of statistically significant change in patterns of ridership and considers each segment of the network in the context of what surrounds it (Anselin 1995; Nelson and Boots 2008). The results of local Moran's I_i is a map of clusters or hot spots of increased ridership and decreased ridership, as well as outlier segments, where the segment showed an increase or decrease and surrounding segments had an opposite change in ridership. To contextualize how the approach can be used to monitor ridership, maps of change in spatial patterns of ridership were annotated with changes that occurred in the cycling network between May 2015 and May 2016.

4.0 Results

The map of normalized change in ridership enabled the identification of segments with an increase or decrease in normalized change in ridership, but it was not possible to identify where statistically significant change had occurred or if clustering was present in the variation in ridership (Figure 1a). Corridors and regions of statistically significant change in normalized ridership by visualizing maps of the network based Local Moran's I*i* applied to change in ridership between years (Figure 1b). Larger clusters were observed when using second order contiguity to define a neighborhood as compared to first order contiguity (Figure 1c). Little difference was observed between weighted and unweighted second order contiguity (Figure 1d). Upon annotating the maps of Local Moran's I_i we observed changes in the distribution of ridership following both the installation of new cycling infrastructure and the closure of existing infrastructure (Figure 2).

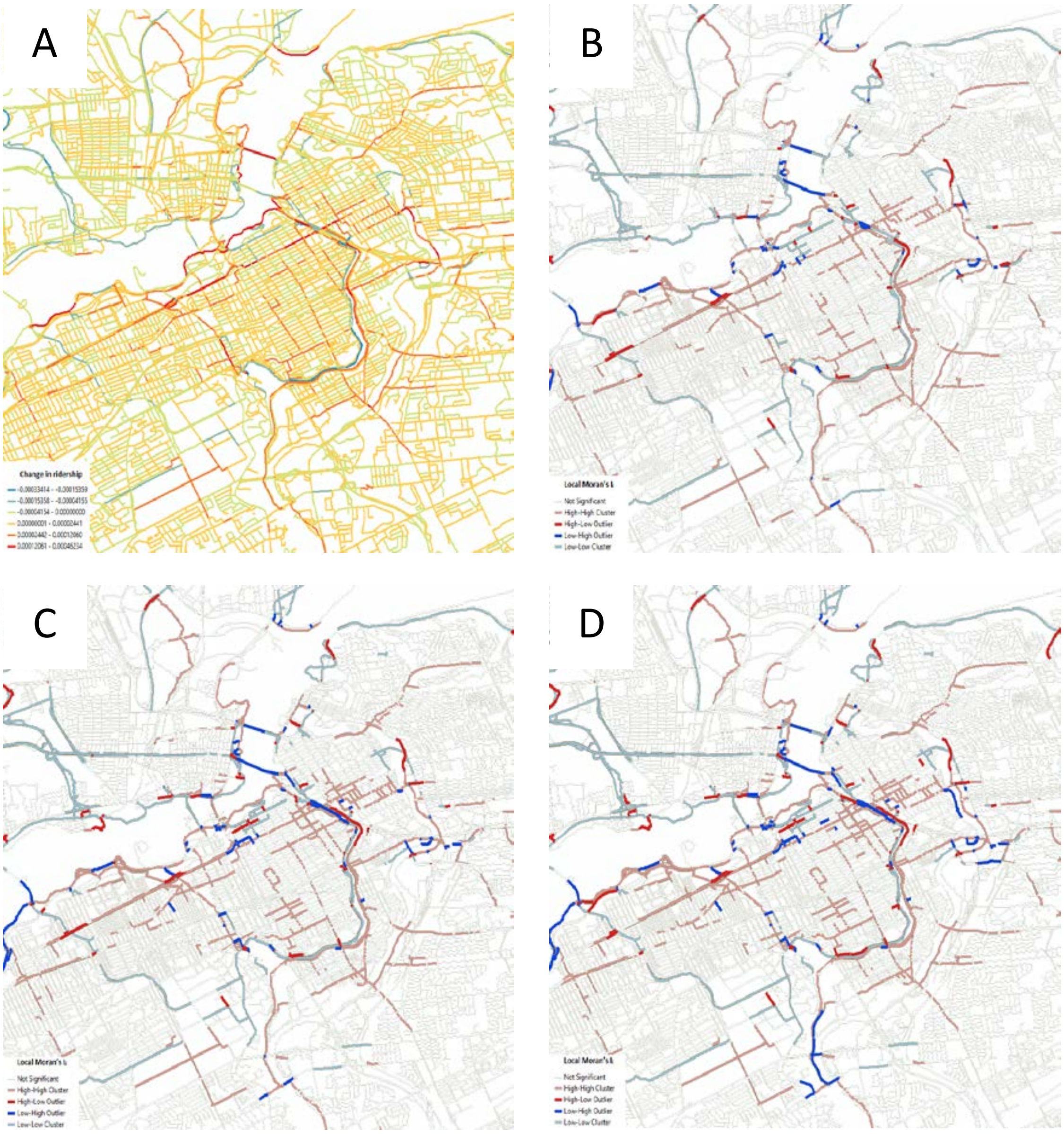

Figure 1: A) Absolute change in normalized ridership, May 2015 to May 2016. B) Local Moran's I_i using the difference of normalized ridership between May 2016 and May 2015. Spatial neighborhood defined by first order contiguity and equal weighting of all neighbors. High-high is a cluster of increased ridership; low-low is a cluster of decreased ridership; high-low is an outlier or a road segment where increased ridership is surrounded by a decrease; low-high is an outlier or a road segment where decreased ridership is surrounded by an increase in ridership. C) Local Moran's I_i using the difference of normalized ridership between May 2016 and May 2015. Spatial neighborhood defined by second order contiguity with varying weights depending on the level of contiguity (first order neighbor = 1, second order neighbor = 0.5). D) Local Moran's I_i using the difference of normalized ridership between May 2016 and May 2015. Spatial neighborhood defined by second order contiguity with equal weighting of all neighbors.

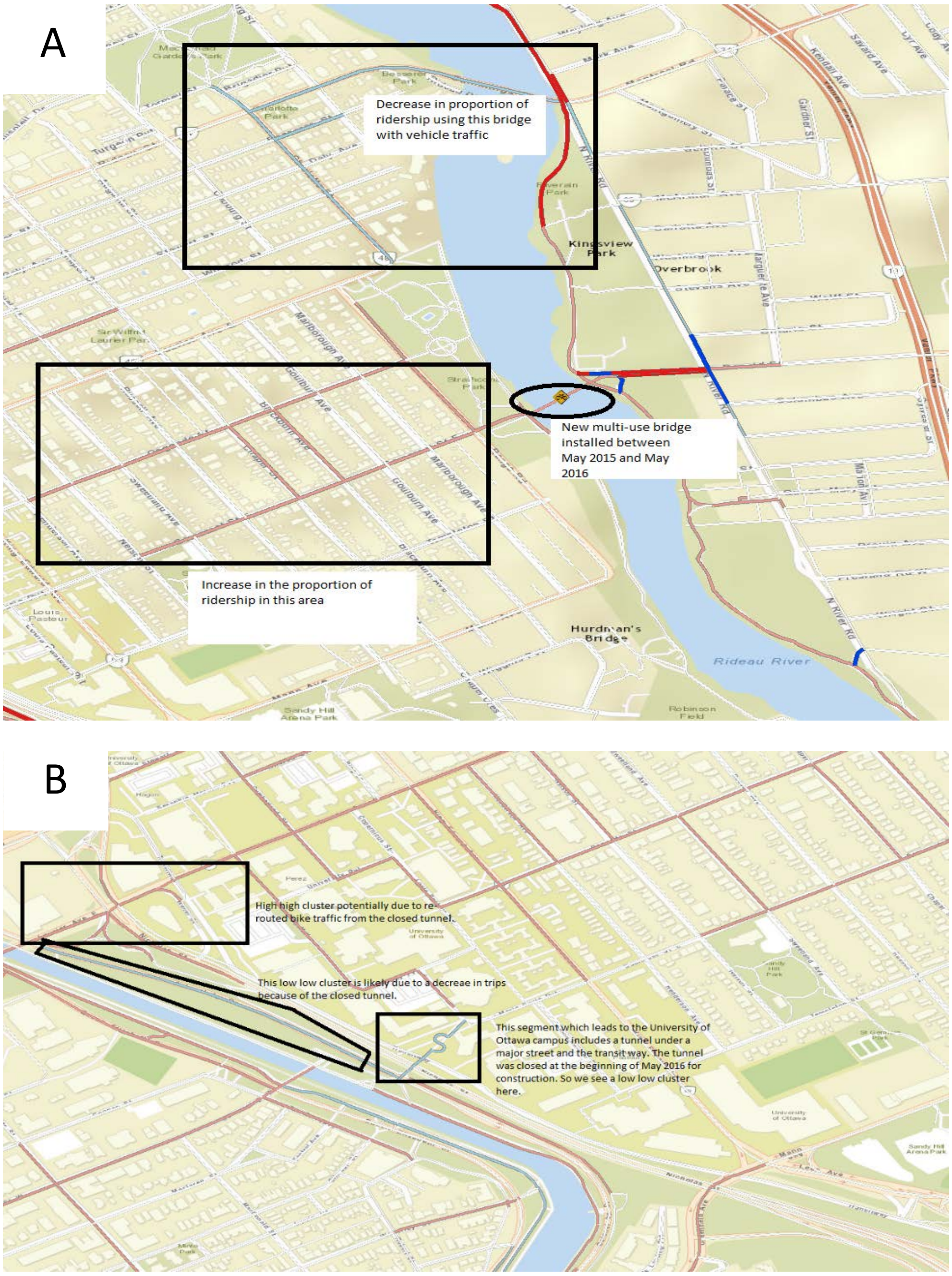

Figure 2. Examples of preliminary annotation of spatial patterns of change in ridership for Ottawa. A) Clusters of increased in ridership and a decreased ridership are associated with a new multi-use bridge path. B) Changes in patterns of ridership associated with a tunnel closure.

5.0 Discussion and Significance

Monitoring and evaluation of the impacts of investment in cycling infrastructure across a city have been limited by a lack of spatially explicit ridership data. However, the results illustrated here demonstrate the importance of consider patterns of change in cycling when infrastructure changes in a city. In both the example of a closed tunnel and new multi-use trail bridge (Figure 2) infrastructure change in one location affects the flow and amount of bicycle traffic in multiple locations. As cities invest more heavily in cycling infrastructure, the need to evaluate how ridership changes is paramount. The next step in our investigation is to quantify if the patterns of change in Strava riders represent patterns for all the cyclists. Anecdotally, in Ottawa the representation of cyclists may be strengthen by a campaign undertaken by the City of Ottawa to encourage cyclists of all ages and abilities to log rides prior to purchasing the Strava data.

While the methods used here are well known to geographers, the application to fitness app data is an important one. A focus on detecting statistically significant change in *spatial pattern* of ridership is paramount to the successful use of Strava data for transportation planning and research. Using Local Moran's I_i we can determine when the change in ridership patterns are unexpected based on chance, which is a threshold that can be defended. The unique aspect of fitness App data is that we sample movement across a city. With millions of users, Strava makes a strong case that fitness Apps are a growing data source and demonstrating how to effectively convert data into useful information will help fill gaps in cycling data. Planning and research will continue to require official and comprehensive count programs to monitor total number of cyclists, but the logistics official counts inherently restrict these to a few locations within the city. Jestico et al (2016) found that there is a strong correlation between Strava and all riders in the core of a mid-sized North American city. Tools to expand use of data to evaluate how patterns of movement are shifting is an additional benefit.

Methodologically, the spatial weights used to implement local methods of spatial autocorrelation is an important consideration (Nelson and Boots 2008). Despite the spatial weights definition, as applied to a network, we found the pattern of change detected similar, and as expected second order contiguity found bigger clusters. In future, setting up weights matrices to consider flow of traffic could be beneficial for monitoring change in traffic pattern.

Strava and other fitness App data represent an important source of VGI that is filling a massive gap in mobility data, particular associated with active transportation. As a geocomputation community we need to lead the demonstration of how to use big spatial data sets on mobility appropriately to better understand mobility in cities. Planners, transportation engineers, and researchers can use the methods outlined here to monitor changes in ridership patterns following investment in infrastructure or other changes to the transportation network.

6. Acknowledgements

Many thanks to the City of Ottawa and Strava.com for providing access to VGI ridership data.

7. References

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

Ferster, C.J., Nelson, T., Winters, M. and Laberee, K., 2017. *Geographic Age and Gender Representation in Volunteered Cycling Safety Data: A Case Study of Bikemaps. org* (No. 17-01439).

Handy, S., Van Wee, B. and Kroesen, M., 2014. Promoting cycling for transport: research needs and challenges. *Transport reviews*, *34*(1), pp.4-24.

Jestico, B., Nelson, T. and Winters, M., 2016. Mapping ridership using crowdsourced cycling data. *Journal of transport geography*, *52*, pp.90-97.

Nelson, T.A. and Boots, B., 2008. Detecting spatial hot spots in landscape ecology. *Ecography*, *31*(5), pp.556-566.

Pucher, J. and Buehler, R., 2008. Making cycling irresistible: lessons from the Netherlands, Denmark and Germany. *Transport reviews*, *28*(4), pp.495-528.

Yamada, I. and Thill, J.C., 2007. Local Indicators of Network-Constrained Clusters in Spatial Point Patterns. *Geographical Analysis*, *39*(3), pp.268-292.