Accessibility and Centrality Based Estimation of Urban Pedestrian Activity

Brendan Murphy
Graduate Research Assistant, Accessibility Observatory
University of Minnesota, Department of Civil, Environmental, and Geo- Engineering 500 Pillsbury Drive SE, Minneapolis, MN 55455 USA
murph677@umn.edu

David Levinson
RP Braun-CTS Chair of Transportation Engineering
Director of Network, Economics, and Urban Systems Research Group
University of Minnesota, Department of Civil, Environmental, and Geo- Engineering
500 Pillsbury Drive SE, Minneapolis, MN 55455 USA
dlevinson@umn.edu `http://nexus.umn.edu`

Andrew Owen
Director of Accessibility Observatory
University of Minnesota, Department of Civil, Environmental, and Geo- Engineering
500 Pillsbury Drive SE, Minneapolis, MN 55455 USA
aowen@umn.edu `http://nexus.umn.edu`

4072 words + 11 figures + 2 tables
August 1, 2015

ABSTRACT

Non-motorized transportation, particularly including walking and bicycling, are increasingly becoming important modes in modern cities, for reasons including individual and societal wellness, avoiding negative environmental impacts of other modes, and resource availability. Institutions governing development and management of urban areas are increasingly keen to include walking and bicycling in urban planning and engineering; however, proper placement of improvements and treatments depends on the availability of good usage data. This study attempts to predict pedestrian activity at 1123 intersections in the Midwestern, US city of Minneapolis, Minnesota, using scalable and transferable predictive variables such as economic accessibility by sector, betweenness network centrality, and automobile traffic levels. Accessibility to jobs by walking and transit, automobile traffic, and accessibility to certain economic job categories (Education, Finance) were found to be significant predictors of increased pedestrian traffic, while accessibility to other economic job categories (Management, Utilities) were found to be significant predictors of decreased pedestrian traffic. Betweenness centrality was not found to be a significant predictor of pedestrian traffic, however the specific calculation methodology can be further tailored to reflect real-world pedestrian use-cases in urban areas. Accessibility-based analysis may provide city planners and engineers with an additional tool to predict pedestrian and bicycle traffic where counts may be difficult to obtain, or otherwise unavailable.

INTRODUCTION

Walking and bicycling are increasingly becoming important transportation modes in modern cities, and for a multitude of reasons, including individual and societal wellness, environmental externalities associated with motorized modes, and resource availability. Planning for biking and walking, and creating societal programs to increase their levels, has been cited as a targeted health need in urban planning going forward (Lumsdon and Mitchell (*1*), Raford and Ragland (*2*), Brownstone (*3*)). Resource limitations, particularly in high-population and developing third-world countries, impose constraints on the maximum level of personal motorized travel allowed, and as a result there is a greater need for viable alternatives. In addressing the viability and availability of alternative modes, high-resolution spatial data on non-motorized transportation behavior patterns is needed.

Rates of walking and bicycling to work in the United States hover around 2.8% and 0.6%, respectively, with public transit use barely higher at 5% nationally Bureau (*6*). Proper placement of pedestrian treatments and improvements has implications to both safety (Schneider et al. (*7*)) and accessibility and mode choice (Iacono et al. (*8*)), but proper information regarding estimated non-motorized traffic levels is needed to locate areas in need of improvement. In determining salient locations for non-motorized improvements, it is important to have accurate records of both existent and potential travel demand (e.g. current levels of walking in a neighborhood, as well as good models of increased demand due to potential treatments); however good quality, high-granularity datasets for non-motorized travel can be difficult to obtain, especially standardized for national spatial inventories (McDaniel et al. (*9*)). For this reason, practitioners and researchers must frequently rely on estimation models for non-motorized traffic, and various methods can suffer from issues of data quality, granularity, and the presence of location-specific variables (Lowry (*10*)).

Many of the issues with the collection of standardized non-motorized transportation data have to do with the factors that influence pedestrian and bicycle behavior. A model of active transport risk assessment is uninformative if the pedestrian and vehicular flows do not accurately represent corresponding levels *in situ*, and many cities do not have dense data sets of active transport flow levels, instead favoring counts of vehicle traffic. As such, active transport flow levels must be extrapolated from sparse data sets using comprehensive methodologies. Land use data are well-documented by the U.S. Census Bureau to the Census Block level of resolution, and general socioeconomic characteristics are maintained as well, and can have significant influence (Schneider et al. (*11*)). However, more specific socioeconomic characteristics are salient in non-motorized travel beyond just adjusted income levels, as well as weather variables (Miranda-Moreno et al. (*12*)) and latent, subjective variables such as visibility and perceptions of lighting, which can be more difficult to obtain at high spatial resolution (Kamargianni (*13*)), and can complicate inter-city comparisons. For these reasons, as well as the overall lack in non-motorized travel counts for many communities, methods of estimating pedestrian and bicycle behavior that do not rely heavily on high-resolution count data area applied in this study.

Aggregate travel behavior studies typically involve analysis at the level of Transit Analysis Zones (TAZs), which are too coarse to allow robust analysis of non-motorized travel (Schneider et al. (*11*)), (Iacono et al. (*8*)); Regional Travel Surveys consider many trip purposes, but are similarly coarse, and typically have too small of sample sizes to allow for robust city-to-city comparison. Census block-level information regarding economic accessibility (access to jobs) via both strictly walking, and via the net accessibility benefit of public transportation, will first be used to explain observed pedestrian traffic at a subset of intersections in the city of Minneapolis, Minnesota. Road network betweeenness centrality will also be used as an explanatory variable, as a proxy of the underlying network structure. A framework for comprehensive pedestrian risk assessment modeling, using pedestrian volume, vehicle volume, and an environmental factor (crosswalk length) on a university campus is provided by Schneider et al. (*7*). The motivation for constructing models of pedestrian and vehicular traffic is in supplementing the sparse data currently available, and deriving a reusable framework to provide a more complete picture of pedestrian activity throughout the city at the level of individual intersections, based on non-location-specific available data.

METHODOLOGY

Data

This section briefly describes the data sources used in the pedestrian estimation models, and the data preparation process.

- **Data Sources**
  1. U.S. Census TIGER 2010 datasets: blocks, core-based statistical area (CBSA) for Minneapolis-St. Paul
  2. U.S. Census Longitudinal Employer-Household Dynamics (LEHD) 2011 Origin-Destination Employment Statistics (LODES)
  3. OpenStreetMap (OSM) North America extract, retrieved April 2014
  4. Turning movement counts (TMC) 2000-2013, City of Minneapolis
  5. Average Annual Daily Traffic (AADT) measurements 2000-2013, City of Minneapolis
  6. GTFS data from Metro Transit
- **Data Preparation**
  1. Construct pedestrian travel network graph for Minneapolis
  2. Geocode pedestrian Turning Movement Count (TMC) and Average Annual Daily Traffic (AADT) data to spatial locations
- **Accessibility & Centrality Calculation**
  1. For each Census block in Minneapolis, calculate travel time to all other blocks within a 5km radius for a single departure time
  2. Calculate cumulative opportunity accessibility to jobs for each census block, using thresholds of 5, 10, . . . , 30
  3. Calculate net transit accessibility benefit using a threshold of 30 minutes
  4. Calculate betweenness centrality for the Minneapolis OSM road network
- **Model estimation**

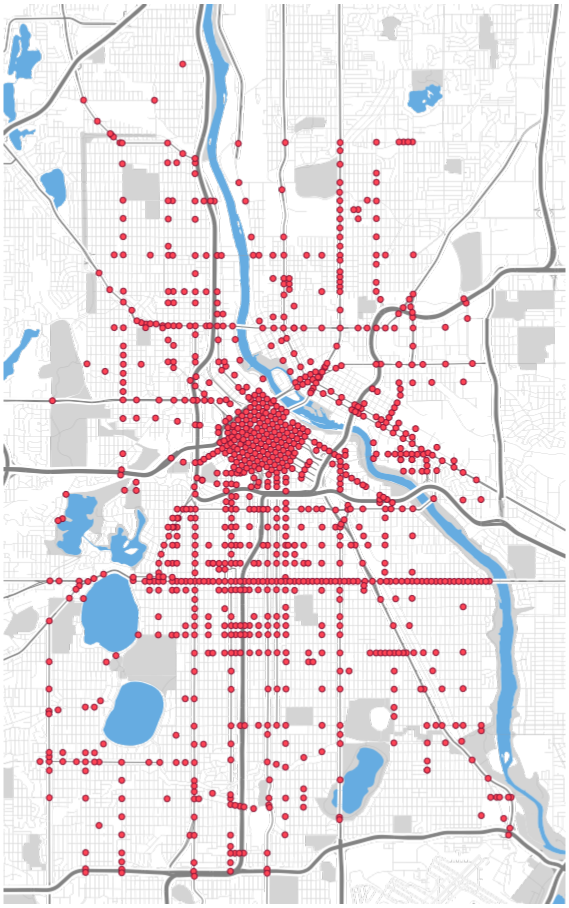

FIGURE 1 : Locations of intersections in Minneapolis with raw pedestrian count data.

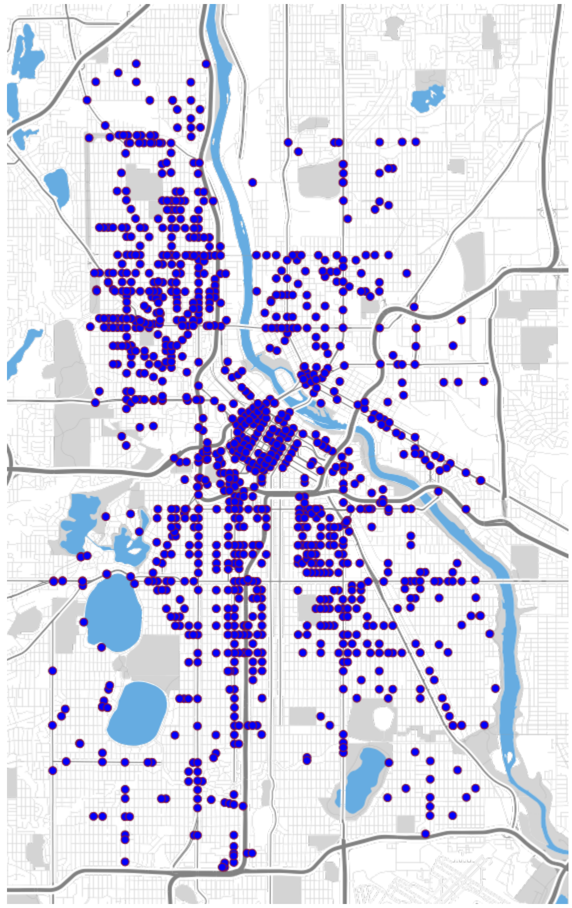

FIGURE 2 : Locations of sampled intersections in Minneapolis, used in estimated pedestrian activity analysis.

1. Construct linear regression of pedestrian behavior on walking accessibility, net transit accessibility, network centrality, and accessibility to job opportunities by sector
2. Assess and validate model on sample of other intersections in Minneapolis

Intersection locations were determined from OSM road centerline data for the Minneapolis-St. Paul CBSA (Core-Based Statistical Area). The subset of intersections for which count data were available is displayed in Figure 1; these intersections were used to construct the predictive models. Accessibility calculations were performed using OpenTripPlanner (OTP) open-source routing software; GIS work performed in QGIS and PostGIS; network centrality measures computed in ArcMap GIS with the Urban Network Analysis Tools toolbox; statistical work done in SQL, Python, and R. Figure 2 displays the locations of intersections in Minneapolis used to estimate pedestrian activity and validate the model.

Accessibility

The first type of explanatory variable used in the model of Minneapolis pedestrian count data is cumulative opportunity accessibility. Using OTP, walking travel times along the network are calculated from each Census block centroid in Minneapolis, to each other block centroid within the travel-time thresholds of 5, 10, ..., 30 minutes. Job opportunities are summed from each block centroid reachable within a given time threshold, yielding an X-minute accessibility measure. Job opportunities are broken down by economic sector, as defined by the North American Industry Classification System. There are two accessibility calculations used in this study:

1. Accessibility to jobs from Census block centroids by walking
2. Accessibility to jobs from Census block centroids by transit & walking

Pedestrian counts are often taken at intersections in either gross counts, or divided by turning movement type. This study uses Turning Movement Count (TMC) data from approximately 750 intersections in Minneapolis; intersection counts were calculated by adding the various TMC types for each intersection in the analysis group, to yield a gross figure of pedestrian activity within an intersection. Two-hour counts for pedestrian activity were used for morning peak (7-9AM), midday (11am-1pm), and evening peak (4-6PM). Accessibility calculations were performed using the following formulation of a gravity-based model:

$$A_i = \sum_j O_j f\left(C_{ij}\right) \tag{1}$$

$$A_i = \text{accessibility for location } i \tag{2}$$
$$O_j = \text{number of opportunities at location } j \tag{3}$$
$$C_{ij} = \text{time cost of travel from } i \text{ to } j \tag{4}$$
$$f\left(C_{ij}\right) = \text{weighting function} \tag{5}$$
$$\tag{6}$$

The choice of weighting function has a large impact on the resulting Accessibility calculations; however, one of the simplest interpretations of cumulative opportunities is an integer count, using the following weighting function:

$$f\left(C_{ij}\right) = \begin{cases} 1 & \text{if } C_{ij} \leq t \\ 0 & \text{if } C_{ij} > t \end{cases} \tag{7}$$

$$t = \text{travel time threshold}$$

This intuitively makes sense when applied to opportunities such as jobs, number of restaurants, transit route departures, and other discrete integer variables in the surrounding environment. We predict that origins exhibiting higher accessibility values would see greater pedestrian activity throughout the day. Accessibility for both walking, and walking + transit modes, are used in the estimation models; subtracting walking accessibility from the multimodal walking + transit accessibility yields the net transit benefit, and including walking and net transit separately in the regression models allows for explicit evaluation of how important transit benefits are to influencing pedestrian activity. Multiple regression in *R* statistical package was then performed to determine the explanatory power of the accessibility measures in predicting pedestrian and vehicular traffic in the AM, midday, PM peaks, as well as for a 6-hour summed count. These additional tables are omitted here. It was expected that origins exhibiting higher walking-accessibility values, and higher centrality values, would see greater pedestrian activity throughout the day.

Centrality

In an attempt to reflect pedestrian activity on the underlying topology of the transportation network, a centrality measure was computed in ArcGIS with the Urban Network Analysis Toolbox, and added to the regression models. Various types of network measures of centrality have been proposed in their applicability to estimation of non-motorized activity levels (McDaniel et al. (*9*), Anciães (*14*), Do et al. (*15*)), and safety and collision rates (Zhang et al. (*16*), Dai et al. (*17*)). One of the most common measures of centrality is "betweenness" centrality, or how "between" other nodes or links a given node or link is. When considering route choice and estimating modal traffic flows, link betweenness centrality is often considered, and consists of the proportion of shortest paths between all node pairs that pass through a link or node (McCahil and Garrick (*18*)). Relatedly, stress centrality consists of counting the number of times each link in a given network is utilized among the set of shortest paths between all node pairs, and is given by:

$$C_s(k) = \sum_{i,j \in V} \sigma_{ij}(k) \tag{8}$$

where σ_{ij} is either 1 if link k is used in shortest path σ_{ij}, and 0 otherwise. This form of stress centrality has been used to spatially assess transportation systems (Derrible (*19*)). In order to adapt stress centrality to the specific characteristics of non-motorized travel, (McDaniel et al. (*9*)) added the following modifications to the link betweenness schematics for the bicycle mode:

1. Restrict shortest paths to preferred bicycle routes
2. Restrict origin-destination (O/D) pairs to only locations reachable by bicycle
3. Modify O/D frequency with trip multipliers

However, for the walking mode, it is not reasonable to include the entire set of road network intersections as possible destinations for a given intersection-origin, due to the lower speed of the walking mode - an assumed 5 km/h. Thus, for the centrality calculations for the walking mode, an on-network radius of 5 kilometers, to represent an hour of walking at average speed, was implemented to increase the saliency and relevance of centrality to actual walking behavior. Additionally, similar modifications to the above for bicycle modes may be implemented for walking, in particular modifying O/D frequency to reflect that a certain subset of nodal origins and destinations exhibit much higher activity levels than others; for simplicity, such modifications were not attempted in this study.

To reflect typical work trips, McDaniel et al. (*9*) chose O/D pairs such that origins were strictly residential parcels, and non-residential parcels were destinations in the morning, and the order was reversed in the evening. However, the authors speculated that allowing for non-residential destinations in the evening to reflect more complex after-work tours could increase model explanatory power (McDaniel et al. (*9*)). Additionally, O/D pairs were limited by a network distance threshold of 5 miles, per the *National Household Travel Survey* (Federal Highway Administration (*21*)). O/D multipliers specified relative magnitude of trip generation, since parcels are heterogeneous in their trip generation capacity; these included density of dwelling units within residential parcels, and square footage density for all other parcels.

These modifications constitute potentially salient areas for further investigation in our model of pedestrian traffic. O/D pairs can be tailored to favor walking trips from residential parcels

TABLE 1 : Dataset Summary Statistics

Description	Value
Intersections with evening ped counts	741
Intersections included in estimation modeling	1123
Intersection-μ total ped activity per day	633.66, $\sigma = 2023.20$
Intersection-μ morning ped activity per day	194.70, $\sigma = 570.34$
Intersection-μ midday ped activity per day	270.74, $\sigma = 994.79$
Intersection-μ evening ped activity per day	264.52, $\sigma = 733.49$

Note: Summary statistics for datasets used in pedestrian activity analysis: pedestrian turning movements between 2000 and 2013 for the City of Minneapolis.

to commercial destinations, as well as limited to reasonable walking distances attained within a 30-minute threshold (2.5 km). Rote stress centrality is first used to evaluate preliminary explanatory power, and feasibility of applying centrality metrics to this model.

Pedestrian Activity Estimation

Multiple regression over the explanatory variables was performed in R for the walking mode. Different time-thresholds of accessibility were compared for explanatory power of pedestrian activity, of which the strongest threshold was chosen for a final parsimonious model to estimate pedestrian traffic throughout the sampled intersections. Iterative stepwise regression was performed using the economic sector accessibility variables, in an attempt to account for the possible differential walking trip generation levels of different job sectors. The parsimonious model is then applied to a broader sample of intersections within Minneapolis, and the estimated pedestrian levels are compared to actual counts for validation, and specific spatial areas of underestimation and overestimation are discussed.

RESULTS

Full tabulation of all bivariate regression models, to determine which time thresholds and peak-hour periods to use for greatest explanatory power in modeling pedestrian traffic levels, are omitted for brevity. It was found that the 15-minute threshold of total accessibility, combined with the PM-peak period pedestrian counts and other variables, yielded the best explanatory power for walking accessibility. A parsimonious model for walking activity, in terms of the strongest explanatory variables, is reported in Table 2. Net transit accessibility benefit was included as an explanatory variable in the pedestrian activity estimation model, to account for the effect of transit in urban cores of increasing pedestrian activity by attracting additional users beyond pure foot traffic. Table 1 lists summary statistics for the datasets used in the following analysis: automobile-pedestrian crashes between 2000 and 2013, and pedestrian turning movement counts between 2000 and 2013 for Minneapolis.

First, the pedestrian counts were modeled in terms of only walking accessibility, for different thresholds and times of day. From this, the strongest explanatory power was determined for PM peak period counts, at a 15-minute accessibility threshold. Pedestrian counts were then modeled in terms of transit & walking accessibility (bimodal accessibility), for different times of day. A 30-minute transit threshold was used, in accordance with the reported data available in the Access Across America: Transit 2014 report (Owen and Levinson (*22*)). Net transit acces-

TABLE 2 : Parsimonious Model Regression Results: With & Without AADT

	Dependent variable:	
	Average PM pedestrians	
	(1)	(2)
Walking accessibility (15-minute)	0.410**	0.649***
	(0.173)	(0.112)
Net transit accessibility (30-minute)	0.320***	0.129**
	(0.093)	(0.053)
Betweenness	0.029	0.487***
	(0.371)	(0.186)
AADT	1.312*	
	(0.679)	
Management jobs 5min	−0.114***	−0.109***
	(0.033)	(0.017)
Education jobs 5min	0.922***	0.700***
	(0.086)	(0.058)
Finance jobs 10min	0.071***	0.054***
	(0.009)	(0.006)
Utilities jobs 15min	−0.968***	−0.729***
	(0.104)	(0.071)
Constant	−15.208	−1.698
	(9.874)	(4.795)
Observations	486	1,016
R^2	0.287	0.226
Adjusted R^2	0.275	0.221
Residual Std. Error	83.830 (df = 477)	72.773 (df = 1008)
F Statistic	23.970*** (df = 8; 477)	42.139*** (df = 7; 1008)
Note:	*p<0.1; **p<0.05; ***p<0.01; (standard error)	

sibility, a measure which looks at the contribution to accessibility from transit service, was also investigated as a potential explanatory variable for walking activity. A threshold of 30-minutes was again used. Betweenness stress centrality was included to relate walking activity to the underlying network structure. Accessibility and betweenness centrality are mapped in Figure 3 and Figure 4, respectively.

Regression results for the two parsimonious models for walking activity, with and without AADT included, are in Table 2. Accessibility by walking, net transit benefit to accessibility, AADT, and accessibility to Finance and Education jobs were all found to be significant predictors of increased pedestrian activity. Accessibility to Management and Utilities jobs were found to be significant predictors of decreased pedestrian activity, relative to other variables. Betweenness centrality was not found to be a significant predictor of pedestrian traffic, but showed weakly positive correlation. A series of maps shows additional views of the data used in the modeling process; Figure 3 shows accessibility to jobs within 30 minutes by walking in Minneapolis, and Figure 4 shows the betweenness centrality of all intersections in Minneapolis calculated with a 5km radius. Accessibility by walking, given the walking mode's uniform nature, shows where economic activity is most concentrated in the region. Centrality gives a sense of the most important nodes in the street network of Minneapolis - that is, the nodes that would affect the highest number of shortest paths, were they to be rendered impassible. Both of these calculations showed positive correlations with pedestrian activity, as shown in Table 2. Figure 5 shows the raw levels of daily pedestrian activity, aggregated from manual pedestrian counts between 2000 and 2013, while Figure 6 shows the estimated levels of evening peak pedestrian activity in Minneapolis, calculated using the model definitions outlined in Table 2. To validate the estimated model, the difference between actual and estimated pedestrian activity is mapped in Figure 7. Additionally, spatial distributions of jobs in categories of Utilities, Finance, Management, and Education are shown in Figure 8, Figure 9, Figure 10, and Figure 11, respectively.

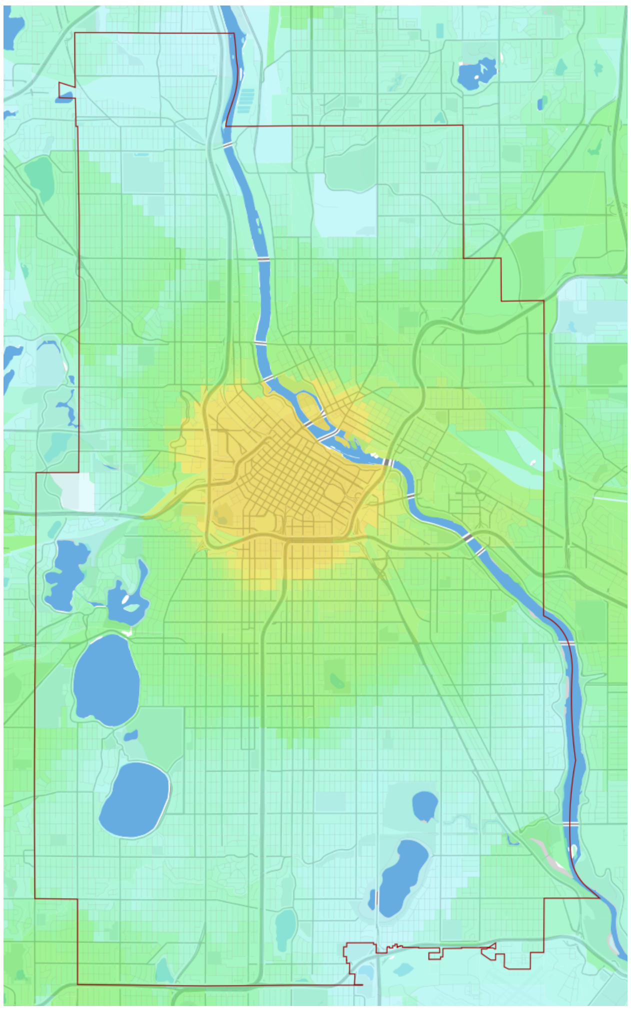

FIGURE 3 : Accessibility to jobs within 30 minutes by walking in Minneapolis.

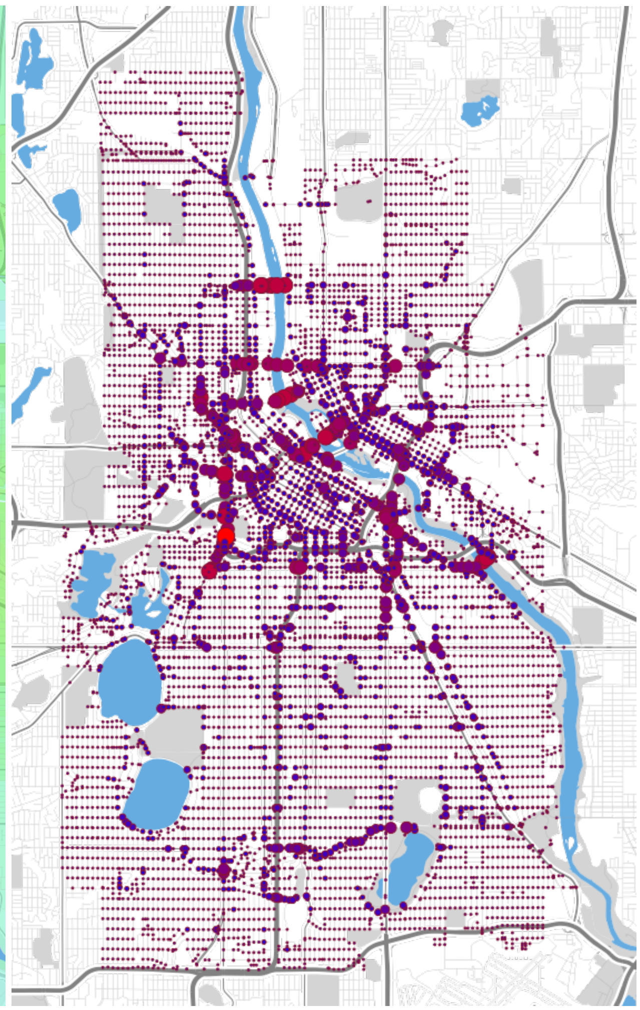

FIGURE 4 : Betweenness centrality of all intersections in Minneapolis; radius of 5km.

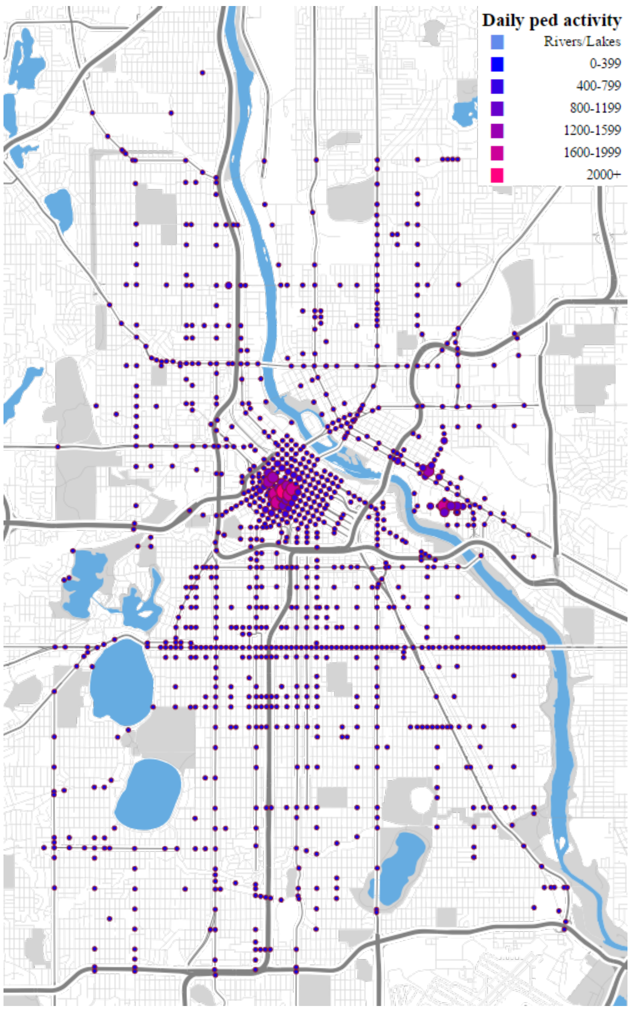

FIGURE 5 : Raw levels of daily pedestrian activity in Minneapolis, 2000-2013.

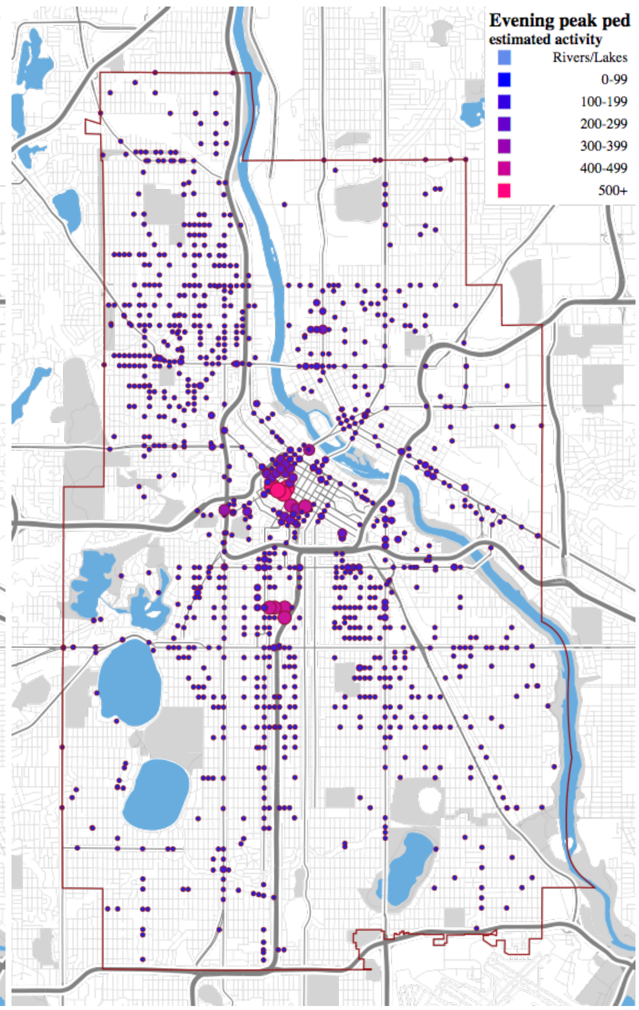

FIGURE 6 : Estimated levels of evening peak pedestrian activity in Minneapolis.

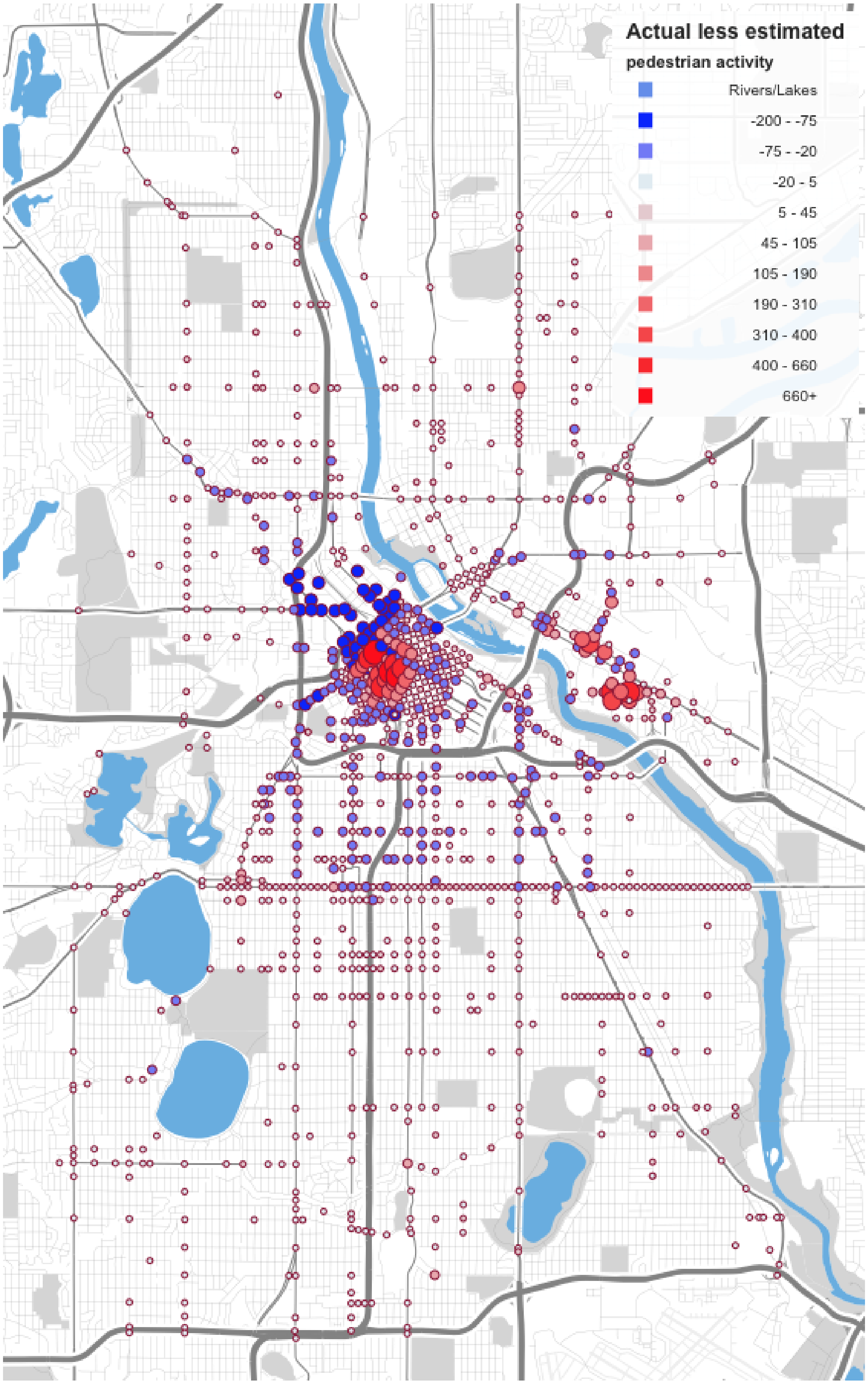

FIGURE 7 : Actual minus estimated pedestrian activity, PM peak period. Reds are areas of underestimation; blues are areas of overestimation.

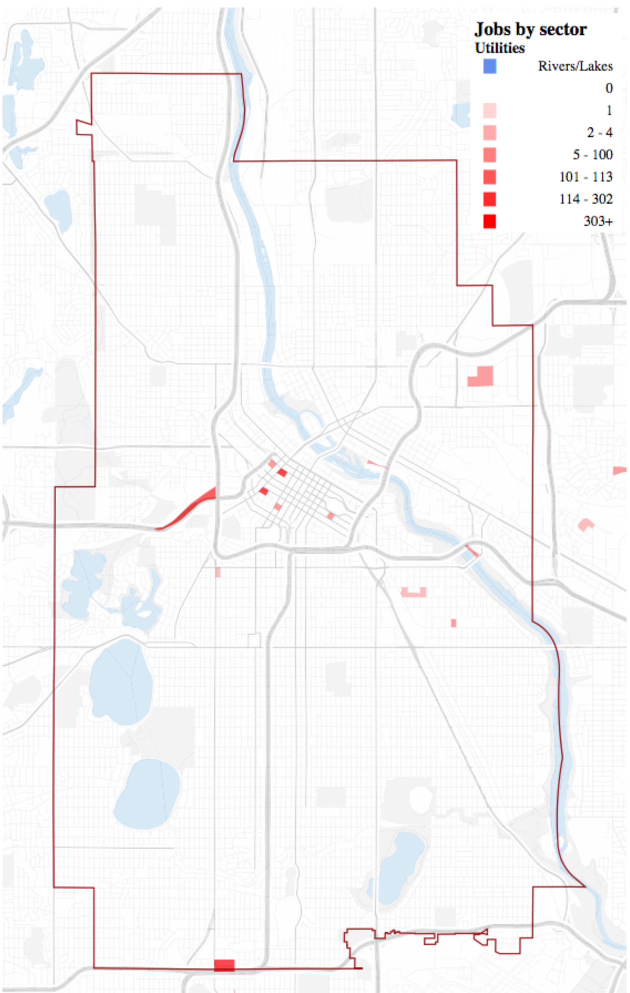

FIGURE 8 : Spatial distribution of utility jobs in Minneapolis, based on LEHD data.

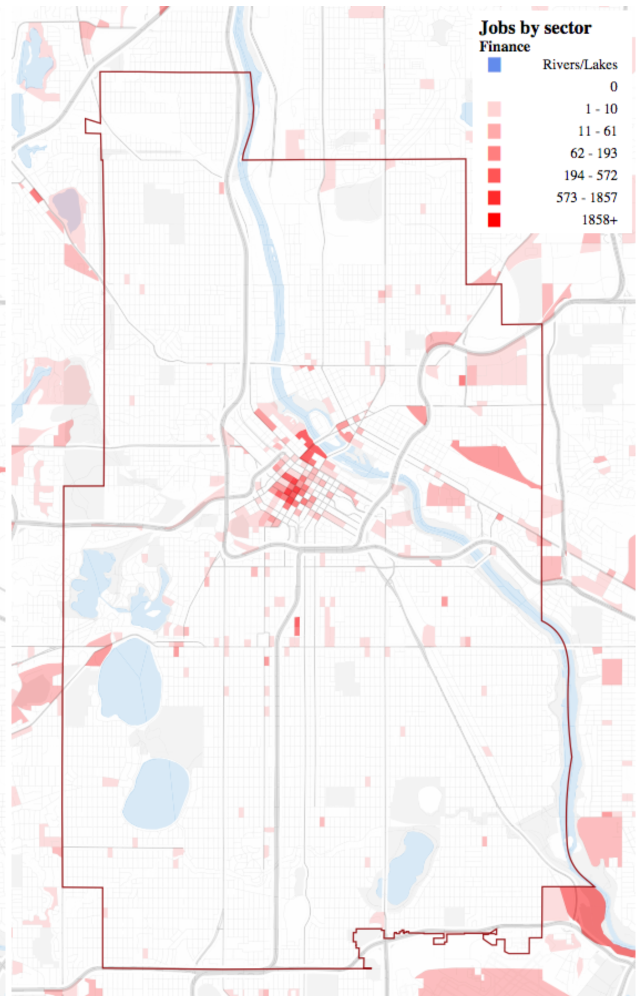

FIGURE 9 : Spatial distribution of finance jobs in Minneapolis, based on LEHD data.

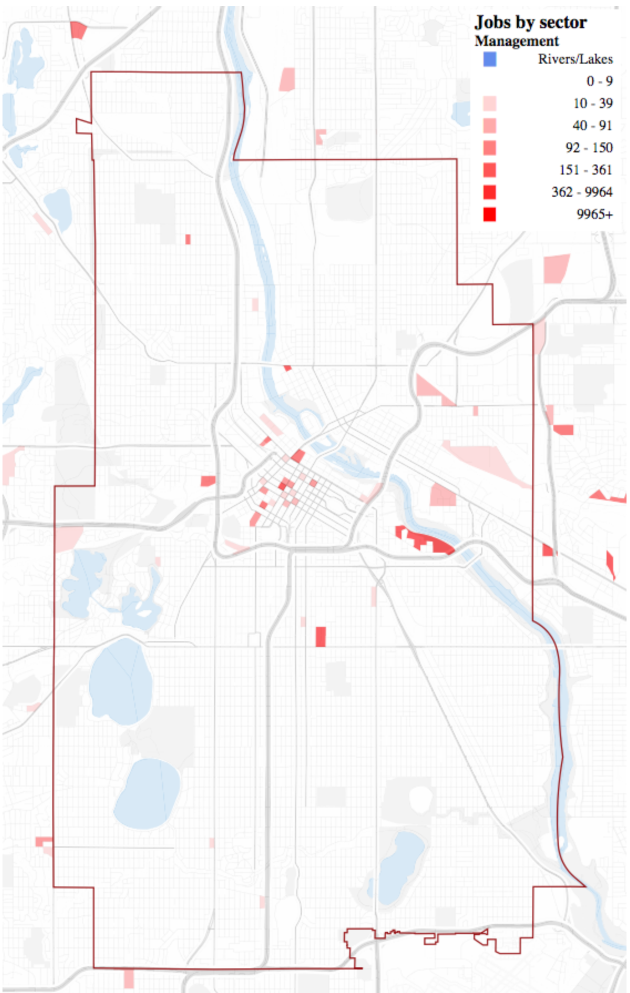

FIGURE 10 : Spatial distribution of management jobs in Minneapolis, based on LEHD data.

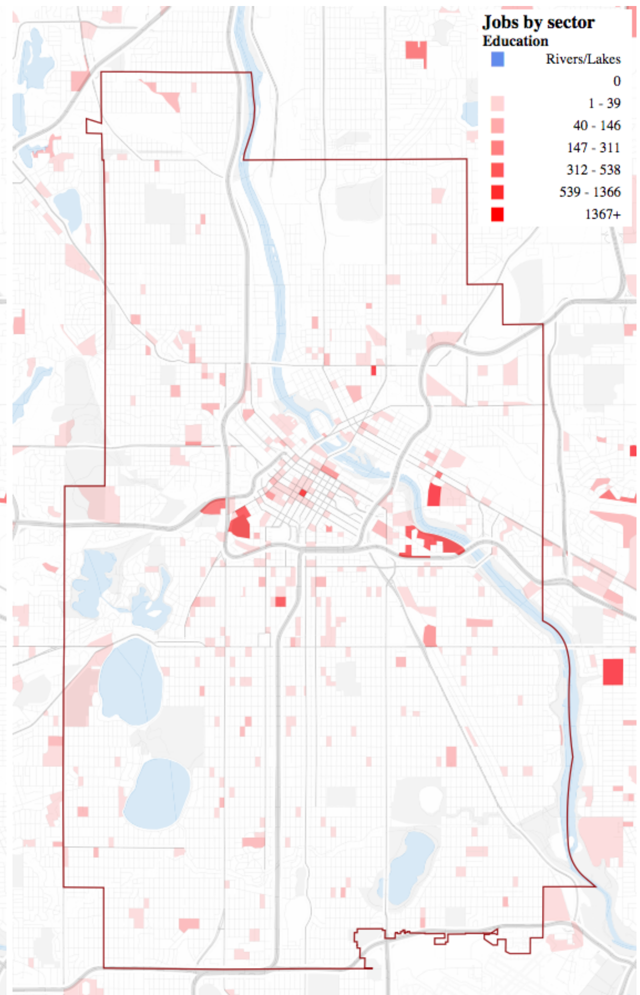

FIGURE 11 : Spatial distribution of education jobs in Minneapolis, based on LEHD data.

DISCUSSION & CONCLUSION

For the bivariate models of pedestrian activity in terms of census block centroid accessibility to jobs via walking, the evening peak period provided the best explanatory power. For all three time periods, as well as the 6-hour total count, R^2 values peaked near 15-minute thresholds, and dropped off in either direction. The correlation between walking accessibility and walking activity is positive. Walking is commonly thought of as a 15-minute-mode, in that the majority of people walking in urban areas will be on trips of duration 15 minutes or less. Further, in dense urban areas, distance matters - a high-threshold measurement of walking accessibility will tend to blur the results and differences between origin points, thus potentially failing to reflect local variabilities in walking patterns. Additionally, accessibility data at the 5-minute threshold level was found to be a consistently less significant predictor of pedestrian activity than higher thresholds.

It was found that pedestrian counts in the evenings exhibited the strongest correlations with the accessibility variables tested, and midday counts exhibited the weakest correlation strengths. It is possible that midday pedestrian traffic is more dispersed in both nature of trip-making and timing, due to variable work schedules. Both the morning and evening periods exhibited stronger correlations with job-based accessibility metrics, in accordance with traditional work commute timings. The subtle difference between the two periods could be explained in part through analysis of individual trip diaries - specifically the distributions of departure and arrival times for morning and evening trips.

As was hypothesized, both the accessibility measures and betweenness centrality exhibited positive influences on pedestrian activity levels, with all the significant variables with strongest R^2 metrics having positive signs. This gives a reasonable framework through which to estimate modal traffic levels at every intersection in Minneapolis and, by extension of the broader framework, in other cities as well. However, betweenness centrality did not exhibit as strong a positive correlation as was predicted. this may have resulted from the specific methodology used - that is, a centrality calculation that takes into account heterogeneous trip generation within an urban area due to varying land use patterns may lead to higher predictive power of centrality measures toward actual pedestrian behavior patterns. Pedestrian behavior in urban areas does not exhibit uniform all-to-all trip generation distribution; rather, there are major sources and attractors, which would shift the distribution of route choices, and thus link and intersection centrality, to favor routes between those origin-destination pairs. Applying techniques analogous to those in McDaniel et al. (*9*) to the walking model may yield more accurate pedestrian behavior estimation based on the centrality metric.

Accessibility to Education and Finance jobs was found to be significantly predictive of increased pedestrian activity, while accessibility to Management and Utilities jobs was found to be significantly predictive of decreased pedestrian activity, relative to other categories; these spatial maps are visible in Figure 8, Figure 9, Figure 10, and Figure 11. Utility jobs tend to be concentrated in areas not immediately in the downtown core, as well as management jobs to a lesser degree; finance jobs are heavily concentrated in the downtown core area, and education jobs are concentrated on walkable campuses. Further, it is plausible that certain categories of jobs attract greater or lesser levels of walking among their workers, dependent on such factors as dress requirements, vehicle needs (e.g. construction and contract workers), and typical density of jobs within each category. Additional cross-comparison analysis among economic job categories is needed to investigate these effects, but initial analysis indicates these spatial distributions correlate to the regression coefficients in Table 2.

A significant and pervasive challenge with analysis dependent on pedestrian, bicycle, and vehicular count data is the issue of data quality and format. Methodologies and data standards can vary from city to city and jurisdiction to jurisdiction; this study used a combination of national (Census, LEHD) datasets and local (Minneapolis traffic) data. Some cities, such as Boston, do not have robust pedestrian and bicycle counting programs throughout the city; others, such as Philadelphia, may have varying data release and non-disclosure agreements between MPOs, cities, and police departments; still other cities may have inconsistent data tracking and release practices, such as Washington, D.C. Such hurdles can make the collection and processing of pedestrian and bicycle spatial use data on a national scale exceedingly difficult. Better standards of practice in data collection, management, and distribution are needed.

However, with pedestrian activity estimation based on sampling existing counts, accessibility analysis, and betweenness centrality of the underlying network, it becomes possible to predict the landscape of pedestrian activity within the urban area. Such techniques may prove important in informing urban planning processes and decisions, pedestrian safety programs, and highlighting areas of the city that experience higher pedestrian activity as salient areas for fine-grained attention to built environment details. An important extension of the identification of intersections with higher potential pedestrian traffic is the visualization of such areas - e.g. downtown. We can reasonably expect certain levels of pedestrian traffic, even where counts may not exist.

There are a few caveats to mention regarding the ability of simply accessibility and centrality to accurately predict pedestrian behavior. Figure 7 highlights sections of the urban area where the model differed significantly from the actual pedestrian counts. For 741 intersections, the number of daily pedestrians was overpredicted, and for 275 intersections the model underpredicted pedestrian activity. The distribution of differences has a mean $\mu = -8.10$ and standard deviation $\sigma = 72.50$; 91.11% of the sampled intersections had $actual - estimated$ differences within 1 standard deviation from the mean. The cases of underestimation and overestimation are geographically interesting to note; the two major areas of underestimation are the inner downtown core, and the East Bank Campus of the University of Minnesota, just east of the Mississippi River, while the major area of overestimation is located west of Hennepin Ave in downtown, near Dunwoody Boulevard and Olson Memorial Highway. The downtown core and the campus of the University are characterized by significant pedestrian activity and are considered walkable areas, whereas the areas just west of downtown are not as walkable; in fact, Dunwoody Boulevard, Olson Memorial Highway, and other roads in the area are multi-lane automobile thoroughfares. While the road network structure and proximity to downtown would predict significant pedestrian activity, physical barriers exist within the built environment. These cases highlight the limitations of centrality and accessibility in capturing elements of the built environment relevant to pedestrian activity where local and hyper-local factors may play significant roles.

Future Directions

Phase II of this investigation will extend the above analysis framework to bicycling activity estimation, as well as extending both pedestrian and bicycle activity estimation out to other metropolitan areas as good data become available. Bicycle activity will be investigated in similar fashion - activity levels at intersections will be modeled with a time-threshold value of bicycle accessibility to jobs, betweenness centrality, net transit accessibility benefit, and accessibility to jobs split by sector. However, we hypothesize that adapting the betweenness measure to use spatial work trip distributions given by LEHD data will more closely reflect actual pedestrian use-cases than all-to-all O/D pair analysis. Bicycle accessibility will be calculated with OpenTripPlanner, and the modeling and analysis process will be exactly analogous to that for the walking accessibility data presented in this report.

REFERENCES

[1] Lumsdon, L. and J. Mitchell, Walking, transport and health: Do we have the right prescription? *Health Promotion International*, Vol. 14, No. 3, 1999, pp. 271–279.

[2] Raford, N. and D. Ragland, *Space Syntax: Innovative Pedestrian Volume Modeling Tool for Pedestrian Safety*, 2004.

[3] Brownstone, D., Key Relationships Between the Built Environment and VMT. *Transportation Research*, , No. Special report 298: Driving and the Built Environment: The Effects of Compact Development on Motorized Travel, Energy Use and CO2 emissions, 2008, pp. 2–4.

[4] World Health Organization, Global status report on road safety. *Injury prevention*, 2013, p. 318.

[5] Jensen, S. r. U., Safety effects of blue cycle crossings: A before-after study. *Accident Analysis and Prevention*, Vol. 40, No. 2, 2008, pp. 742–750.

[6] Bureau, U. S. C., Modes Less Traveled Bicycling and Walking to Work in the United States : 2008 2012. , No. May, 2014, pp. 2008–2012.

[7] Schneider, R. J., R. M. Ryznar, and A. J. Khattak, An accident waiting to happen: A spatial approach to proactive pedestrian planning. *Accident Analysis and Prevention*, Vol. 36, No. 2, 2003, pp. 193–211.

[8] Iacono, M., K. J. Krizek, and A. El-Geneidy, Measuring non-motorized accessibility: issues, alternatives, and execution. *Journal of Transport Geography*, Vol. 18, No. 1, 2010, pp. 133–140.

[9] McDaniel, S., M. B. Lowry, and M. Dixon, Using Origin-Destination Centrality to Estimate Directional Bicycle Volumes, 2013, p. 15.

[10] Lowry, M., Spatial interpolation of traffic counts based on origin-destination centrality. *Journal of Transport Geography*, Vol. 36, 2014, pp. 98–105.

[11] Schneider, R. J., L. S. Arnold, and D. R. Ragland, *Pilot Model for Estimating Pedestrian Intersection Crossing Volumes*, 2010.

[12] Miranda-Moreno, L. F., P. Morency, and A. M. El-Geneidy, The link between built environment, pedestrian activity and pedestrian-vehicle collision occurrence at signalized intersections. *Accident Analysis and Prevention*, Vol. 43, No. 5, 2011, pp. 1624–1634.

[13] Kamargianni, M., Generation s Y Travel Behavior and Perceptions Towards Walkability Constraints. *Transportation Research Board 93rd Annual Meeting. January 12-16, Washington, D.C.*, 2014.

[14] Anciães, P. R., Urban transport , pedestrian mobility and social justice A GIS analysis of the case of the Lisbon Metropolitan Area. , No. July, 2011.

[15] Do, M. T., O. Grembek, D. Ragland, and C.-Y. Chan, Weighting integration by block heterogeneity to evaluate pedestrian activity, 2013, pp. 14p, schmas,\rtabl., ill., bibliogr.

[16] Zhang, Y., J. Bigham, D. Ragland, and X. Chen, Investigating the associations between road network structure and non-motorist accidents. *Journal of Transport Geography*, Vol. 42, 2015, pp. 34–47.

[17] Dai, D., E. Taquechel, J. Steward, and S. Strasser, The impact of built environment on pedestrian crashes and the identification of crash clusters on an urban university campus. *The western journal of emergency medicine*, Vol. 11, No. 3, 2010, pp. 294–301.

[18] McCahil, C. and N. W. Garrick, The Applicability of Space Syntax to Bicycle Facility Planning. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2074, No. -1, 2008, pp. 46–51.

[19] Derrible, S., No Title. In *Using Network Centrality to Determine Key Transfer Stations in Public Transportation Systems*, 2012.

[20] Dowling, R. G. and B. K. Ostrom, *Highway Capacity Manual 2010*. Transportation Research Board, Washington, D.C., 2010.

[21] Federal Highway Administration, 2009 National Household Travel Survey, 2010, p. 82.

[22] Owen, A. and D. Levinson, *Access Across America: Transit 2014*. University of Minnesota Center for Transportation Studies, http://www.its.umn.edu/Publications/ResearchReports/pdfdownloadl.pl?id=2506, 2014.