FINAL REPORT

Multi-City, National Scale Direct-Demand Models of Peak-Period Bicycle and Pedestrian Traffic

Date: May 2017

Huyen Le, PhD Student, Virginia Tech
Ralph Buehler, PhD, Associate Professor, Virginia Tech
Steve Hankey, PhD, Assistant Professor, Virginia Tech

Prepared by:
School of Public and International Affairs, Virginia Tech
140 Otey St., Blacksburg, VA 24061
1021 Prince St., Alexandria, VA 22314

Prepared for:
Virginia Center for Transportation Innovation and Research
530 Edgemont Road
Charlottesville, VA 22903

1. Report No.	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Multi-City, National-Scale Direct-Demand Models of Peak-Period Bicycle and Pedestrian Traffic		**5. Report Date** May 2017 **6. Performing Organization Code**
7. Author(s) Huyen Le, Ralph Buehler, Steve Hankey		**8. Performing Organization Report No.**
9. Performing Organization Name and Address School of Public and International Affairs, Virginia Tech 140 Otey St., Blacksburg, VA 24061 1021 Prince St., Alexandria, VA 22314		**10. Work Unit No. (TRAIS** **11. Contract or Grant No.** DTRT13-G-UTC33
12. Sponsoring Agency Name and Address US Department of Transportation Office of the Secretary-Research UTC Program, RDT-30 1200 New Jersey Ave., SE Washington, DC 20590		**13. Type of Report and Period Covered** Final 6/1/16 – 5/31/17 **14. Sponsoring Agency Code**

15. Supplementary Notes

16. Abstract

Direct-demand models are potentially useful tools for generating spatial estimates of pedestrian and bicycle traffic volumes to help plan for active transport facilities and target infrastructure investments. To date, most direct-demand models are city-specific; lack of spatial and temporal coverage of traffic counts on a national scale has precluded generalizability and transferability of city-specific models.

This project aims to address this limitation by sourcing peak-period non-motorized traffic counts at 6,342 locations across 20 U.S. metropolitan statistical areas (MSAs) to estimate spatial patterns of bicycle and pedestrian traffic. We developed models to estimate bicycle and pedestrians traffic at intersections and segments during two-hour morning and afternoon peak periods. Our models have reasonable goodness of fit for both bicycle traffic (adjusted R^2: 0.19 to 0.56) and pedestrian traffic (adjusted R^2: 0.45 to 0.72). We found a number of land-use and network variables that were correlated with bicycle and pedestrian traffic, for example, multimodal network density, presence of water bodies, nearby offices, industrial area, zero-car households, as well as bicycle and walking commuting mode shares. Intersection density is also a strong predictor for pedestrian volume; off-street and on-street bicycle facilities are strong predictors of bicycle volume.

Our count data have good spatial and temporal coverage across a variety of cities and regions in the US. Estimating models across cities allows for estimating non-motorized traffic in cities where counts are inadequate or unavailable with higher reliability. Our models could be used to inform decisions on where to locate non-motorized transportation facilities and to assess exposure to accidents with motor vehicles or other environmental hazards.

17. Key Words Non-motorized transportation, direct-demand model, bicycle and pedestrian counts, active travel, built environment	**18. Distribution Statement** No restrictions. This document is available from the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 28	**22. Price**

Acknowledgement

The authors would like to thank the Transportation Research and Education Center at Portland State University for their effort cleaning data from the National Bicycle and Pedestrian Documentation Project. We also thank Bryce Johnson, student in the Master of Urban and Regional Planning program at Virginia Tech, for his assistance in this project.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Executive Summary

Many cities in the U.S. have stated goals to increase walking and bicycling to improve public health, reduce emissions, and increase livability. Quantifying the relationship between the built environment on walking and bicycling levels can help plan for active transport facilities and target infrastructure investments. Direct-demand models are potentially useful tools for generating spatial estimates of pedestrian and cyclist traffic volumes to inform these decisions.

To date, most studies have used city-specific traffic counts to build city-level direct-demand models. However, lack of spatial and temporal coverage of traffic counts on a national scale have precluded generalizability from these models. This limitation underscores the need for development of multi-city, national-scale direct-demand models of bicycle and pedestrian traffic in the continental US.

This project aims to address this limitation by sourcing peak-period non-motorized traffic counts at 6,342 locations across 20 U.S. metropolitan statistical areas (MSAs) to estimate the spatial patterns of bicycle and pedestrian traffic. The count locations were geocoded and surrounding land-use types and transportation network features were tabulated as potential predictor variables. Based on the count data (i.e., dependent variables) and land-use variables (i.e., independent variables) we developed direct-demand models to estimate spatial patterns of bicycle and pedestrian traffic across the 20 MSAs.

We developed (1) base case models (including all data from 20 MSAs) and (2) alternative models for spatially concentrated count locations (including data from MSAs that have 100 locations or more) to predict bicycle and pedestrians traffic at intersections and segments during two-hour morning and afternoon peak periods. Bicycle facility (i.e., bicycle lane or off-street trail) data were not available across all cities. We developed an additional set of models for bicycle traffic for the nine MSAs where bicycle facility data to assess the impact of bicycle infrastructure.

Our models demonstrate reasonable goodness of fit for both bicycle traffic (adjusted R^2: 0.19 to 0.56) and pedestrian traffic (adjusted R^2: 0.45 to 0.72). We found a number of land-use and network variables that were correlated with bicycle and pedestrian traffic, such as multimodal network density, presence of water, nearby offices, industry, and zero-car households, as well as greater bicycle and walking commuting mode shares. Intersection density is also a strong predictor for pedestrian volume, while off-street and on-street bicycle facilities are strong predictors of bicycle volume. Despite data and modeling limitations, our models produced robust outcomes, which were validated using a cross validation and sub-sampling method (i.e., modeling a subset of MSAs with spatially dense traffic count locations). However, some variables showed mixed results (positive and negative effects among models) including

household density, types of housing units, retail, and service land use. This indicates that practitioners should be cautious when using these land use types as predictors for bicycle and pedestrian traffic.

A strength of our approach is that our count data have good spatial and temporal coverage across a variety of cities and regions in the US. The coverage of the count data combined with the use of predictor variables that are available at the national-scale allows for estimating non-motorized traffic in cities where counts are inadequate or unavailable (a limitation of previous work). Our models could be used to inform decisions on where to locate non-motorized transportation facilities and to assess exposure to accidents with motor vehicles or other environmental hazards.

Table of Contents

List of Figures

List of Tables

Problem Statement

Many cities in the U.S. have stated goals to increase walking and bicycling to improve public health, reduce emissions, and increase livability (Hankey et al., 2016; Jackson et al., 2013). Providing generalizable information about the effect of the built environment on walking and bicycling levels may help cities plan for active transport facilities and target infrastructure investments to improve walkability and bikeability. Direct-demand models, a statistical tool that allows modelers to predict traffic volume based on land-use and transportation network attributes, are potentially useful tools for generating spatial estimates of pedestrian and cyclist traffic volumes to inform these decisions.

To date, most studies have used city-specific traffic counts to assess correlates of active travel and the built environment (Hankey & Lindsey, 2016; Tabeshian & Kattan, 2014; Miranda-Moreno & Fernandes, 2011). However, lack of spatial and temporal coverage of pedestrian and bicycling traffic counts on a national scale have precluded generalizability from these city-level studies. To the best of our knowledge, no study has utilized pedestrian and bicycle count data across multiple metropolitan areas to develop direct-demand models and predict non-motorized traffic at locations without counts.

The lack of generalizability among previous city-specific models underscores the need for development of multi-city, national-scale direct-demand models of bicycle and pedestrian traffic in the continental US. These statistical models could be used to identify potential locations of future bicycle and pedestrian facilities or to develop estimates of crash rates. They could also be used to estimate non-motorized traffic volumes in cities where there are few counts or counts are unavailable.

Approach

Our work aims to address this research gap by developing a set of direct-demand models to estimate non-motorized traffic using bicycle and pedestrian traffic counts (i.e., dependent variable) based on land use, transportation network, and temporal data (i.e., independent variables). The overall workflow is shown in Figure 1; detailed descriptions of each task are in the main body of the report.

Task 1: Non-motorized traffic count aggregation

For 20 US metropolitan statistical areas (MSAs) we sourced and aggregated pedestrian and bicycle traffic counts that were collected over a span of 15 years. The count data were obtained from the National Bicycle and Pedestrian Documentation Project (NBPDP) database or requested directly from each MSA. We selected and aggregated counts during morning and afternoon peak periods (two hours each) since those two time periods represent the most frequently counted hours-of-day. Although we collected counts across seasons, we focused mainly on fall counts since that was the most common season data were collected in the MSAs in our database, which resulted in 9,870 observations for bicycle counts and 7,644 observations for pedestrian counts.

Task 2: Tabulating independent variables

For each count location, we tabulated surrounding land use, traffic, street network, destination accessibility, and socio-demographic variables at 12 buffer sizes (100-3,000m). Weather data were also included as control variables. We sourced this data from nationally available sources including: American Community Survey (ACS), Smart Location Database (SLD), and the National Oceanic and Atmospheric Administration (NOAA).

Task 3: Developing base-case direct-demand models

We used forward stepwise linear regression models to develop four base-case models for morning and afternoon peak-period bicycle and pedestrian traffic volumes. Independent variables were selected from the database created in Task 2 based on statistical inclusion criteria described below (see Methodology).

Task 4: Developing alternative models

We developed two alternative models that may inform efforts to improve direct-demand modeling in the future:

Alternative model 1 – Bicycle facility data: We developed direct-demand models of bicycle traffic using a subset of cities where bicycle facility data are available. Then, we compared the model results to the base-case models that did not include bicycle facilities as an independent variable.

Alternate model 2 – Spatially concentrated traffic counts: We developed alternate models that include only cities with a large number of traffic counts (i.e., >100 count locations). We compared model performance for models that use spatially concentrated traffic count data as compared to the base-case models.

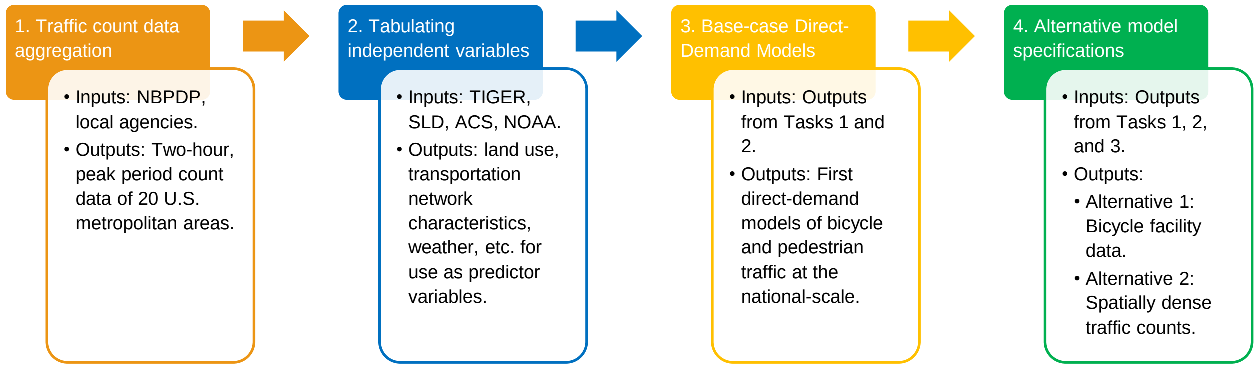

Figure 1. Workflow for the database management and modeling approach.

Methodology

Our work employs stepwise linear regression to develop direct-demand models of bicycle and pedestrian traffic based on neighborhood-level (i.e., Census block group) land use, transportation network, and socio-demographic characteristics. This section describes methods used for collecting, processing, and modeling the bicycle and pedestrian traffic count data.

Data Sources

Data used in this study were collected from various sources (Table 1). We acquired bicycle and pedestrian count data from the National Bicycle and Pedestrian Documentation Project database and by contacting individual local agencies to obtain updated counts for subsequent years. Other data, such as land-use and transportation network characteristics, were collected from the ACS 5-year summary, EPA's SLD, and TIGER. All data are publicly available at the national level. We also collected weather variables, such as temperature and precipitation, from the NOAA website, which is also publicly available at the national level. Bicycle facility data were obtained from Google Earth imagery with historical view, which allows us to track the changes in facility types over time.

Figure 2. Map of metropolitan areas with available bicycle or pedestrian count data.

Table 1. Data description

Type of Data	Source	Unit of measurement	Areal unit of base data	Tabulation method	Year
Bicycle and pedestrian traffic counts	NBPDP; Local agencies	AM/PM peak-hour	Point	-	2000-2016
Land use data					
Industrial	ACS/SLD	Job	Block group	Buffer	2010
Service	ACS/SLD	Job	Block group	Buffer	2010
Retail	ACS/SLD	Job	Block group	Buffer	2010
Office	ACS/SLD	Job	Block group	Buffer	2010
Water	TIGER shapefile	Square meter	Polyline	Buffer	2014
Housing unit	ACS/SLD	Unit	Block group	Buffer	2010
Number of households	ACS/SLD	Household	Block group	Buffer	2010
Transportation-related data					
Number of zero-vehicle households	*ACS/SLD*	Household	Block group	Buffer	2010
Bicycle commute mode share	ACS	Percent	Block group	Buffer	2014
Walking commute mode share	ACS	Percent	Block group	Buffer	2014
Public transport commute mode share	ACS	Percent	Block group	Buffer	2014
Number of public transit stops	SLD/GTFS	Stop	Point	Buffer	2010
Total road network density	SLD	Miles/sq mile	Block group	Buffer	2010

Type of Data	Source	Unit of measurement	Areal unit of base data	Tabulation method	Year
Network density in terms of facility miles of multi-modal links per square mile[1]	SLD	Miles/sq mile	Block group	Buffer	2010
Street intersection density (weighted, auto-oriented intersections eliminated)[2]	SLD	Intersections/sq mile	Block group	Buffer	2010
Bicycle facility	Google Earth	Type	Point estimate	Point	
Socioeconomics					
Median household income	ACS	US dollar	Block group	Buffer	2014
Population below 18 years of age	ACS	Percent	Block group	Buffer	2014
Population from 18 to 45 years of age	ACS	Percent	Block group	Buffer	2014
Population from 46 to 65 years of age	ACS	Percent	Block group	Buffer	2014
Population above 65 years old	ACS	Percent	Block group	Buffer	2014
Weather data	NOAA				2000-2016
Precipitation		Inch			
Temperature		Degree Fahrenheit			

[1] For a full description of these data, please see the SLD User's Guide (2014, pp.20-23) at https://www.epa.gov/sites/production/files/2014-03/documents/sld_userguide.pdf
[2] Ibid.

Data Processing

A core task in this project was to assemble a national-scale database of bicycle and pedestrian traffic counts and to assemble corresponding land-use and transportation network variables. Below we describe our method for data cleaning and aggregation as well as give descriptive statistics of the database.

Count Data

We obtained traffic count data from each jurisdiction and from the NBPDP database. It is noteworthy that the count data used in this study is not exhaustive; that is, more count data might be available for each MSA, however, they were not available when we requested.

Count data were obtained in two different formats, depending on their availability for each jurisdiction: (1) raw count data or (2) aggregated count data. Raw count data are typically recorded in 15-minute intervals. Aggregated count data are the total traffic counts at each location in two hour peak periods. Most count campaigns focus on morning (7-9AM) and/or afternoon (4-6PM or 5-7PM) peak-periods on weekdays, and the lunch hour (11AM-1PM) peak-period on weekends; however, counts were also collected to a lesser degree during other times of day in a number of metropolitan areas. For the purpose of building our models, we focused on weekday, morning and afternoon peak-period counts. We aggregated all 15-minute counts into two-hour counts to allow for comparison across geographies.

When counts were reported for unidirectional flows we converted to bidirectional counts at each location by adding unidirectional counts at the same segments. Our dataset includes counts at street segments (i.e., screenline counts) and intersections. We separated the two types of count locations due to their different nature and thus difference in absolute volumes. For intersection counts with turning movements, we separated the counts into segment counts for each leg for use in the segment models.

Finally, we aggregated counts by season (focusing on the fall because fall counts are the most abundant cross jurisdictions). Specifically, we took the average of all counts at the same location, in the same peak period, in the same year, for each season (Figure 3). Since the participating MSAs are in different geographic regions, we generally grouped counts conducted in August to November as fall counts. The final counts were then normalized by log-transforming for modeling. The zero values of counts were dropped as log of zero is undefined, although the zero values only account for 2.9% of the total number of observations in our dataset. In this report, we only use count data collected in fall to simplify the model building process.

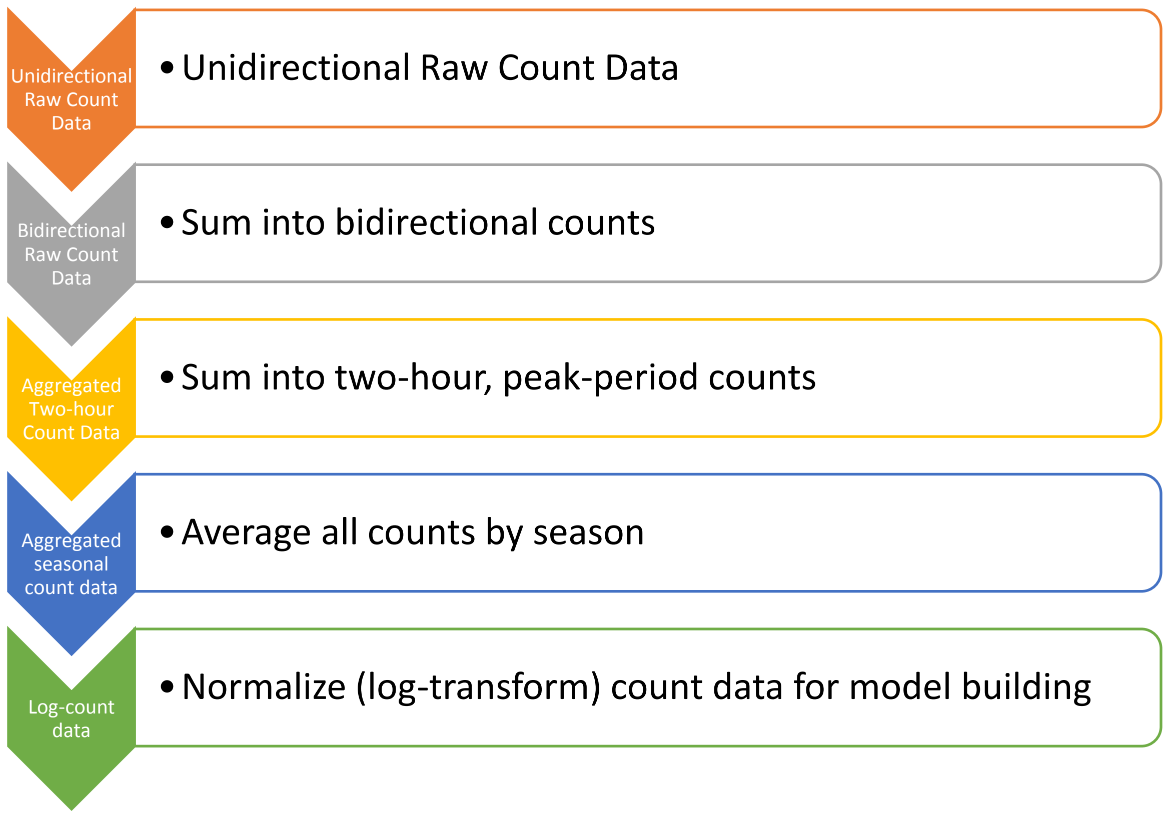

Figure 3. Data processing workflow

Table 2 provides the number of count observations in the fall season by peak period. In total, 9,870 observations were collected for bicycle traffic, of which, afternoon peak counts account for 64% of the sample. Pedestrian traffic counts totaled 7,644 observations; 60% represent afternoon peak periods. Most MSAs collected traffic counts during the afternoon peak; about half of the MSAs did not count bicyclists and pedestrians during morning peak hours. Traffic counts collected at intersections account for 47% and 49% of the total sample for bicycle and pedestrian traffic, respectively.

Table 2. Bicycle and pedestrian counts by peak period and location type (fall counts only)

	Bicycle				Pedestrian			
	AM		PM		AM		PM	
Metropolitan Area	Intersctn	Segment	Intersctn	Segment	Intersctn	Segment	Intersctn	Segment
Blacksburg, VA		101		101		72		72
Boston, MA	4	39	4	41	4	8	1	8
Champaign Urbana, IL	66	255	66	255	121		121	
Cleveland, OH				82				81
Columbus, OH	7	213			7	213		
Denver, CO	48		74					
Hartford, CT	3	1	61	11	3	1	61	11
Lawrence, KS				100				100
Los Angeles, CA	462	520	461	428				
Madison, WI	91	73	144	73	0	73	0	73
Manhattan, KS			112				112	
Minneapolis, MN				950				950
New York City, NY						1,022		1,022
Philadelphia, PA		198		190		158		162
Portland, OR			36	55			36	55
San Francisco, CA			308	1,241			79	1
Seattle, WA	303		305		305		305	
St Louis, MO			142				238	
Tucson, AZ	1,058		1,060		1,075		1,074	
Washington, DC	10	54	10	54	10		10	
Total	**2,052**	**1,454**	**2,783**	**3,581**	**1,525**	**1,547**	**2,037**	**2,535**

Table 3 and Table 4 show bicycle and pedestrian counts by year, respectively. Most jurisdictions conducted counts during the 2010-2016 period. Bicycle traffic was counted repeatedly at 3,141 locations (i.e., each location were counted at least twice within 15 years), or was measured once at 1,650 locations during the study period from 2001 to 2016. Similarly, pedestrian traffic was counted repeatedly at 1,717 locations, and counted once at 1, 274 locations in 15 years.

Table 3. Bicycle counts by year

Bicycle	Year													
Metropolitan Area	2001	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Blacksburg, VA													202	
Boston, MA	8	2	34			29	20	39	16	50	31	12	9	
Champaign Urbana, IL										620				40
Cleveland, OH									24	40	45	24		
Columbus, OH			39	17	33	28	43	77	38	43	34	35	38	32
Denver, CO					22			2	7	4	10	37	59	
Hartford, CT													28	48
Lawrence, KS										17	26	20	22	24
Los Angeles, CA						56	32	221	525	171	415	141	347	162
Madison, WI		12	12	12	12	12	12	98	60	48	157	119	74	18
Manhattan, KS												26	24	62
Minneapolis, MN					51	78	153	84	131	142	144	167		
New York City, NY					48									
Philadelphia, PA							136	136	197	22	41	73	235	120
Portland, OR						39	52	52						
San Francisco, CA				150	133	146	150	146	197		244	388		
Seattle, WA								608						
St Louis, MO										33	32	40	37	
Tucson, AZ						106	188	180	194	149	272	506	525	4
Washington, DC						74	74	106	100	115	140	96	228	46

Table 4. Pedestrian counts by year

Pedestrian	Year												
Metropolitan Area	2001	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Blacksburg, VA												144	
Boston, MA	4	18			4	3	7	4	9	8	7	3	
Champaign Urbana, IL							10	18	170	4	6	18	16
Cleveland, OH								23	40	45	24		
Columbus, OH		39	17	33	28	43	77	38	43	34	35	38	32
Hartford, CT												28	48
Lawrence, KS									17	26	20	22	24
Madison, WI							66	28		44	42	14	18
Manhattan, KS											26	24	62
Minneapolis, MN				51	78	153	84	131	142	144	167		
New York City, NY				476	450	454	456	456	456	456	456	456	228
Philadelphia, PA							100	237	28	37	65	51	118
Portland, OR					39	52							
San Francisco, CA											80		
Seattle, WA							608				1	1	
St Louis, MO									39	50	66	83	
Tucson, AZ		8	78				173	190	513	274	504	499	4
Washington, DC									4	36		32	10

Land-Use and Transportation Network Data

For each count location, we tabulated land-use and transportation network data using the Land Use Regression (LUR) tools (Akita, 2014) for ArcGIS. This set of tools allows us to measure areas of polygons, number of points, or distance of lines that fall inside a buffer. Using LUR tools, we measured the area of each land-use type, weighted average income, percentage of population by age group (and other socio-economic characteristics), and other variables listed in Table 1 at 12 different buffer sizes: 100m, 200m, 300m, 400m, 500m, 750m, 1000m, 1250m, 1500m, 2000m, 2500m, and 3000m. The sizes of the buffers were chosen based on a similar study by Hankey & Lindsey (2016).

Weather Data

We obtained weather data from NOAA for each count date and each city. We assigned the lowest temperature of each day for morning peak counts, and the highest temperature of each day for afternoon peak counts. Only daily average precipitation was used for each count location, regardless of peak periods. We then aggregated temperature and precipitation data by season along with the count data.

Analysis

Once the data were cleaned and aggregated, we developed a set of direct-demand models for bicycle and pedestrian traffic during morning and afternoon peak periods. We applied the forward stepwise regression approach to select the variables most correlated with active travel among a set of possibly relevant variables listed in the previous section. This method has been applied in previous studies (Hankey & Lindsey, 2016).

We modeled bicycle and pedestrian traffic (dependent variable) using land use, transportation network, weather, and socio-demographic variables as predictors (independent variables). The choice of independent variables was based on the existing literature (Jones et al., 2010; Miranda-Moreno & Fernandes, 2011; Schneider et al., 2011; Tabeshian & Kattan, 2014; Fagnant & Kockelman, 2016; Hankey & Lindsey, 2016) and professional judgement. Each variable was measured at 12 buffer sizes and included for selection in the model building process. However, each variable was only allowed to be selected once among all buffer sizes.

We used forward stepwise linear regression to select statistically meaningful variables for the final models. Specifically, the independent variable with the highest correlation with the dependent variable (log of bicycle and pedestrian count) was selected first. The process continues by searching for the independent variable with the highest correlation with the model residuals. To avoid multicollinearity, our procedure did not select variables that are highly correlated with one of the previously chosen independent variables (using a check for Variance

Inflation Factor). The process continued until the last coefficient of the independent variable included in the model was statistically insignificant (at 0.05 level) or violated criteria for multicollinearity; this variable was then removed and the model was completed. Descriptive statistics of the dependent and independent variables are provided in Table 5 and Table 6.

Table 5. Descriptive statistics for number of pedestrians and bicyclists counted (not log-transformed)

MSA		Bicycle				Pedestrian			
		AM		PM		AM		PM	
		Intersctn	Segment	Intersctn	Segment	Intersctn	Segment	Intersctn	Segment
Blacksburg, VA	Mean		6.8		10.2		33.5		65.0
	SD		8.9		13.9		48.5		132.3
	Median		4		5		20		23.5
Boston, MA	Mean	21.0	639.0	128.0	591.0	37.9	319.7	726.5	186.7
	SD	33.6	954.8	223.4	945.2	57.0	413.2		327.1
	Median	6	192	18	181.5	14	210	726.5	57.5
Champaign Urbana, IL	Mean	20.6	10.5	68.2	35.0	82.3		185.4	
	SD	27.3	15.9	31.6	21.9	126.0		280.3	
	Median	11	4	70.5	31	30		62	
Cleveland, OH	Mean				26.9				84.6
	SD				28.1				103.63
	Median				14				49.5
Columbus, OH	Mean	2.6	25.0			63.6	150.9		
	SD	2.5	35.7			75.0	189.9		
	Median	3	12			12	63		
Denver, CO	Mean	42.1		70.5					
	SD	38.5		52.7					
	Median	29		60					
Hartford, CT	Mean	6.3	25.0	17.8	19.3	90.3	116.0	86.8	50.9
	SD	5.1		17.3	16.5	70.1		139.4	59.2
	Median	5	25	13	15	122	116	36	19
Lawrence, KS	Mean				17.4				47.9
	SD				17.3				95.6
	Median				12				16
Los Angeles, CA	Mean	45.3	41.9	56.3	69.8				
	SD	43.9	56.3	61.8	85.8				
	Median	35	24	42	43				
Madison, WI	Mean	163.9	110.7	224.8	139.7		58.2		96.1
	SD	168.7	96.9	253.3	121.2		37.5		80.6
	Median	87	97	127	104		51		68
Manhattan, KS	Mean			20.4				87.3	
	SD			23.5				136.4	
	Median			12				23	

MSA		Bicycle				Pedestrian			
		AM		PM		AM		PM	
		Intersctn	Segment	Intersctn	Segment	Intersctn	Segment	Intersctn	Segment
Minneapolis, MN	Mean				108.9				149.4
	SD				140.0				279.4
	Median				58				68
New York City, NY	Mean						2127.4		3812.3
	SD						2389.4		3860.4
	Median						1259		2688
Philadelphia, PA	Mean		27.0		37.4		268.7		472.5
	SD		53.5		59.4		399.2		687.7
	Median		9		21		99		179
Portland, OR	Mean			46.2	176.8			61.1	112.0
	SD			42.2	246.2			47.0	160.4
	Median			30	53			47.5	57
San Francisco, CA	Mean			304.6	149.9			1639.8	9773.0
	SD			320.3	216.9			1908.7	
	Median			179	61			907	9773
Seattle, WA	Mean	37.1		38.4		72.9		88.2	
	SD	78.4		78.3		123.1		187.5	
	Median	11		13		36		42	
St Louis, MO	Mean			25.12				69.4	
	SD			35.95				115.4	
	Median			13				29	
Tucson, AZ	Mean	52.7		54.1		64.7		84.9	
	SD	88.4		95.8		135.3		221.7	
	Median	25		24		20		21	
Washington, DC	Mean	80.0	76.1	55.1	68.8	7.9		2.7	
	SD	72.7	108.7	71.9	106.3	20.1		8.2	
	Median	82	48	32	30	0		0	
Total	**Mean**	**52.2**	**50.2**	**86.4**	**108.5**	**67.3**	**1459.7**	**149.5**	**1639.4**
	SD	**85.7**	**191.3**	**165.1**	**197.1**	**131.4**	**2159.6**	**518.7**	**3049.6**
	Median	**26**	**15**	**31**	**44**	**24**	**597**	**31**	**181**

Table 6. Descriptive statistics of independent variables

	100m buffer size			1000m buffer size			3000m buffer size		
Variable	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Total jobs	35093	5812	91444	22030	3215	59484	122798	37192	243977
Retail jobs	30	3	97	2041	460	5295	9784	3309	18258
Office jobs	139	4	522	12304	581	38833	42833	5893	93649
Households	89	37	132	8723	4174	10452	63147	29409	72320
Housing units	100	41	150	9724	4567	11912	69799	32322	81080
Zero-car households	32	17	39	3189	2000	2953	23495	14115	20532
Industry jobs	50	3	454	3159	478	8943	16340	5429	28199
Entertainment jobs	62	6	263	3956	638	11055	17100	4688	34899
Service jobs	178	18	514	13655	2817	31872	65307	27401	109553
Median household income	58068	50926	36571	58221	54294	29169	59044	54054	24521
Transit mode share	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Bicycle mode share	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Walking mode share	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1
Below 18 years old	14.2	13.5	9.5	14.9	14.4	7.9	16.2	15.4	5.5
18 to 45 years old	50.3	46.1	20.5	48.2	45.3	16.1	43.8	42.6	11.5
46 to 65 years old	21.5	22.5	9.8	22.0	23.0	7.1	22.3	22.3	5.2
Above 65 years old	14.0	11.1	14.1	14.9	12.2	10.7	17.7	13.9	11.1
# of transit stops	1.2	0.0	1.9	55.8	37.0	60.7	382.8	272.0	390.0
Water	1034	0	4331	148768	1594	328392	2529278	561853	4004732
Total road network density	24.4	23.4	10.7	24.3	23.2	11.1	24.2	21.9	14.7
Intersection density	123.2	103.0	90.8	119.1	104.4	68.1	105.1	96.2	50.0
Multimodal network density	4.1	3.2	3.8	3.6	3.2	2.4	3.1	3.0	1.4
Temperature (no buffer)	70.5	70.0	12.3						
Precipitation (no buffer)	0.3	0.0	1.2						

Note: Three illustrative buffer sizes of small, medium, and large buffers from the count locations. Variables were tabulated for other buffer sizes and were not described in this table.

Findings

Results

We developed three sets of direct-demand models: full models (with all count data from the 20 metropolitan areas), bicycle facility models (with nine metropolitan areas where bicycle facility data are available), and spatially concentrated models (with jurisdictions that have 100 or more count locations for each mode).

Full Models

We first developed models that included (1) all MSAs listed in our database and (2) all nationally available variables as candidates for selection. All models except one (the bicycle AM segment model) show relatively good fit, with adjusted R^2s ranging from 0.38 – 0.56 for bicycle traffic, and 0.45 to 0.72 for pedestrian traffic. Given the low model fit, we do not recommend using the bicycle AM segment model until further work has been completed to improve the model. The full models are shown in Table 8; coefficients are shown in each cell, with the buffer size selected in the model building process in parentheses. Since the dependent variables (bicycle and pedestrian counts) were log-transformed, the coefficients are interpreted as percentage change in traffic volume.

Variables describing characteristics of the transportation network entered all models. Multimodal network density, measured as miles of streets that accommodate various modes of transport per square mile, was a strong positive predictor for bicycle traffic, with a stronger relationship to traffic volume at segments as compared to intersections (17 – 40% at segment count locations versus 4-9% at intersections), holding other variables constant. However, multimodal network density is negatively correlated with pedestrian traffic in the segment count models, indicating that higher multimodal network density is associated with a lower pedestrian volume. A similar effect on pedestrian traffic is found for total road network density. Further, higher intersection density has a positive association with pedestrian volume. An increase of one intersection per square mile is associated with a 0.6-0.8% increase in pedestrian traffic. The mixed findings for pedestrian traffic volumes may indicate confounding effects among different measures of network density.

Land-use features entered most models. For example, areas of water bodies within 100 meter to one kilometer buffers of the count locations has a strong positive association with bicycle traffic (one hectare of water is associated with 8-70% more bicycle traffic), and a mixed, modest effect on pedestrian traffic (i.e., decrease 0.4% or increase up to 0.8% of pedestrian traffic). Office land use has a negative association with bicycle traffic, and a strong positive relationship with pedestrian traffic. Industrial area generally has a positive relationship with both bicycle and pedestrian traffic in six of the eight models. Retail and service show mixed associations with walking and cycling. A higher number of households within a buffer is correlated with higher non-motorized traffic across the models. In general, the direction of effect from these variables are similar to findings from other studies, although the magnitudes of effect are different (Hankey & Lindsey, 2016).

Our models also show that neighborhoods with more bicycle and walking commute mode share are associated with higher bicycle traffic counts. In contrast, neighborhoods with high public transit commuting mode share and high density of bus stops generally have lower non-motorized traffic. Similar to the relationships for various measures of network density this could signal confounding effects among these variables, for example, areas with higher levels of transit service could also be areas with high rates of walking and bicycling.

Rain, temperature, and socio-demographic variables (e.g., age and income) were included in the models as control variables. The sign of coefficients for these variables changed across models and would benefit from further refinement in future iterations of the models.

Table 7. List of MSAs that have traffic count data used in the full models

	AM Peak		PM Peak	
	Segment	Intersection	Segment	Intersection
Bicycle Models	Blacksburg	Boston	Blacksburg	Boston
	Boston	Champaign Urbana	Boston	Champaign Urbana
	Champaign Urbana	Columbus	Champaign Urbana	Denver
	Columbus	Hartford	Cleveland	Hartford
	Hartford	Los Angeles	Hartford	Lawrence
	Los Angeles	Madison	Lawrence	Los Angeles
	Madison	Seattle	Los Angeles	Madison
	Philadelphia	Tucson	Madison	Manhattan
	Tucson	Washington, DC	Minneapolis	Portland
	Washington, DC		Philadelphia	San Francisco
			Portland	Seattle
			San Francisco	St Louis
			Tucson	Tucson
			Washington, DC	Washington, DC

	AM Peak		PM Peak	
	Segment	Intersection	Segment	Intersection
Pedestrian Models	Blacksburg Boston Columbus Hartford Madison New York City Philadelphia Tucson	Boston Champaign – Urbana Columbus Hartford Seattle Tucson Washington, DC	Blacksburg Boston Cleveland Hartford Lawrence Madison Minneapolis New York City Philadelphia Portland San Francisco Tucson	Boston Champaign – Urbana Hartford Lawrence Manhattan Portland San Francisco Seattle St Louis Tucson Washington, DC

Table 8. Full direct-demand models of bicycle and pedestrian traffic in 20 MSAs

	Bicycle				**Pedestrian**			
	AM		PM		AM		PM	
Full models (Fall)	Segment	Intersection	Segment	Intersection	Segment	Intersection	Segment	Intersection
Multimodal network density	0.402 (3000)	0.039 (500)	0.166 (2000)	0.087 (750)	-0.130 (3000)	0.192 (1500)	-0.104 (3000)	0.080 (100)
Total network density				0.015 (1500)	-0.004 (2000)		-0.006 (1500)	
Intersection density					0.006 (750)	0.006 (750)	0.006 (500)	0.0089 (1500)
Water (hectare)		0.697 (100)	0.141 (200)	0.076 (300)	0.005 (750)	-0.004 (1000)	0.030 (400)	0.006 (750)
Office (thousand)		-0.146 (300)		-0.674 (100)	0.250 (100)	2.068 (100)	0.255 (100)	2.361 (100)
Industry (thousand)		0.016 (3000)	0.127 (300)	0.565 (200)	0.028 (750)	-0.041 (1500)	0.020 (750)	-0.462 (300)
Retail (thousand)		-0.048 (2500)	-0.709 (100)	0.032 (3000)				0.133 (750)
Service			0.005 (3000)			-0.00001 (3000)	0.253 (100)	-0.021 (1500)
Entertainment (thousand)					-0.233 (300)			
Housing Unit		-0.002 (100)						
Zero-car household (thousand)				0.0021 (100)	1.365 (300)	0.333 (1000)	0.039 (3000)	0.222 (1000)
Household (thousand)		0.022 (3000)			0.006 (3000)	-1.683 (200)	0.643 (200)	
Bicycle mode share		16.475 (3000)	12.171 (1000)	13.578 (3000)	-4.523 (200)			

	Bicycle				Pedestrian			
	AM		PM		AM		PM	
Full models (Fall)	Segment	Intersection	Segment	Intersection	Segment	Intersection	Segment	Intersection
Walking mode share		1.776 (100)	2.17 (1000)	1.474 (500)	4.260 (100)	4.133 (200)	2.881 (100)	3.932 (400)
Transit mode share		-2.483 (3000)	-4.539 (2000)	-4.409 (3000)	-7.123 (3000)	-2.080 (3000)	-10.192 (3000)	-5.799 (3000)
Transit stops		-0.001 (2500)	-0.003 (1500)	0.042 (100)	-0.003 (3000)	0.088 (100)	-0.002 (3000)	0.087 (100)
Precipitation		0.112	-0.064	0.124			0.106	
Temperature					0.004	0.013		-0.006
Income (thousand dollars)	0.008 (3000)	0.008 (2500)	-0.004 (750)	0.005 (3000)	-0.019 (2500)		-0.013 (3000)	-0.004 (100)
Under 18 years old		-0.024 (3000)	-0.016 (100)	-0.026 (3000)	-0.029 (1250)		0.015 (100)	-0.014 (1000)
18-45 years old		-0.018 (3000)		0.009 (300)	0.012 (2000)		0.009 (100)	
45-65 years old		-0.013 (750)	0.019 (3000)		0.052 (1500)			-0.022 (1000)
Above 65 years old			-0.021 (750)		-0.012 (200)		0.015 (2500)	-0.011 (300)
Number of MSAs	10	10	14	14	8	7	12	11
N	1348	1902	3271	2604	1471	1149	2525	1600
Adj-R2	0.19	0.43	0.38	0.56	0.66	0.45	0.72	0.58

Note: Buffer sizes are shown in parentheses. All dependent variables were log-transformed. All variables were significant at 0.05 level.

To validate and test the robustness of the models, we performed cross validation by employing a Monte Carlo-based hold-out analysis. Briefly, we randomly separated the dataset into two parts: a training set (containing a random selection of 90% of the original sample) to build the models, and a test set (containing the remaining 10% of the original sample) to validate the results of the training models. The process was repeated 100 times, resulting in eight sets of 100 training models each (for bicycle/pedestrian AM/PM peak at segment/intersection). The average adjusted R^2 from the training models were similar to the values displayed in Table 8. The training models were used to estimate traffic counts at locations in the test dataset and compare differences in estimates vs. actual counts; in general, the validation results were robust with only modest changes in adjusted R^2 of 0.01 to 0.04. Common variables selected in the training models were similar to the variables displayed in Table 8. This indicates that our model has reasonable out-of-sample prediction, and that the above selected variables are the most statistically significant factors in our dataset to consider when modeling bicycle and pedestrian traffic (other important variables may exist that we did not include in our dataset). However, we lacked a true external validation dataset to more rigorously test our models. Future work could use systematic hold out procedures of specific MSAs to test model performance.

Alternative Models with Bicycle Facility Data

In order to account for the impact of bicycle facilities on bicycle traffic, we developed another set of models that included bicycle facilities as an independent variable. At the time of writing this report, only nine MSAs in our sample had such data: Columbus, OH, Blacksburg, VA, San Francisco, CA, Madison, WI, Los Angeles, CA, Minneapolis, MN, Philadelphia, PA, Cleveland, OH, and Lawrence, KS. Count data from these areas were collected at street segments only (i.e., these models do not include intersection counts).

Bicycle facilities were categorized as follows: on-street facilities include sharrows (shared lane markings), bike lanes, buffered bike lanes, protected bike lanes, and bike boulevards; off-street facilities include trails and shared use paths that are completely separated from vehicular traffic. When none of the above bicycle facilities are present, the street segment is coded as no facility. Since network data on bicycle facilities was not available, our approach (see above) included assessing whether bicycle facilities existed based on Google Earth imagery. As such, on-street and off-street facility types are introduced in the models as dummy variables.

Two models shown in Table 9 perform fairly well in predicting bicycle traffic (Adjusted $R^2 = 0.46 – 0.51$). Bicycle facility variables were strong predictors of bicycle volume, which aligns with findings from previous studies using bicycle traffic counts (Hankey & Lindsey, 2016) and bicycle commuting mode share (Buehler & Pucher, 2012; Buehler & Dill, 2016). Compared to a street segment without a facility, having an on-street bicycle facility is associated with a 43.8 to 54.2% increase in bicycle traffic, while having an off-street bicycle facility is associated with a 77.3 to 95.9% increase in bicycle traffic. The presence of entertainment venues, water bodies, zero-car households, and high bicycle commuting mode share are positively associated with higher bicycle volume.

Weather and socio-demographic variables were included as control variables. Some of these variables exhibit mixed effects (positive and negative) on bicycle traffic. In general, the result from this alternative set of models is similar to findings from the full models.

Table 9. Alternative models including bicycle facility as a predictor

Facility Models (Bicycle, Fall, Segment)	AM Peak	PM Peak
Multimodal NW Density	0.520 (3000)	0.057 (500)
Off-Street Bike Facility	0.773	0.959
On-Street Bike Facility	0.438	0.542
Temperature	-0.011	0.012
Precipitation		-0.219
Entertainment	0.003 (100)	
Water (hectare)	0.147 (200)	0.131 (200)
Zero-car household (thousand)	0.125 (750)	
Bicycle Mode share	18.015 (1250)	10.077 (1000)
Transit Mode share	-2.875 (400)	
Transit Stops	-0.002 (3000)	
Below 18 years old		-0.014 (100)
45-65 years old	-0.026 (100)	
Above 65 years old	-0.017 (750)	-0.034 (1000)
Income (thousand dollars)	0.004 (3000)	-0.005 (1250)
N	878	2,980
Adj-R2	0.51	0.46
MSAs included	Blacksburg, Los Angeles, Madison, Philadelphia	Blacksburg, Cleveland, Columbus, Lawrence, Los Angeles, Madison, Minneapolis, Philadelphia, San Francisco

Note: Buffer sizes are shown in parentheses. All dependent variables were log-transformed. All variables were significant at 0.05 level.

Alternative Models with Spatially Concentrated Count Locations

To explore whether the number of count locations within a jurisdiction impacted model performance, we developed another set of alternative models for jurisdictions that have 100 or more count locations. A list of MSAs included in these models is shown in Table 10; regression results are shown in Table 11.

Table 10. List of MSAs with high number of count locations

	AM Peak		PM Peak	
	Segment	Intersection	Segment	Intersection
Bicycle Models	Blacksburg (101)	Los Angeles (462)	Blacksburg (101)	Los Angeles (461)
	Champaign – Urbana (255)	Seattle (303)	Champaign – Urbana (255)	Madison (144)
	Columbus (213)	Tucson (1,058)	Lawrence (100)	Manhattan (112)
	Los Angeles (520)		Los Angeles (428)	Seattle (305)
	Philadelphia (198)		Minneapolis (950)	San Francisco (308)
			Philadelphia (190)	St Louis (142)
			San Francisco (1,241)	Tucson (1,060)
Pedestrian Models	Columbus (213)	Champaign – Urbana (121)	Lawrence (100)	Champaign – Urbana (121)
	New York City (1,022)	Seattle (305)	Minneapolis (950)	Manhattan (112)
	Philadelphia (158)	Tucson (1,075)	New York City (1,022)	Seattle (305)
			Philadelphia (162)	St Louis (238)
				Tucson (1,074)

Note: The number of observations (i.e., traffic counts) for each city are displayed in parentheses.

Adjusted-R^2 and the variables that entered the spatially dense models are similar to those of the full models presented above. In general, for the land-use and transportation variables, the sign and magnitude of the coefficients are similar, except for retail (direction of effect reversed) and households (substantial change in magnitude). As such, the models are robust for this sensitivity analysis and analysts may be more confident in use of these models for estimating traffic volumes at locations without counts in regions throughout the country. However, the marginally significant effects of household and retail variables should be interpreted with caution.

Table 11. Models to predict bicycle and pedestrian traffic in areas with spatially dense count locations

	Bicycle				Pedestrian			
	AM		PM		AM		PM	
Spatially Dense Models	Segment	Intersection	Segment	Intersection	Segment	Intersection	Segment	Intersection
Multimodal network density	0.472 (3000)		0.022 (100)	0.213 (2500)	-0.062 (1000)	0.181 (2000)	-0.078 (3000)	0.063 (100)
Total network density			0.032 (2000)	0.011 (1500)			-0.010 (2000)	
Intersection density					0.021 (3000)	0.006 (750)	-0.009 (3000)	0.010 (1250)
Water (hectare)	0.970 (100)	0.802 (100)	0.169 (200)	0.216 (200)	0.017 (500)	-0.003 (1000)	0.008 (1000)	0.007 (750)
Office (thousand)		-0.419 (300)			0.024 (300)	1.907 (100)	0.361 (100)	2.855 (100)
Industry (thousand)			0.034 (2000)	0.043 (1500)	0.030 (500)	-0.029 (3000)	0.040 (200)	-0.132 (750)
Retail (thousand)			-1.018 (100)					0.086 (3000)
Service (thousand)		0.077 (750)	0.0017 (750)					-0.036 (1500)
Entertainment (thousand)							-0.185 (100)	1.041 (300)
Housing Unit					0.047 (1000)			0.151 (1000)
Zero-car household (thousand)			0.037 (1000)			0.239 (1000)	0.002 (200)	
Household (thousand)		0.014 (3000)				-1.409 (200)	0.042 (1250)	
Bicycle mode share	6.371 (1000)	15.995 (3000)	8.601 (750)	11.142 (3000)				

	Bicycle				Pedestrian			
	AM		PM		AM		PM	
Spatially Dense Models	Segment	Intersection	Segment	Intersection	Segment	Intersection	Segment	Intersection
Walking mode share					4.305 (100)	3.784 (200)	3.829 (100)	4.239 (400)
Transit mode share			-2.980 (2000)	-4.562 (3000)	-5.388 (3000)	-3.520 (3000)	-13.282 (3000)	-4.705 (3000)
Transit stops			-0.003 (1500)		-0.003 (3000)	0.120 (100)	0.039 (200)	0.077 (100)
Precipitation			-0.076	0.192				
Temperature				0.01		0.013	0.01	
Income (thousand dollars)		0.008 (2500)	-0.004 (750)	0.007 (3000)	-0.014 (2500)	-0.004 (100)	-0.020 (3000)	-0.003 (100)
Under 18 years old			-0.017 (100)	-0.036 (3000)			0.019 (100)	
18-45 years old			0.005 (100)	0.017 (200)			0.012 (100)	
45-65 years old					0.030 (300)		0.011 (100)	-0.021 (1000)
Above 65 years old		0.012 (3000)	-0.029 (1250)				0.017 (2000)	-0.008 (100)
Number of MSAs	5	3	7	7	3	3	4	5
N	1,185	1,760	2,963	2,433	1,377	1,133	2,209	1,533
Adj-R2	0.29	0.39	0.38	0.57	0.6	0.46	0.71	0.54

Note: Buffer sizes are shown in parentheses. All dependent variables were log-transformed. All variables were significant at 0.05 level.

Strengths and Limitations

Our study has several limitations pertaining to data availability and modeling methods. In terms of data, some predictors of bicycle and pedestrian traffic, as shown in previous literature, such as slope, vehicular traffic (AADT), speed limit, park and open space, etc. were unavailable and thus were not included in the model. In the future, we will refine the models by including some of these variables as they become available. A general challenge when developing national-scale models is that refined land-use data vary across jurisdictions; future efforts to develop more specific land-use patterns in a consistent format across the country would benefit the type of modeling described here. Similarly, methods for counting bicycles and pedestrians vary across the country; work to make bicycle and pedestrian traffic counts more consistent across jurisdictions would be useful for spatial modeling.

With regards to modeling methods, we have not yet accounted for temporal and spatial dependence (e.g., autocorrelation). For example, count locations that are close in proximity to each other may exhibit spatial autocorrelation. Similarly, locations where counts were collected multiple times over the years could potentially exhibit temporal dependence when modeling counts at these locations. We aim to address these issues in future work.

Despite these limitations, our work adds to the growing body of work on direct-demand modeling in several ways. A key contribution of our work is the development of models that do not rely on data from a single city. Previous direct-demand models were not able to transfer to other jurisdictions making their usefulness limited for practitioners outside of the study area. Our models rely on count data from 20 MSAs across the country and therefore may be better suited for estimating traffic patterns in locations with few or no bicycle and pedestrian counts. Expanding on our work (and addressing some of the limitations above) would allow for more reliable models that could be applied across the country.

Conclusions and Recommendations

This study examined the relationship between the built environment and non-motorized traffic volumes in 20 US Metropolitan Statistical Areas. Our models showed reasonable goodness of fit for both bicycle traffic (adjusted R^2: 0.19 – 0.56) and pedestrian traffic (adjusted R^2: 0.45 – 0.72). We found a number of land-use and network variables that were correlated with bicycle and pedestrian traffic, such as multimodal network density, presence of water bodies, offices, industry, zero-car household, as well as bicycle and walking commuting mode shares. Intersection density is also a strong predictor for pedestrian volume, while off-street and on-street bicycle facilities are strong predictors of bicycle volume. Despite data and modeling limitations, our models produced robust outcomes, which were validated using a cross validation and sub-sampling method (i.e., modeling a subset of MSAs with spatially dense traffic count locations). However, some mixed results (variables with positive and negative coefficients across models) for household density, housing unit, retail, and service land use indicate that practitioners should be cautious when using these land use types as predictors for bicycle and pedestrian traffic.

A strength of our approach is that our count data have good spatial and temporal coverage across a variety of cities and regions in the US (most previous studies rely on single-city models). The coverage of the count data combined with the use of predictor variables that are available at a national-scale allows for estimating non-motorized traffic in cities where counts are inadequate or unavailable (a limitation of previous work). Our models could be used to inform decisions on where to locate non-motorized transportation facilities and to assess exposure to accidents with motor vehicles or other environmental hazards.

References

Akita, Y. (2014) LUR Tools. The University of North Carolina at Chapel Hill. Retrieved from http://www.unc.edu/~akita/lurtools/index.html on May 25, 2017.

Buehler, R., & Pucher, J. (2012). Cycling to work in 90 large American cities: new evidence on the role of bike paths and lanes. *Transportation*, *39*(2), 409–432. https://doi.org/10.1007/s11116-011-9355-8

Buehler, R., & Dill, J. (2016). Bikeway Networks: A Review of Effects on Cycling. *Transport Reviews*, *36*(1), 9–27. https://doi.org/10.1080/01441647.2015.1069908

Fagnant, D. J., & Kockelman, K. (2016). A direct-demand model for bicycle counts: the impacts of level of service and other factors. *Environment and Planning B: Planning and Design*, *43*(1), 93–107. https://doi.org/10.1177/0265813515602568

Jackson, R., Dannenberg, A., & Frumkin, H. (2013). Health and the built environment: 10 years after. *American Journal of Public Health, 103*(9), 1542-1544. doi:10.2105/AJPH.2013.301482

Hankey, S., Lindsey, G., Wang, X., Borah, J., Hoff, K., Utecht, B., & Xu, Z. (2012). Estimating use of non-motorized infrastructure: Models of bicycle and pedestrian traffic in Minneapolis, MN. *Landscape and Urban Planning*, *107*(3), 307–316. https://doi.org/10.1016/j.landurbplan.2012.06.005

Hankey, S., & Lindsey, G. (2016). Facility-demand models of peak period pedestrian and bicycle traffic: Comparison of fully specified and reduced-form models. *Transportation Research Record, 2586*(2586), 48-58. doi:10.3141/2586-06

Hankey, S., Lindsey, G., & Marshall, J. D. (2016). Population-level exposure to particulate air pollution during active travel: Planning for low-exposure, health-promoting cities. *Environmental Health Perspectives, 125*(4), 527-534. doi:10.1289/EHP442

Miranda-Moreno, L. F., & Fernandes, D. (2011). Modeling of pedestrian activity at signalized intersections. *Transportation Research Record,* (2264), 74-82. doi:10.3141/2264.09

Pulugurtha, S., & Repaka, S. (2008). Assessment of Models to Measure Pedestrian Activity at Signalized Intersections. *Transportation Research Record: Journal of the Transportation Research Board*, *2073*, 39–48. https://doi.org/10.3141/2073-05

Schneider, R., Arnold, L., & Ragland, D. (2009a). Methodology for Counting Pedestrians at Intersections: Use of Automated Counters to Extrapolate Weekly Volumes from Short Manual Counts. *Transportation Research Record: Journal of the Transportation Research Board*, *2140*, 1–12. https://doi.org/10.3141/2140-01

Schneider, R., Arnold, L., & Ragland, D. (2009b). Pilot Model for Estimating Pedestrian Intersection Crossing Volumes. *Transportation Research Record: Journal of the Transportation Research Board*, *2140*, 13–26. https://doi.org/10.3141/2140-02

Tabeshian, M., & Kattan, L. (2014). Modeling non-motorized travel demand at intersections in calgary, canada use of traffic counts and geographic information system data. *Transportation Research Record, 2*(2430), 38-46. doi:10.3141/2430-05