THE JOURNAL OF TRANSPORT AND LAND USE http://jtlu.org
VOL. 6 NO. 2 [2013] pp. 47–58 doi: 10.5198/jtlu.v6i2.296

Spatial modeling of bicycle activity at signalized intersections

Jillian Strauss [a]
McGill University

Luis F Miranda-Moreno [b]
McGill University

Abstract: This paper presents a methodology to investigate the link between bicycle activity and built environment, road and transit network characteristics, and bicycle facilities while also accounting for spatial autocorrelation between intersections. The methodology includes the normalization of manual cyclist counts to average seasonal daily volumes (ASDV), taking into account temporal variations and using hourly, daily, and monthly expansion factors obtained from automatic bicycle count data. To correct for weather conditions, two approaches were used. In the first approach, a relative weather ridership model was generated using the automatic bicycle count and weather data. In the second approach, weather variables were introduced directly into the model. For each approach, the effects of built environment, road and transit characteristics, and bicycle facilities on cyclist volumes were determined. It was found that employment, schools, metro stations, bus stops, parks, land mix, mean income, bicycle facility type (bicycle lanes and cycle tracks), length of bicycle facilities, average street length, and presence of parking entrances were associated with bicycle activity. From these, it was found that the main factors associated with bicycle activity were land-use mix, cycle track presence, and employment density. For instance, intersections with cycle tracks have on average 61 percent more cyclists than intersections without. An increase of 10 percent in land-use mix or employment density would cause an increase of 8 percent or 5.3 percent, respectively, in bicycle flows. The methods and results proposed in this research are helpful for planning bicycle facilities and analyzing cyclist safety. Limitations and future work are discussed at the end of this paper.

1 Introduction

Active transportation is an essential component of every city's transportation system, whether it is used as the only mode or in combination with other modes, providing access to or from the transit network. The use of the bicycle as a mode choice is currently on the rise in some Canadian cities, such as Montreal (Statistics Canada 2003; Miranda-Moreno and Nosal 2011; Vélo Québec 2005). The many recognized benefits of cycling as well as the continuously increasing bicycle activity give rise to some practical and research questions related to the built environment, bicycle facilities, and data collection techniques. Among these are three key issues:

1. What is the link between built environment (BE), bicycle facilities, and road designs on bicycle activity at the microlevel (e.g., at intersections)? Intuitively, one would expect to witness higher cyclist concentrations at intersections with bicycle facilities and appropriate designs, but what is the magnitude of the impact of different facility designs (e.g., bicycle lanes versus cycle tracks)? Knowledge of the factors that increase or decrease bicycle activity, referred to here also as bicycle flows or volumes at an intersection, is an essential component in the planning, operation, and design of bicycle facilities, road safety analysis, etc. Transportation engineers and planners are interested in estimating the impact of built-environment changes and the impact that installing new bicycle facilities has at intersections.

2. Is it possible to estimate bicycle activity based on statistical methods? Bicycle activity is a key variable in safety analysis at intersections. Bicycle flows are required to provide a complete definition of risk exposure measures. However, many agencies refrain from collecting such data since carrying out manual cyclist counts is expensive and time consuming. To try to solve this issue, previous studies have proposed cyclist and pedestrian activity estimation methods to quantify volumes at intersections with missing data. Among the proposed approaches are the models developed by Haynes and Andrzejewski (2010), Jones et al. (2010) and Griswold et al. (2011). These models apply linear or log-linear regression to evaluate cyclist and pedestrian volumes as functions of built environment, road and transit characteristics, and other variables.

[a] jillian.strauss@mail.mcgill.ca
[b] luis.miranda-moreno@mcgill.ca

3. How can manual bicycle counts be standardized to obtain average annual daily volumes accounting for temporal patterns and weather effects? At sites where bicycle counts are available, such counts are often only provided for short periods of time. A method to standardize a few hours of manual counts taken any time during the day is required since such data is a necessary input for safety studies and for transportation agencies and governments to guide the design, location, and allocation of resources for new bicycle facilities.

Despite these recent developments, previous work on cyclist activity models have not used normalized average seasonal daily flows correcting for the potential effects of weather conditions on manual bicycle counts. One of the few studies using average daily volumes is that done by Schneider et al. (2009); however, counts were not corrected for weather conditions. Moreover, most studies have employed simple statistical methods that do not correct for spatial autocorrelation. In addition, due to the data limitations, small to moderate sample sizes have been used in previous research. Accordingly, this paper has two main objectives:

1. Propose a methodology to normalize manual cyclist counts at intersections to average seasonal bicycle flows correcting for temporal trends and weather effects.

2. Develop spatial bicycle ridership models to investigate the link between cyclist volumes and built environment, road and transit network attributes, and bicycle facilities. These models account for the presence of spatial autocorrelation among intersections.

This paper is broken down into several sections. Section 2 offers a literature review. Section 3 presents the modeling framework. Section 4 reveals the case study for this paper and describes the different sources of data used as well as the models and results, and Section 5 presents the conclusions drawn from this study and directions for future work.

2 Literature review

Research looking into different aspects of non-motorized transportation is growing in interest. Among other research topics, several studies have focused on the impact of socio-demographic and built-environment factors as well as the presence of bicycle infrastructure on bicycle usage. In many cases, this has been studied using survey data and travel-behavior models (Dill and Carr 2003; Dill and Gliebe 2008; Garrard et al. 2008; Hillman 1993; Kim et al. 2007; Krizek et al. 2009; McCahill and Garrick 2008; Moudon et al. 2005; Pucher and Buehler 2006; Pucher et al. 2010; Stinson and Bhat 2003; Xing et al. 2010). An extensive study carried out in Dutch, Danish, and German cities found that providing separate bicycle facilities along popular roads as well as implementing traffic-calming measures were a necessity to promoting convenience and safety for cyclists (Pucher and Buehler 2008). Xing et al. (2010) also concluded that improving and expanding the bicycle network as well as making changes to the physical environment, mainly encouraging mixed land use, can have a positive effect on bicycle activity in a given area. Krizek et al. (2009) were interested in comparing bicycle activity in 1990 to that in 2000 as a function of the presence and proximity of bicycle facilities at both on- and off-street facilities. Overall it was found that areas closer to facilities witness larger cyclist volumes. The study done by Moudon et al. (2005) found similar results; providing exclusive bicycle facilities can yield an increase in cyclist numbers. Concerning this result, one cannot infer causality. The presence of the facility can be either the cause or the effect of elevated cyclist flows, and therefore further work is required to determine causality.

Three recent studies in California have reported bicycle activity models linking observed bicycle volumes to built-environment factors, road characteristic, and socio-demographic attributes. These studies developed regression models without accounting for the potential presence of spatial autocorrelation. In the city of Santa Monica, afternoon bus frequency, land-use mix, density of residents under the age of 18, and proximity to the bicycle network were used as variables to predict weekday afternoon peak-hour bicycle volumes through intersections (Haynes and Andrzejewski 2010). Another study in San Diego calibrated a log-linear regression model using employment density and the length of nearby multi-use trails to predict weekday 7 a.m. to 9 a.m. bicycle volumes (Jones et al. 2010). A different study carried out in Alameda County developed models based on two-hour bicycle counts performed at a sample of 81 intersections in 2008 and 2009 during spring (Griswold et al. 2011). The explanatory variables were generated by extracting land use, transportation system, and socio-demographic characteristics within different radial distances from the intersections. The models tested had adjusted R-squared values ranging from 0.39 to 0.60. These models showed that bicycle volumes tended to be higher at intersections surrounded by more commercial retail properties within 0.1 mile, closer to a major university, with a marked bicycle facility on at least one leg of the intersection, surrounded by less hilly terrain within 0.5 mile, and surrounded by a more connected roadway network.

In Cambridge, Massachusetts, a space syntax method was applied to model cyclist activity (McCahill and Garrick 2008). This method is a tool capable of assessing the quality of existing bicycle infrastructure as well as the significance of that facility within the entire network. With only a small sample of intersections, the model still performed well and identified population and employment density as two variables explaining most of the variation in bicycle flows. Also in the United States, Stinson and Bhat (2003) carried out a stated preference survey and identified that cyclists prefer, among others things: shorter travel times, local roads instead of arterials, bicycle facilities, streets where parallel parking is prohibited, and fewer stop signs and intersections to cross.

Griswold et al. (2011) provide a summary of the previous bicycle studies and the BE, road and bicycle network, and socio-demographic factors associated with bicycle activity. To mention a few, population and employment density as well as the presence of bicycle facilities, bus frequency, and smooth pavement all have a positive effect on cycling. On the other hand, some of the factors that would induce lower bicycle volumes are: slope, motor-vehicle volume, parallel parking, and major roads.

Despite these previous studies, methods for obtaining and studying bicycle volumes at and along different facilities are only now beginning to gain popularity, and more work is still required. In general, past studies (mainly cross-sectional) have been based on relatively small samples of intersections, due to being restricted to locations with counts. Also, past studies have not used normalized average daily counts, using hourly, daily, and monthly expansion factors and have not corrected for the effect of weather when the manual counts were taken. In addition, spatial autocorrelation has not yet been addressed.

3 Modeling framework

Consider C_i as the outcome of interest representing the number of cyclists observed at a certain intersection *i* during a given period of time (e.g., one day or several hours). A proportion of the cyclists (trips) have their origins or destinations in the proximity of a given intersection *i*; however, other cyclists have their origins or destinations out of the proximity of the intersection—cyclists merely passing through.

The attraction or production of cyclist trips in the area surrounding an intersection should be highly correlated with the land-use and road-network characteristics as well as demographic factors. For instance, as residential or commercial densities increase, the production and attraction of cyclist trips in this area are expected to be higher. On the other hand, bicycle flows through an intersection that neither originate nor end within its vicinity should be associated with road connectivity as well as the presence and types of bicycle facilities. In addition, there are other factors (observed and unobserved) such as travel demand patterns (hourly, daily, and monthly) and weather or seasonal conditions that could have an important affect on observed flows (C_i).

Then, to identify the main factors associated with bicycle volumes and their order of magnitudes, a spatial lag regression model is formulated. This is given by Equation 1.

$$ln(C_i) = \alpha+\beta LU_i + \gamma UF \,\&\, D_i + \phi GD_i + \tau CF_i + \eta W_i + \rho \omega C_j + \varepsilon_i \quad (1)$$

where *i* stands for the intersection *i* (*i=1,...,n*) and *j* stands for all intersections in the vicinity of *i*.

LU_i= Land-use characteristics. These variables include residential, commercial, governmental, industrial, and parks, and recreational.

$UF \,\&\, D_i$ = Urban form and demographics. These attributes cover area demographics such as population, employment, income levels, and presence of schools as well as road and transit characteristics, such as the presence of bus stops and length of bus routes, presence of metro stations, number of intersections, portion of major roads, etc.

GD_i = Geometric design at the intersection. This includes the presence of a median, presence of one-way approaches, number of approaches, number of lanes, and typology of intersecting streets.

CF_i = Cyclist infrastructure such as presence and length of bicycle lanes or cycle tracks.

W_i= Hourly weather conditions (temperature, humidity, and precipitation) during the period when the manual counts were taken. Data from the same weather station was used for the entire area.

$\omega C_j = \sum_{j=1}^{n} \omega_{ij} C_j$, where ωC represents the influence (spillover effect) of the bicycle activity at neighboring intersections. This depends on the bicycle activity at intersections *j* (with $j \neq i$) and the inverse of the distance from *i* to *j*, represented by ωj. Cyclists observed at a given intersection can be divided in those starting or ending a trip in its vicinity or those passing through the intersection. The spatial influence of neighboring intersections can help to capture this effect.

ε_i = Independent error term representing unobserved heterogeneities.

$\alpha, \beta, \lambda, \phi, \tau, \eta, \rho$ = Parameters to be estimated from the data.

This model stipulates that bicycle activity at intersection *i* also depends on the bicycle activity at surrounding intersections *j*. The expected bicycle activity at intersection *i* no longer depends solely on microlevel factors at *i* but also on a spatial correlation component that depends on the level of bicycle activity among neighboring intersections *j*. A proportion of cyclist trips at a given intersection start or end somewhere else in the city, and therefore part of the cyclist traffic at a given intersection *i* is simply through traffic.

To represent spatial autocorrelation, two types of proximity measures have been proposed in the literature: 1) adjacency-based and 2) distance-based (Drukker et al. 2011). An adjacency-based measure specifies $\omega_j = 1$ if intersections *i* and *j* are neighboring sites and 0 otherwise. For instance, the neighboring intersections of location *i* can be those that are located in the same administrative area (e.g., in the same borough). Alternatively there are more general weight functions based on the distances between intersections that have also been suggested for ω_j. This paper employs the more general distance-based method where the entries in the distance-based vector, ω_j, acquire non-zero values for all intersections *j* that are defined as connected to *i*. Applying one of these weight matrices to the model creates what is known as a spatial lag model. To take into account the effects that nearby intersections have on the bicycle activity through intersection *i*, an inverse distance matrix needs to be created. This matrix can be generated using the X and Y coordinates of the adjacent intersections. This matrix, ω_j, is $n \times n$ and each entry represents the inverse of the distance between intersections *i* and *j*, $1/d_{i,j}^{\alpha}$, where α is the scale parameter. Different values of α can then be tested including 0.8, 1.0, 1.2.

To estimate the model parameters, different steps and sources of data are needed including average daily cyclist volumes, built environment, intersection design attributes, and information on bicycle infrastructure. The modeling framework proposed in this work is illustrated using a rich dataset that has been assembled for the island of Montreal.

4 Case study: The island of Montreal

The island of Montreal, Quebec, Canada, was used as the application environment. In this study, a sample of 758 intersections is used, and a significant proportion of these intersections are located in the central neighborhoods of the island. This sample represents more than one-third of the total number of signalized intersections on the island and considers a range of land-use and urban-form attributes to capture the factors affecting bicycle activity. These specific intersections were selected since they meet the following criteria: 1) recent cyclist counts were available (mostly obtained in 2009 by the Montreal Department of Transportation), and 2) counts were done during the cycling season from April 1 to November 30, when bicycle facilities are open. See Figure 1 for the locations of these intersections and their respective annual average daily bicycle flows (after being adjusted for temporal trends as discussed later in the document).

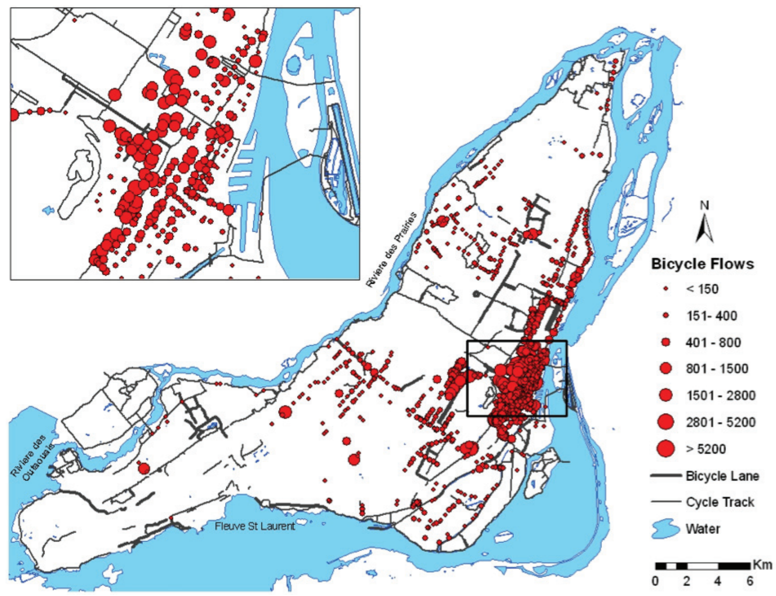
4.1 Built environment and geometric design

Several geometric design and built-environment characteristics have been identified for each intersection individually using Geographic Information Systems (GIS), namely ArcMap and Google Street View and through field visits. The geometric design variables that were included in the analysis are:

- Number of approaches per intersection
- Presence of a median in one or all approaches
- Presence of one-way streets in one or all approaches
- Maximum speed limit of all approaches in an intersection
- Typology of the intersecting streets (combinations of arterial, collector, and local streets)
- Presence of bicycle facilities, bicycle lanes, or cycle tracks, within 15 meters of the intersection

Motor-vehicle flows passing through each intersection aggregated by movement type, left turning, right turning, and through moving motor-vehicle flows.

Figure 1: Intersections and their respective bicycle flows.



To generate the land-use, urban-form and bicycle-facility characteristics, four buffer dimensions, 50 meters, 150 meters, 400 meters and 800 meters, were defined for each intersection. Different buffer dimensions were used to test the impact that certain features such as land use, demographics, and transit and bicycle-facility network availability have on bicycle activity. As mentioned previously, the majority of the intersections are located in the central neighborhoods of Montreal, which is characterized as being a dense area with a good land-use mix as well as having good transit coverage. This reality justifies the use of small buffer sizes. The 50-meter buffer serves to determine how the area in the immediate vicinity of the intersection affects bicycle activity. The 150-meter buffer scans a bit farther past the intersection but is still quite close to the intersection itself. The 400-meter buffer reaches even farther and covers an area of short travel distance from the intersection. The 800-meter buffer expands farther past the intersection while still presenting a fairly short travel distance by bicycle. The land-use and urban-form characteristics came from two sources, Statistics Canada and DMTI Spatial Inc. Census tract data, demographics, and road and transit network data were all intersected with the buffers of differing dimensions. These categories were then decomposed into the specific data needed for this analysis, further described in Table 1. In addition to several road and geometric design characteristics, bicycle infrastructure attributes in the vicinity of the intersections were incorporated in the analysis such as the presence of a bicycle lane or cycle track in one or more of the approaches and the total length of bicycle facilities within the buffer. In this paper, bicycle lane is the term used to describe any bicycle facility that is not separated from the roadway by any physical barrier and often just has painted lines or sharrows. Cycle tracks often have concrete medians or bollards separating them from car traffic. Bicycle facilities slightly offset from the road are also considered cycle tracks.

An additional variable, land mix, was generated for each buffer based on the different land-use types present. Land mix is modeled after an entropy index, which measures the level of homogeneity or mix in a given buffer (Kockelman 1997), achieved by applying Equation 2,

$$E_j = -\Sigma_{i=1}^{n} \frac{\frac{A_{ij}}{D_j} \ln\left(\frac{A_{ij}}{D_j}\right)}{\ln(n)} \qquad (2)$$

where E_j = land mix entropy index, A_{ij} = the area of land use i in buffer j, D_j = area of buffer j excluding open space and water body, and n = the number of land-use types. In this analysis, n is 5, representing residential, commercial, industrial, governmental, and parks and recreational. The value of the land-mix index ranges from 0 to 1, representing the case of a homogenous buffer, only one land-use type present, to a well-mixed and diverse buffer, respectively.

For population and employment variables, Equation 3 was applied to approximate the value of those variables within the buffer,

$$X_i = \Sigma_j \frac{A_{ji}}{A_j} * X_j \qquad (3)$$

where X_i = population or number of jobs within buffer i, X_j = population or number of jobs in the census block j, A_{ji} = area of the census block inside the buffer i and A_j = area of census block j.

4.2 Daily bicycle volumes

Manual and automatic cyclist counts were combined to obtain average seasonal daily flows taking into account temporal trends. Manual bicycle counts used in this study were collected by the city of Montreal on a weekday during 2008 and 2009 over an eight-hour period. These eight hours include morning peak from 6 a.m. to 9 a.m., noon period from 11 a.m. to 1 p.m.. and evening peak from 3:30 p.m. to 6:30 p.m. However, these counts were not all collected on the same date, as is usually done in traffic studies. To normalize these flows, average seasonal daily bicycle flows were computed using expansion factors estimated from permanent automatic bicycle count stations. This was accomplished following the standard procedure used in traffic engineering (World Road Association 2003).

To convert hourly flows into annual average daily flows, automatic bicycle count data were used. Automatic counts were collected for long periods of continuous time (years) from automatic bicycle counters, loop detectors (permanent count stations), located in specific areas along five of Montreal's bicycle facilities running alongside specific streets. These counters are located: 1) along de Brébeuf between Rachel and Marie-Anne, 2) along Berri between de Maisonneuve and Ontario, 3) along de Maisonneuve between Berri and Saint Denis, 4) along de Maisonneuve between Peel and Stanley, and 5) along Saint Urbain between Mont-Royal and Villeneuve. These counters are continuously obtaining bicycle count data; however, for the purpose of this study, the dataset was filtered to only include the months of interest for which Montreal's bicycle facilities are open (April 1 to November 30). Data is provided in 15-minute intervals for both directions, but for the sake of this study, these flows have been aggregated per hour and path (direction of travel is irrelevant). Based on the automatic count data, expansion factors were obtained to correct for the hour, day, and month in which the manual counts were recorded.

Table 1: Summary statistics.

Category	Variable	Units	Buffer: 50 m		Buffer: 150 m		Buffer: 400 m		Buffer 800 m	
			Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Land Use	Residential		3.58	2.38	35.10	19.53	301.44	2766.61	953.52	349.29
	Commercial		0.79	1.34	5.52	9.91	22.96	49.27	112.52	131.20
	Governmental	m^2	0.54	1.20	5.65	10.04	33.36	47.13	181.00	152.54
	Industrial	(1000s)	1.33	1.85	13.69	16.72	86.01	105.58	413.60	335.42
	Park and Recreational		0.31	0.94	3.05	6.95	22.74	33.68	125.68	122.42
	Open Space		1.28	1.71	7.56	11.10	31.54	47.73	129.95	121.65
	Land-Mix Index	value 0 to 1	0.342	0.242	0.429	0.227	0.515	0.194	0.661	0.148
Demographics	Population	Count	0.052	0.039	0.469	0.337	3.221	1.922	12.202	6.034
	Workers	(1000s)	0.024	0.019	0.218	0.164	1.497	0.940	5.663	2.979
	Median Income		42.47	21.03	42.52	20.50	42.91	18.58	44.71	16.91
	Average Income	$ (1000s)	57.93	33.91	58.14	32.99	59.09	30.00	62.52	28.27
			0.021	0.161	0.145	0.439	1.033	1.228	4.169	2.901
	Number of Schools	Count								
Transit	Number of Metro Stops	0, 1	0.008	0.089	0.073	0.279	0.406	0.673	1.418	1.793
	Number of Bus Stops	Count	0.021	0.161	0.145	0.439	1.033	1.228	4.169	2.901
Road	Number of Intersections	Count	1.94	1.36	7.58	4.57	46.34	22.28	168.80	71.50
	Average Street Length		0.12	0.05	0.25	0.12	0.43	0.14	1.66	4.47
	Length of Bicycle Facilities	km	0.022	0.044	0.099	0.167	0.583	0.639	2.072	1.721
Intersection	Number of Approaches*	0,1	0.77	0.42	Attributes of the intersections are independent of buffer sizes (* 0 if three-leg intersection and 1 if four or more legs)					
	Presence of a Median		0.46	0.50						
	Maximum Posted Speed Limit	km/h	61.57	10.01						
	Presence of One-Way Approach		0.74	0.44						
	Presence of Parking Entrance		0.15	0.35						
	Presence of Bicycle Lane	0, 1	0.09	0.29						
	Presence of Cycle Track		0.13	0.33						
Weather	Temperature	°C	13.93	6.81	Weather conditions are independent of buffer sizes					
	Humidity	%	63.74	17.40						
	Presence of Precipitation	0, 1	0.27	0.44						
Bicycle Flow	Adjusted for Weather-Hourly		283.8	624.4	Dependent Variables					
	Adjusted for Weather-Daily	Count	276.8	587.9						
	Weather as Variable		208.9	447.7						

a. Hourly, daily, and monthly expansion factors

Using the automatic count dataset, expansion factors for hour, day, and month were obtained, independent of weather. Since the vast majority of the counts used were obtained in 2009 and only during weekdays, only the 2009 weekday data are used to generate the hourly, daily, and monthly factors. The process of developing expansion factors is in accordance with the one used for developing motor-vehicle expansion factors (World Road Association 2003). Hourly factors were obtained by computing the average bicycle flows per hour. The hourly factors were then calculated by dividing each individual hourly average by the average of the averages for all 24 hours. The first step in obtaining both the daily and monthly factors was aggregating the full 24-hour flows for each day. The daily factors were then calculated by dividing the average individual daily values by the average of the averages for the five weekdays. The monthly factors were obtained in the same way by dividing the average individual monthly values by the average of the averages for all eight months (April to November). Using expansion factors, average seasonal daily volumes (ASDV) were obtained by dividing the eight-hour flows by the hourly, daily, and monthly factors. These factors have been reported in Strauss and Miranda-Moreno (2011). The eight hours for which we have bicycle counts account for about 51 percent of the daily ridership. Looking at the days of the week, Monday and Friday experience slightly fewer bicycles than Tuesday, Wednesday, and Thursday, which witness roughly the same ridership. In terms of monthly ridership, April has a fairly low ridership, which continuously increases from May to September and then drops even lower in October and November.

b. Weather correction method

Previous studies have addressed the sensitivity of weather conditions on bicycle activity (Brandenberg et al. 2007; Nankervis 1999a; Nankervis 1999b; Thomas et al. 2009; Winters et al. 2007; Richardson 2000; Richardson 2006). Counts taken in April for instance, during the same hour and the same day of the week can be very different depending on weather conditions such as temperature, humidity, and precipitation. To take into account the effects of weather conditions on the observed bicycle flows, two approaches were used:

Approach 1 consists of the development and application of a relative weather model generated from automatic count data. Automatic hourly bicycle counts were matched to their corresponding hourly weather conditions and a relative ridership model was calibrated—for more detail, we refer to Miranda-Moreno and Nosal (2011). The weather variables in this dataset, temperature, relative humidity, and precipitation were applied to generate a model to determine the percent increase or decrease in hourly ridership due to relative changes in weather conditions from their mean values for the same path, month, year, day of the week, and hour. The relative changes (in percentage, represented by ΔR) were measured as the deviation of absolute ridership (represented by R) from the average hourly value for the same bicycle facility, month, and day of the week (represented as $\bar{R}$) as expressed in Equation 4. A similar procedure was carried out to generate a daily weather model; however, in this case the total ridership for the day was compared to the average ridership for that same path, month, year, and day.

$$\Delta R = \frac{(R - \bar{R})}{\bar{R}} \tag{4}$$

The relative ridership model to account for weather effects on cycling was then defined as shown in Equation 5. The impact of certain extreme weather conditions, such as very warm (temperatures greater than 32° C), very cold (lower than minus 20° C), a combination of warm temperatures and high humidity, and the presence of rain within three hours prior to the count as well as other lag effects were tested. In accordance with Nankervis (1999a), morning weather conditions can affect a cyclist's decision to bike or not. Using the automatic count data, complex models were generated containing temperature, humidity, precipitation, and other variables, where ΔR = Deviation in percentage of ridership due to the relative changes in temperature, humidity, and amount of precipitation from the normal conditions for that month and year, $\Delta Temperature$ = Relative change in temperature from the average monthly conditions for that year, $\Delta Humidity$ = Relative change in humidity from the average monthly conditions for that year, *Rain1* = Less than 15 millimeters of rain per hour or less than 45 millimeters of rain during the day, *Rain2* = Between 15 and 50 millimeters of rain during the hour or between 45 and 100 millimeters of rain during the day, and *Rain3* = Greater than 50 millimeters of rain during the hour or greater than 100 millimeters of rain during the day.

These rain values represent low, moderate, and high amounts of precipitation and the range of values is based on the National Weather Service's convention that has been adjusted to suit Montreal (Nosal and Miranda-Moreno 2012).

β_0, β_1, β_2, β_3, β_4, and β_5 are model parameters to be estimated from the automatic count data.

Using the relative ridership model, the observed manual hourly bicycle counts at each intersection were adjusted for weather. Then, applying the expansion factors, manual hourly flows were converted to average seasonal daily flows.

Approach 2 directly introduces weather conditions into the bicycle activity model. In this case, temperature, humidity, and precipitation, observed at the time of the count, obtained from the same source as in Approach 1, are input into the bicycle activity model.

The hourly weather conditions during the times in which counts were taken were obtained from Environment Canada's National Climate Data (ECNCD) Information Archive. This paper controls for weather conditions but does not focus on estimating its impact. The McGill University weather station was used to obtain data regarding temperature, relative humidity, and amount of precipitation. This station is located within four kilometers of all the intersections, which are mostly located in the central neighborhoods of the island of Montreal. The Montreal Pierre Elliot Trudeau Airport weather station was used when the McTavish data was missing.

$$\Delta R = \beta_0 + \beta_1\, \Delta Temperature + \beta_2\, \Delta Humidity + \beta_3\, Rain1 + \beta_4\, Rain2 + \beta_5\, Rain3 \tag{5}$$

Table 2: Hourly and daily weather model results.

Variables	Hourly Model			Daily Model		
	Coef.	P-value	Elasticity*	Coef.	P-value	Elasticity*
Temperature	0.235	0.000	2.4	0.288	0.000	2.9
Humidity	-0.657	0.000	-6.6	-0.643	0.000	-6.4
Rain 1	-13.459	0.000	-134.6	-6.197	0.000	-62.0
Rain 2	-22.892	0.000	-228.9	-17.829	0.000	-178.3
Rain 3	-24.041	0.000	-240.4	-24.646	0.000	-246.5
Constant	0.949			4.160		
R-Squared	0.390			0.554		

*Elasticities are expressed in terms of a 10 percent change in the independent variable.

After model calibration, the parameters in Equation 5, for both the hourly and daily models can be seen in Table 2. These parameters reveal that cyclists prefer cycling in warm temperature conditions. A 10 percent increase in temperature is expected to increase cyclist flows by more than 2 percent. Humidity and precipitation have the opposite effect of temperature. A rise in humidity by 10 percent would cause ridership to decrease by about 6.5 percent. As anticipated, cyclists do not like biking in the presence of rain, and cyclist numbers are expected to drop by an increasing amount the more rain there is.

4.3 Bicycle activity model

Bicycle activity is a positively skewed count variable, and therefore both log-linear and negative binomial regression model settings are tested. Equation 1 shows the log-linear formulation with the natural log of bicycle activity, *ln(Ci)*, as the dependent variable. This analysis was carried out separately for the three dependent variables of interest: bicycle flows adjusted for weather conditions using the hourly model, bicycle flows adjusted for weather conditions using the daily model, and bicycle flows with weather conditions as variables in the model. This study was carried out at an intersection level to capture the specific design attributes of the intersection, such as presence of a median, number of approaches, typology of the intersecting streets, and the presence of a bicycle facility, as well as the characteristics of the area surrounding each intersection. Again, to test the sensitivity and impact of the buffer dimensions, different buffer sizes were used: 50 meters, 150 meters, 400 meters, and 800 meters. Note that when using model Approach 2, weather conditions, temperature, humidity, and precipitation during the hourly period when manual counts were obtained were also added to the model.

The number of correlated variables has been observed to increase with increasing buffer size, and therefore the variable selection was done very carefully. Although bicycle activity may be better predicted using larger buffer sizes, caution must be taken when selecting variables for the model since the proportion of correlated variables is very likely to increase with increasing buffer size, which has occurred with the four buffer sizes in this analysis. Correlation matrices are not reported in this paper; however, they were useful in identifying which variables and to what extent the variables were related to bicycle activity. For this analysis, to prevent problems caused by multicollinearity, if two variables were highly correlated, correlation greater than 0.4, the variable with the stronger relationship with bicycle activity was retained while the other variable was omitted from the model. Correlation matrices were built for every buffer dimension separately and following every proposed model to double check the correlation between the model variables. After exhaustively trying different combinations of variables, the best model was selected for each dependent variable and these results are reported in Table 3. The best model was selected based on the following criteria: i) all variables in the model must be significant, ideally to the 5 percent level, ii) variables in the model cannot be correlated with one another (correlation greater than 0.4), and iii) the model meeting the first two criteria and with the greatest adjusted R-squared value was chosen.

Table 3 shows the variables that have an effect on bicycle flows. These include employment, presence of schools, metro stations, bus stops, land mix, mean income, presence of bicycle paths and cycle tracks, length of bicycle facilities, average street length, and presence of parking entrance, regardless of how weather is accounted for. The effects of these variables are at different buffer sizes. Note that temperature is not significant in the model with weather as variables, possibly since most of the counts were done in similar temperature conditions resulting in insufficient variability in the temperature to see an effect. Note that since the results for the models accounting for weather using the hourly and daily weather models are very similar, Table 3 is only reporting the results based on the hourly model.

In terms of goodness of fit based on Akaike's Information Criterion (AIC), the spatial models perform better overall compared to the models without spatial autocorrelation. Based on AIC values, the spatial model generated using an α value of 0.8 performed the best for all models with AIC values of 2144 and 2145 for the model adjusted using the hourly model and with weather as variables, respectively.

In general, the models accounting for spatial lag have lower standard error values than models that neglect spatial effects. This highlights the importance of accounting for spatial correlation. Lagged spatial models not only fit the data better than non-spatial models but also provide appropriate estimates and levels of significant values. However, it is also worth noting that while all the variables included in the models are significant to the 5 percent level in the model without a spatial effect, in some cases they are no longer significant to this level once the spatial effect is introduced.

The elasticities reported in Table 3 are those corresponding to the model with the best fit, which, in every case, is the model with a friction value of 0.8. The following discussion

Table 3: Results.

Variables	Hourly Weather Model												
	No Spatial Effect			α=1			α=1.2			α=0.8			Elasticity
	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z	
400m Employment (1000)	0.378	0.044	0.000	0.359	0.044	0.000	0.363	0.044	0.000	0.363	0.044	0.000	5.30%
400m Presence of Schools	0.251	0.082	0.002	0.246	0.081	0.002	0.249	0.081	0.002	0.249	0.081	0.002	1.38%
800m Metro Stations	0.256	0.023	0.000	0.245	0.023	0.000	0.248	0.023	0.000	0.248	0.023	0.000	3.44%
150m Bus Stops	0.087	0.016	0.000	0.086	0.016	0.000	0.087	0.016	0.000	0.087	0.016	0.000	2.94%
800m Land Mix	1.244	0.279	0.000	1.215	0.275	0.000	1.222	0.276	0.000	1.222	0.276	0.000	7.98%
50m Mean Income (1000s)	0.005	0.001	0.000	0.005	0.001	0.000	0.005	0.001	0.000	0.005	0.001	0.000	2.80%
Presence of a Bicycle Lane	0.451	0.133	0.001	0.453	0.131	0.001	0.453	0.132	0.001	0.453	0.132	0.001	36.4%
Presence of a Cycle Track	0.989	0.111	0.000	0.955	0.110	0.000	0.966	0.110	0.000	0.966	0.110	0.000	61.0%
800m Length of Bicycle Facilities	0.070	0.025	0.006	0.064	0.025	0.011	0.066	0.025	0.009	0.066	0.025	0.009	1.39%
800m Average Street Length	-0.016	0.008	0.055	-0.015	0.008	0.062	-0.015	0.008	0.058	-0.015	0.008	0.058	-0.25%
Presence of Parking Entrance	-0.301	0.106	0.005	-0.304	0.104	0.004	-0.305	0.105	0.004	-0.305	0.105	0.004	-3.53%
Constant	2.592	0.202	0.000	3.006	0.236	0.000	2.854	0.225	0.000	2.854	0.225	0.000	
Rho				-0.0004	0.000	0.001	-0.0002	0.000	0.012	-0.0002	0.000	0.012	
R-Squared*	0.4851			0.5000			0.4970			0.5030			
AIC	2155.87			2149.25			2153.58			2143.86			
Variables	Weather as Variables												
	No Spatial Effect			α=1			α=1.2			α=0.8			Elasticity
	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z	
400m Employment (1000)	0.374	0.044	0.000	0.362	0.044	0.000	0.366	0.044	0.000	0.357	0.044	0.000	5.35%
400m Presence of Schools	0.229	0.082	0.005	0.228	0.080	0.005	0.229	0.081	0.004	0.225	0.080	0.005	1.28%
800m Metro Stations	0.258	0.023	0.000	0.250	0.023	0.000	0.253	0.023	0.000	0.247	0.023	0.000	3.51%
150m Bus Stops	0.079	0.016	0.000	0.079	0.016	0.000	0.079	0.016	0.000	0.078	0.016	0.000	2.69%
800m Land Mix	1.313	0.279	0.000	1.287	0.276	0.000	1.296	0.276	0.000	1.277	0.275	0.000	8.43%
50m Mean Income (1000s)	0.005	0.001	0.000	0.004	0.001	0.000	0.005	0.001	0.000	0.004	0.001	0.000	2.57%
Presence of a Bicycle Lane	0.448	0.136	0.001	0.448	0.134	0.001	0.449	0.134	0.001	0.446	0.133	0.001	36.0%
Presence of a Cycle Track	0.991	0.111	0.000	0.967	0.110	0.000	0.977	0.110	0.000	0.955	0.109	0.000	61.5%
800m Length of Bicycle Facilities	0.068	0.025	0.007	0.064	0.025	0.011	0.065	0.025	0.009	0.062	0.025	0.013	1.38%
Average Street Length	-0.015	0.008	0.058	-0.015	0.008	0.063	-0.015	0.008	0.060	-0.015	0.008	0.067	-0.24%
Presence of Parking Entrance	-0.316	0.106	0.003	-0.319	0.105	0.002	-0.319	0.105	0.002	-0.318	0.104	0.002	-3.75%
Temperature	0.002	0.006	0.676	0.002	0.006	0.748	0.002	0.006	0.714	0.001	0.006	0.793	0.21%
Humidity	-0.009	0.002	0.000	-0.009	0.002	0.000	-0.009	0.002	0.000	-0.008	0.002	0.000	-5.39%
Presence of Precipitation	-0.323	0.096	0.001	-0.305	0.095	0.001	-0.311	0.095	0.001	-0.299	0.095	0.002	-3.49%
Constant	2.925	0.286	0.000	3.202	0.307	0.000	3.074	0.299	0.000	3.403	0.322	0.000	
Rho				-0.0003	0.000	0.021	-0.0002	0.000	0.124	-0.0004	0.000	0.002	
R-Squared*	0.5174			0.5210			0.5190			0.5230			
AIC	2150.19			2148.92			2151.82			2144.81			

* Elasticities are expressed in terms of a 10 percent change in the independent variable or a 0 to 1 change in the case of a dummy variable. **Squared correlation for spatial models.

will be based solely on the model containing weather conditions as variables. From this, one can observe that increasing the number of working people within 400 meters of the intersection by 10 percent would cause a 5.35 percent increase in bicycle flows. The presence of a school or metro station would cause a 1.28 percent and 3.51 percent increase in bicycle activity, respectively. Average street length and the presence of a parking entrance at an intersection have a negative effect on bicycle activity. Shorter street lengths imply greater connectivity due to shorter block distances. Perhaps one of the most important results to mention is the important effect of the presence of bicycle facilities, which varies according to the facility type. Introducing a bicycle lane or a cycle track at an intersection would cause an increase in bicycle volumes by 36 percent and 61 percent, respectively. While all the variables discussed above are indeed significant across all models, some of these results do not imply causality. The effect of transit variables such as metro stations and bus stops can be interpreted as having a direct effect on bicycle activity through mode transfers during the trip (which might be marginal). An indirect link can also be observed since, in general, neighborhoods that use transit also tend to walk and cycle more, and therefore these central neighborhoods may have a high bicycle modal share.

5 Conclusion

Knowledge of bicycle activity in critical locations, such as through intersections, serves many purposes and uses by engineers, city planners, and designers as well as public health professionals. Bicycle volumes can be used: i) by transportation agencies and governments to guide the design and location of new bicycle facilities, ii) in safety analysis to provide complete risk exposure measures, and iii) to evaluate the impact of new developments or built-environment changes and prioritize available resources. More uses are likely to be developed as transportation agencies and governments continue to emphasize the importance of active transportation in any sustainable transportation system.

This study has identified and quantified the effects of built environment and road and transit network characteristics as well as bicycle facilities on bicycle activity through signalized intersections.

A methodology to normalize manual counts and take into account spatial autocorrelation is introduced. The main conclusions drawn from this study are:

1. The models accounting for spatial autocorrelation have been shown to provide slightly better results than when spatial effects are neglected. Spatial correlation, however, has a small effect on the parameter estimates, which are very similar across models.

2. The explanatory power of the developed models is in the range of those found in the literature. In all cases, the R-squared value is between 0.48 and 0.52. This means that although the quality of the model is acceptable, the prediction capability still needs to be improved. This can be done by increasing the sample size and adding other variables (geometry, topography, and presence of bike-sharing stations). Compared to pedestrian activity models, it seems that bicycle activity is more difficult to predict.

3. As land mix increases so does bicycle activity. A 10 percent increase in land mix is expected to cause around an 8 percent increase in bicycle activity. Intersections located in diverse areas with residential, commercial, governmental, and industrial zones as well as parks and recreational areas have higher cyclist volumes than intersections without.

4. Intersections with bicycle facilities such as bicycle lanes and cycle tracks have over 36 percent and 61 percent more cyclists, respectively, than intersections without these facilities. These results highlight the importance of paying particular attention to bicycle activity when installing bicycle facilities or making changes to the built environment. Intersections should be redesigned with the appropriate interventions to handle the increase in bicycle activity. For instance, additional space and traffic controls may be necessary (e.g., bicycle boxes, exclusive bicycle phase, etc.). An increase in the number of cyclists without appropriate interventions and without any reduction in motor vehicle traffic flows can cause an increase in the number of cyclist injuries. This highlights the importance of bicycle facility location and the need for traffic calming measures when installing bicycle facilities.

As part of the future work, the location and proximity of Bixi stations (Montreal's public bike system) in relation to Montreal's intersections will be added as a factor to predict bicycle flows through intersections. Another factor to consider is the elevation and slope of the intersection and its surrounding area; steep inclines may discourage cyclists from choosing a certain route. Also, a larger sample of intersections will be used—additional data collection will be undertaken. Finally, larger buffer dimensions could be used to account for greater travel distances achieved by cyclists in comparison to pedestrians. To measure the effect of bicycle facilities on bicycle activity, before-and-after observational studies will be used to quantify the volume changes.

References

Brandenberg, C., A. Matzarakis, and A. Arnberger. 2007. Weather and cycling—a first approach to the effects of weather on cycling. *Meteorological Applications*, 14:61–67, doi:10.1002/met.6.

Dill, J. and T. Carr. 2003. Bicycle commuting and facilities in major U.S. cities: If you build them, commuters will use them. *Transportation Research Record: Journal of the Transportation Research Board*, 1828: 116–123, doi:10.3141/1828-14.

Dill, J. and J. Gliebe. 2008. *Understanding and measuring cycling behavior: A focus on travel time and route choice* (OTREC-RR-08-03). Oregon Transportation Research and Education Consortium, Portland, Oregon. http://otrec.us/project/33.

Drukker, D. M., H. Peng, I. R. Prucha, and R. Raciborski. 2011. *Creating and managing spatial-weighting matrices using the Spmat command*. StataCorp, College Station, Texas, USA.

Garrard J., G. Rose, and S. K. Lo. 2008. Promoting transportation cycling for women: The role of bicycle infrastructure. P*reventative Medicine*, 46: 55–59. doi:10.1016/j.ypmed.2007.07.010.

Griswold, J. B., A. Medury, and R. J. Schneider. 2011. Pilot models for estimating bicycle intersection volumes, Transportation Research Board 90th Annual Meeting, Washington, DC. doi:10.3141/2247-01.

Haynes, M. and S. Andrzejewski. 2010. GIS based bicycle and pedestrian demand forecasting techniques. Travel model improvement program webinar. http://tmip.fhwa.dot.gov/sites/tmip.fhwa.dot.gov/files/fehr_and_peers.pdf (Accessed Oct. 11, 2010).

Hillman, M. 1993. Cycling and the promotion of health. *Policy Studies*, 14: 49–58. doi:10.1080/01442879308423639.

Jones, M., S. Ryan, J. Donlon, L. Ledbetter, D. R. Ragland, and L. Arnold. 2010. Seamless travel: Measuring bicycle and pedestrian activity in San Diego County and its relationship to land use, transportation, safety, and facility type. Caltrans Task Order 6117. California Department of Transportation, Sacramento, CA.

Kim, J. K., S. Kim, G. F. Ulfarsson and L. A. Porrello. 2007. Bicyclist injury severities in bicycle-motor vehicle accidents. *Accident Analysis and Prevention*, 39: 238–251. doi:10.1016/j.aap.2006.07.002.

Kockelman, K. M. 1997. Travel behavior as function of accessibility, land-use mixing and land-use balance evidence from San Francisco Bay Area. *Transportation Research Record: Journal of the Transportation Research Board*, 1607: 116–125. doi:10.3141/1607–16.

Krizek, Kevin J., G. Barnes, and K. Thompson. 2009. Analyzing the effect of bicycle facilities on commute mode share over time. *Journal of Urban Planning and Development*, 135: 66–73. doi:10.1061/(ASCE)0733-9488(2009)135:2(66).

McCahill, C. and N. W. Garrick. 2008. The applicability of space syntax to bicycle facility planning. *Transportation Research Record: Journal of the Transportation Research Board*, 2074: 46–51. doi:10.3141/2074-06.

Miranda-Moreno, L., and T. Nosal. 2011. A look at weather's impact on physically separated cycling facilities and temporal trends in an urban environment. Transportation Research Board 90th Annual Meeting, Washington, DC.

Moudon, A. V., C. Lee, A. D. Cheadle, C. W. Collier, D. Johnson, T. L. Schmid, and R. D. Weather. 2005. Cycling and the built environment, a US perspective. *Transportation Research Part D*, 10: 245–261. doi:10.1016/j.trd.2005.04.001.

Nankervis, M. 1999a. The effect of weather and climate on bicycle commuting. *Transportation Research Part A*, 33: 417–431. doi:10.1016/S0965-8564(98)00022–6.

Nankervis, M. 1999b. The effects of weather and climate on urban bicycle commuters' decision to ride—a pilot study. *Road and Transportation Research*, 8: 85–97.

Nosal, T. and L. Miranda-Moreno (2012). Cycling and weather: A multi-city and multi-facility study in North America. Transportation Research Board 91st Annual Meeting, Washington, DC.

Pucher, J. and R. Buehler. 2006. Why Canadians cycle more than Americans: A comparative analysis of bicycling trends and policies. *Transport Policy*, 13: 265–279. doi:10.1016/j.tranpol.2005.11.001.

Pucher, J. and R. Buehler. 2008. Making cycling irresistible: Lessons from the Netherlands, Denmark, and Germany. *Transport Reviews*, 28: 495–528. doi:10.1080/01441640701806612.

Pucher, J., J. Dill, and S. Handy. 2010. Infrastructure, programs, and policies to increase bicycling: An international review. *Preventive Medicine*, 50: S106¬–S125. doi:10.1016/j.ypmed.2009.07.028.

Richardson, A. J. 2000. *Seasonal and weather impacts on urban cycling trips.* The Urban Transit Institute Report 1-2000, Victoria, Australia.

Richardson, A. J. 2006. Estimating bicycle usage on a national cycle network. *Transportation Research Record: Journal of the Transportation Research Board*, 1982: 166–173. doi:10.3141/1982-22.

Schneider, R. J., L. S. Arnold, and D.R. Ragland. 2009. Pilot model for estimating pedestrian intersection crossing volumes. *Transportation Research Record: Journal of the Transportation Research Board*, 2140: 13–26. doi:10.3141/2140-02.

Statistics Canada. 2003. Where Canadians work and how they get there. Statistics Canada, Ottawa, Canada.

Stinson, M. A. and C. R. Bhat. 2003. Commuter bicyclist route choice: Analysis using a stated preference survey. *Transportation Research Record: Journal of the Transportation Research Board*, 1828: 107–115. doi:10.3141/1828-13.

Strauss, J. and L. Miranda-Moreno. 2011. A methodology for the analysis of cyclist injury occurrence at signalized intersections. *Proceedings of the 21st Canadian Multidisciplinary Road Safety Conference*, Halifax, Nova Scotia.

Thomas, T., R. Jaarsma, and B. Tutert. 2009. Temporal variations of bicycle demand in the Netherlands: The influence of weather on cycling. Transportation Research Board 88th Annual Meeting, Washington, DC.

Vélo Québec. 2005. L'État du vélo au Québec en 2005. Vélo Québec Association, Montréal, Québec.

Winters, M., M. C. Friesen, M. Koehoorn, and K. Teschke. 2007. Utilitarian bicycling: A multilevel analysis of climate and personal influences. *American Journal of Preventative Medicine*, 32: 52–58. doi:10.1016/j.amepre.2006.08.027.

World Road Association (PIARC), Technical Committee on Road Safety. 2003. *Road safety manual.* Route 2 Market Limited, Swanley, United Kingdom.

Xing, Y., S. L. Handy, and P. Mokhtarian. 2010. Factors associated with proportions and miles of bicycling for transportation and recreation in six small US cities. *Transportation Research Part D: Transport and Environment*, 14: 73–81. doi:10.1016/j.trd.2009.09.004.