Evaluating methods for measuring daily walking to public transport: Balancing accuracy and data availability

Emily Grisé
School of Urban Planning, McGill University
E-mail: emily.grise@mail.mcgill.ca

Rania Wasfi
Centre de Recherche du CHUM, Université de Montréal
Department of Geography, McGill University
E-mail: rania.wasfi@mail.mcgill.ca

Nancy Ross
Department of Geography, McGill University
E-mail: nancy.ross@mcgill.ca

Ahmed El-Geneidy
School of Urban Planning, McGill University
E-mail: ahmed.elgeneidy@mcgill.ca

July 2018

Word count: (6,549 words, 4 tables)

For Citation please use: Grisé, E., Wasfi, R., Ross, N., & El-Geneidy, A. (2019). *Evaluating methods for measuring daily walking to public transport: Balancing accuracy and data availability.* Paper presented at the 98th Annual Meeting of the Transportation Research Board, Washington D.C., USA.

ABSTRACT

Accurately quantifying physical activity accumulated through daily commuting is challenging due to the scarcity of detailed data, especially for public transport trips. Using Montreal, Canada as a case study, this paper measures and compares an individuals' daily amount of walking to and from public transport in their regular commute to work using two datasets and two methods. The first method uses urban level detailed origin-destination survey data. Distances walked to and from public transport stations are measured using trip details provided from the survey. The second method uses open data including commuting flows obtained from census data. Public transport trips for each flow are modeled using General Transit Feed Specification (GTFS) data obtained from public transport operators, and through applying a fastest route algorithm. Walking distances are then extracted from these trip routes. Multilevel mixed-effect regression modeling is used to identify the determinants of total walking when using both methods. A sensitivity analysis is then used to derive an adjustment table for those who wish to use open data. For commuter train users 471 meters must be added to walking estimates obtained from the commuting flows data, while negative adjustments are required for subway users (122 meters), city bus users (366 meters), suburban bus users (516 meters), and peripheral bus users (1186 meters). Findings from this study provide professionals without access to detailed origin-destination survey data with a guide to accurately use open data to estimate total walking distances accumulated in a daily commute by public transport.

Keywords: Transit, walking, travel behaviour, daily physical activity

INTRODUCTION

Being active for at least 150 minutes a week can help reduce the risk of heart disease, type 2 diabetes, high blood pressure, and obesity (*1*). However, only two in ten Canadian adults and one in ten children met the Canadian physical activity guidelines in 2013 (*2*). While the most commonly cited barrier to exercise is lack of time (*3-7*), the good news is that individuals can achieve the recommended physical activity guidelines through small sessions of at least ten minutes of activity throughout the day (*8*). Walking accumulated through daily travel is one way that individuals can meet physical activity targets. Compared to public health interventions that aim to promote active lifestyles and change the behaviour of individuals, policies that aim to increase the proportion of trips that are made on active modes of travel such as public transport, bicycling and walking, are likely to increase the activity levels of a given population in a way that can be sustained in the long term (*9*).

While walking is regarded as one of the most accessible forms of physical activity for all ages, it is limited in its range as a mode of transport. The ability to walk to work is relatively uncommon and only for individuals who can afford to live in close proximity to their place of work. This has meant that increasing scholarly attention is being paid to walking to public transport, as almost all trips by this mode require some physical activity. Researchers have looked at how much physical activity is accumulated when commuting by public transport, and what characteristics of the built environment are associated with higher levels of walking (*10-14*).

Most studies to date have measured walking to public transport using either travel survey data (*10; 13; 14*), or accelerometer readings (*11*). However, city or municipal planners often lack either the access to such detailed data or the resources to monitor physical activity levels of the population. Biosensor-assessed step counts are also burdensome for respondents and can induce non-routine behaviour. Veillette et al. (*15*) recently calculated levels of walking to public transport using detailed and non-detailed travel survey data. While this study provided a better understanding of how modeled walking routes differ from reported routes, it did not present a methodology for municipalities without access to travel survey data to measure walking to public transport. The goal of our study is to simulate a scenario where a planner only has access to aggregate, publicly available commuting flow census data at the census tract (CT) level (*16*), which provides details on the number of residents that live and travel between different CTs. We propose that such census data can be used to measure walking to public transport when combined with General Transit Feed Specification (GTFS) data obtained from public transport operators. We undertake a multilevel regression analysis to identify the determinants of walking using both methods, then use a series of sensitivity analyses to compare both methods of calculating walking distances and generate an adjustment table that can be used by professionals who are estimating walking levels from widely available census commuting flows and open GTFS data. This study is of relevance to planners, researchers, and policy-makers from municipalities of all sizes who are wishing to measure the relationship between physical activity and public transport use in areas where access to detailed travel survey information is limited.

LITERATURE REVIEW

In response to rising obesity levels observed in many Westernized countries, research from leading medical and public health journals have advocated for the increased uptake of walking and bicycling as a means of achieving daily moderate-intensity physical activity goals for improved health (*17; 18*). The specific level of physical activity achieved through daily public transport use is mediated by many factors. Potentially the most significant factor determining physical activity levels of public transport users is the mode being accessed. Distances walked to buses are generally shorter compared to commuter trains and subway, and thus catchment areas have traditionally been defined as 400 meters for bus and 800 meters for rail (for example *19; 20; 21*). However, several authors have observed walking distances to public transport beyond the 400 and 800 meter threshold, thus presenting an opportunity for operators to increase stop spacing (*22-24*). Variations have also been observed between downtown and suburban locations with suburban commuters tend to walk further to more comfortable and faster modes (*23; 25*).

Using detailed origin-destination survey data in Montreal, Canada, Wasfi, Ross and El-Geneidy (*14)* modeled the underlying factors that contribute to achieving the recommended 30 minutes of daily physical activity while commuting by public transport. The authors concluded that approximately 11% of commuters achieved their recommended minutes of daily physical activity, and suburban train users walked the greatest number of minutes per day. Similar findings were observed by Besser and Dannenberg (*10)*, who used the National Household Travel Survey to assess walking times to public transport across America. The authors observed that 29% of public transport users achieved more than 30 minutes of daily physical activity from walking to and from stops, and that rail users were 1.67 times more likely than bus users to achieve their daily physical activity target. Although these studies find that not all public transport riders regularly meet their daily physical activity guidelines solely through walking to public transport, it is important to note that individuals walking to public transport are reaching some level of physical activity that they would not have if commuting by a sedentary form of transport, such as driving.

More recently, researchers have been interested in understanding the exertion level of public transport related walking to ascertain whether these walkers are in fact meeting the required moderate-to-vigorous physical activity (MVPA) intensity level set within physical activity guidelines (*26*). Using accelerometer data collected from participants residing within a half mile of a new light-rail stop in Salt Lake City, USA, Brown and Werner (*11)* measured the length and intensity of walking trips to and from light rail stops, concluding that these walking trips are of moderate-intensity, and collectively can contribute to attaining the WHO-recommended physical activity level. MacDonald et al. (*12)* studied the impact of a new light rail transit (LRT) line in Charlotte, USA on obesity and physical activity level. These authors found that commuting by LRT was associated with a reduction in BMI and a reduced odds of becoming obese over time, compared to a control group of individuals living in similar neighbourhoods but did not commute by LRT.

DATA

Detailed Origin-Destination Survey

A common method for cities to evaluate the travel behaviour of residents is through the collection of origin-destination (OD) surveys. OD survey data is used by cities in many ways including modeling travel demand and evaluating the modal split. In Montreal, Canada, the OD survey is conducted every five years by the Agence Métropolitaine de Transport (AMT) (*27*). Five percent of households are sampled and asked to recall all trips made in the past day by each household member. This data includes coordinates of the location of the home and each location visited in that day, mode taken for each trip, and personal and household characteristics. For trips taken by public transport, each respondent was asked to provide route details including bus route(s) used, subway or train station of access and egress, and subway used. After selecting all home-based trips for the purpose of work or school that were taken by adults (over the age of 18) that began and ended in Greater Montreal, our sample consisted of 9,588 trips that were made using bus, subway or commuter rail. We omitted trips with missing route information, as well as trips where individuals drove to a transit station.

Commuting Flows

As part of the Canadian Census that is collected every five years, Statistics Canada collects information on the commuting modes and patterns of all Canadians. This data is available at the CT level across Canada and provides detail regarding the number of residents and their travel mode. Similar data exist in other countries around the world, making this approach appealing for international comparative research. We selected all commuting flow pairs within the Greater Montreal Region where there were at least 10 individuals who commute by public transport between CT pairs. Note, information on flows with between 0 and 10 individuals commuting were suppressed for confidentiality. Our sample consisted of 2,755 commuting flows.

METHODOLOGY

Origin-Destination Microdata Scenario

We modeled walking distances from public transport for each work-related trip in ArcGIS using two shapefiles. The first shapefile was a pedestrian street network where highways were removed from the road network. The second contained all stops serving the subway, commuter train and bus services. The objective here was to calculate the walking distance from the individuals' home to the first public transport stop and the walking distance from the last public transport stop used to the final destination. As we only knew the route(s) used in each trip, we used the closest facility tool to generate routes from each home and work location to their respective nearest 200 public transport stops. We then selected only the routes connecting each home or work location to a nearby stop that matched the route that the respondent reported using. The two distances were summed to determine the total walking distance for each public transport trip.

Using the stop where each individual began their trip, we measured the headway (time between subsequent vehicles) of the first public transport route used at the beginning of the trip during the morning peak (7 and 9 AM). We also measured the distance travelled in-vehicle, to provide a complete picture of the individual's daily public transport commute. This was done using the previously identified first and last stop used in their trip, and by determining the length of the public transport route between these points using a network built from GTFS data from all public transport operators in the Greater Montreal region. GTFS data produced in 2013 was used to represent the level of service that was present at the time of data collection. The number of transfers that occurred during this trip was calculated as the total number of distinct public transport routes used during this trip minus one.

Commuting Flows Scenario

Using the CT centroids, we determined the public transport route of all flows according to the shortest travel time. GTFS data from all public transport operators in the Greater Montreal region were obtained and a public transport network was constructed in ArcGIS using the 'Add GTFS to a network dataset' toolbox available in ArcMap. The fastest route between each CT pair was determined, and this travel time includes time for walking, waiting, boarding and alighting, and in-vehicle travel time. We solved these routes using a departure time of 8 AM on a regular Tuesday in December of 2013. After all routes were obtained, a series of GIS computations were applied to obtain the walking distances for each trip, first mode used, distance traveled in-vehicle, and number of transfers.

The portion of each route that required the individual to walk to the first and last public transport stop was calculated by intersecting each modeled trip flow with the street network. After routes were intersected with the street network, we applied the unsplit line tool in ArcGIS to merge coinciding walking segments. We selected only the walking segments that represented either walking from origin to first stop or walking from last stop used to final destination. After selecting the two routes that were connected to the origin and destination point, we summed the two walking distances to represent total walking during a trip.

The number of transfers were obtained by subtracting the number of walking segments by two, which represent walking segments at the origin and destination. However, to represent transfers occurring between subways and trains we determined the number of subway and rail transfer stations each user passed through and adjusted the number of transfers to account for transferring subway or train lines. The first mode taken was determined by isolating the walking segment from each respondents' home to the first stop and identifying the portion of the public transport route that intersects with the first walking segment. Lastly, we determined whether a trip began on a city, suburban, or a peripheral bus by determining which agency boundary each trip originated from.

Statistical analysis

Two multilevel mixed-effect regression models were constructed to uncover the key determinants of daily walking to public transport as well as to explore how the two datasets, detailed OD data and commuting flows, differ in their estimation of walking levels (under the assumption that the walking distances obtained from detailed OD data are the gold standard). The first model evaluated the determinants of walking to public transport according to detailed OD survey data, to provide a baseline understanding of what factors influence walking levels. In this model, we nested each observation in its respective origin CT, to account for characteristics in the built environment unique to each neighbourhood that would likely influence walking behaviour. To account for characteristics of each respondents' local built environment that are not directly controlled for in our model, the multilevel model approach adopted in this study reduces the potential biases that may be imposed on the model due to not controlling for such characteristics. We interpret the model estimates under different scenarios, to estimate walking distances specific to the first public transport mode taken.

In our second model, trip details obtained from the detailed OD data scenario and the commuting flows scenario are merged using a dummy variable to differentiate the method used to obtain the details of each trip. In this model we nested routes first according to their origin CT and second by their census flow. Nesting each observation in the census flow accounts for similarities that are likely to arise as a result of sharing the same origin and destination CTs, such as walkability, level of service and partially self-selection. Since these variables are not directly controlled for in our model, adopting a multilevel approach reduces estimation biases. For routes obtained from the detailed OD data, the census flow was obtained by concatenating the origin and destination CT that are provided in the dataset. According to the frequency of travel between CTs, some census flows had multiple routes from the detailed OD data. Finally, the estimates of this model, are used to conduct a second sensitivity analysis, to obtain estimates of daily walking distances according to first mode taken for both methods. Differences between estimates of walking distance obtained by the two scenarios were observed and presented as recommended adjustment values, capturing how much the commuting flows method either underestimates or overestimates walking distances relative to the detailed OD scenario.

Both models use home to work trip walking distance (distance walked from home to first stop and from last transit stop to work) as their dependent variable. The independent variables included in the two models vary slightly according to data availability. Trip details, including number of transfers, first mode taken, in vehicle distance, and neighbourhood characteristics are included in both models (Table 1). Additional variables included in the first model (detailed OD scenario) include personal characteristics such as income level, age and sex. In the second model (detailed OD scenario and commuting flows scenario merged), interaction terms were included in the model to test the effect of first mode taken on total walking distance between the two methods of analysis.

Table 1 Description of variables and summary statistics

		Detailed OD			Commuting flows	
Variable	**Definition**	**Mean**	**Std. Dev**		**Mean**	**Std. Dev**
Walking distance (m)	Individuals' total walking distance during home to work trip (includes distance to access first transit stop and distance from final transit stop to work)	1007.31	1003.47		1034.63	944.14
Individual Characteristics						
Sex	A dummy variable equaling one if the individual reported their sex as male	45%	50%		N/A*	N/A
Age	Age of the individual in years	43.13	12.26		N/A	N/A
Middle income ($20K - 79K)	A dummy variable equaling one if the individual reported a household income between $20,000 - 79,000	46%	50%		N/A	N/A
Low income (< 20K)	A dummy variable equaling one if the individual reported a household income under $20,000	11%	31%		N/A	N/A
Trip characteristics						
In-vehicle distance (km)	Total distance spent in transit vehicle(s)	10.00	7.02		10.08	7.04
Headway of first route used (min)	The headway (time between subsequent vehicles) of the first transit route used at the beginning of the trip	16.59	33.89		N/A	N/A
First mode taken						
Subway	A dummy variable equal to one if the subway was the first mode used	30%	46%		29%	45%
Train	A dummy variable equal to one if the train was the first mode used	7%	25%		5%	22%
Peripheral bus	A dummy variable equal to one if a peripheral bus was the first mode used	5%	22%		5%	22%
Suburban bus	A dummy variable equal to one if a suburban bus was the first mode used	14%	35%		14%	35%
GTFS Dummy * First mode taken subway	An interaction term representing a trip modeled using GTFS data, where the first mode taken was a subway	N/A	N/A		--	--

GTFS Dummy * First mode taken train	An interactive term representing a trip modeled using GTFS data, where the first mode taken was a commuter train	N/A	N/A		--	--
GTFS Dummy * First mode taken suburban bus	An interactive term representing a trip modeled using GTFS data, where the first mode taken was a suburban bus	N/A	N/A		--	--
GTFS Dummy * First mode taken peripheral bus	An interactive term representing a trip modeled using GTFS data, where the first mode taken was a peripheral bus	N/A	N/A		--	--
Number of transfers						
1 transfer	A dummy variable equal to 1 if 1 transfer was undertaken during the trip	37%	48%		34%	47%
2 transfers	A dummy variable equal to 1 if 2 transfers were undertaken during the trip	19%	39%		23%	42%
3 or more transfers	A dummy variable equal to 1 if 3 or more transfers were undertaken during the trip	4%	19%		10%	30%
Characteristics of home neighbourhood						
Percent of the population with a university degree	Percent of people within the individuals' home CT with a university degree	27.88	13.66		27.83	13.61
Population density (km^2) of home CT	Population density of the individuals' home CT	7471.64	5090.53		7547.27	5117.05
Number of bus stops in home CT	Number of bus stops located in the CT of the individuals' home	79.36	32.46		79.65	32.26
GTFS Dummy	A dummy variable equal to one if the trip characteristics were modeled using GTFS data	N/A	N/A		0.22	0.42

* N/A represents a variable was not included in the respective model

RESULTS

Comparing trip details by scenario

The average walking distance of a home to work commute as measured by the detailed OD scenario is 1007 meters, compared to 1035 meters as estimated by the commuting flows method. Therefore, on average the commuting flows method underestimates daily walking distance by 3% when compared to reported trip details in the OD survey. However, important differences in average walking distances between the two scenarios are apparent when comparing route characteristics (Table 2). Looking at average walking distances by first mode taken, we observe that commuting flow trips that begin on a peripheral bus overestimate walking by 44% when compared to detailed OD data trips. Interestingly, estimates of daily walking distances for trips beginning on a subway were most closely estimated by the commuting flows method (underestimated by 142 meters). In line with the overall goal of increasing the proportion of the population that meets the minimum daily physical activity goals (30 minutes a day), we calculated what proportion of individuals will meet their daily activity goals through walking to and from public transport stations. Important differences are observed across all modes. Results of the commuting flows scenario show an underestimation of the proportion of train users that meet the recommended daily walking target, while this scenario overestimates the proportion of city bus, suburban and peripheral bus users who meet the recommended physical activity targets through walking to public transport.

Other considerable differences between the two scenarios are likely contributing to differences in walking distances estimates across the two methods are likely a result of routing differences. In the detailed OD survey, 6% of the sample reported the commuter train as their first mode, whereas the commuting flows data estimates that 1% of trips begin on a train. This result can be explained by a few possible reasons. One reason being that the centroid of a CT may not coincide with where the majority of residents within a CT live and thus the travel time associated with using that train exceeds another mode of public transport, or it routes that individual on a bus which will later connect to the train. Alternatively, since all trips were routed at 8 AM on a weekday, this departure time may not align with the train schedule thus another public transport mode was a faster option. Furthermore, waiting time for a public transport vehicle is included in the calculation of travel time, therefore for individuals who take the commuter train and adjust their departure time according to the train schedule, the fastest route algorithm may route them first on a bus if a shorter travel time to their destination is available. One solution to this issue that has been demonstrated in the literature, is to calculate travel time for every minute during a time period of interest (*28*), which can then be used to obtain the fastest route. Yet such approach is time consuming and cannot be easily adopted, accordingly we opted not to use such approach in this paper as we are more concerned about the transferability of the method to different regions.

The second noticeable difference in routing details between these methods are the number of transfers taken for each trip. Most notably, only ten percent of trips estimated with the commuting flows method involved zero transfers, whereas 41 percent of detailed OD trips did not involve transfers. Public transport users are often averse to transferring and in reality may choose to either walk further distances or choose alternative travel routes that might be longer in time to avoid transferring between either the same modes or switching transit modes. The shortest time travel algorithm used to route public transport trips from the commuting flows data does not include preferences or penalties for transfers, accordingly future research can be conducted to asses how different routing preferences and/or penalties can be incorporated into routing algorithms with the aim of more closely modeling travel behaviour.

TABLE 2 Comparing route characteristics between detailed OD data and commuting flows scenario

	Detailed OD data scenario	Commuting flows scenario
First mode taken (home-work trip) (%)		
Train	6	1
Subway	30	27
City bus	45	51
Suburban bus (STL and RTL)	14	15
Peripheral bus (CIT)	5	6
Number of transfers taken (home-work trip) (%)		
0 transfers	41	10
1 transfer	37	24
2 transfers	19	35
3 or more transfers	4	31
Average daily walking distance by first mode taken (m) †		
Train	4,632	2,932
Subway	2,255	2,113
City bus	1,519	2,010
Suburban bus (STL and RTL)	1,681	2,499
Peripheral bus (CIT)	2,471	4,410
Proportion of individuals meeting recommended daily walking target* (%)		
Train	71	58
Subway	24	18
City bus	10	17
Suburban bus (STL and RTL)	13	31
Peripheral bus (CIT)	32	68

*30 minute daily physical activity target

† Daily walking distance was obtained by multiplying total walking distance during home to work commute by 2

Results of the statistical models

Among the individual characteristics that were included in Model 1 (modeling total walking distance from detailed OD scenario), age was the only variable associated with walking (Table 3). A one-year increase in age was associated with a decrease in distance walked by 2.5 meters, while holding all other variables constant. Income was not a significant predictor of walking. This contradicts previous literature, which observed that lower income individuals walked greater distances (*14*). Our contrasting findings are potentially mediated when controlling for first mode taken, due to the network structure and high cost of living in areas surrounding subway or train stations. All variables describing trip characteristics included in our model were found statistically significant, except for the headway of the first route taken. Previous research has observed a negative relationship between service frequency and walking distance, where service with larger waiting times reduces total walking distance (*14*). The relationship between service frequency and walking distance is likely better captured in this model by the first mode taken variable, as distances walked to access a transit stop vary according to differences in service such as speed, comfort, frequency (*19; 23*). Every additional kilometer traveled in vehicle, is associated with an additional 14 meters of walking, while all other variables in the model are constant.

The first mode taken variables show that train users walk on average 1.3 kilometers farther than individuals who began their trip on a city bus, while subway users walk 581 meters further than city bus users. On the other hand, suburban bus users walk 186 meters less than city bus users, and similarly peripheral bus users walk 254 meters less than city bus users. Using the mean value of all continuous variables, and assuming a no transfer situation, we found that individuals who begin their daily commute on a train, walk on average 2.3 kilometers on their home-work commute, which is the farthest walking distance of all modes and translates to a walking time of 25 minutes at a moderate pace (5.47 km/h or 3.4 miles/hour). This is followed by individuals who began their trip on a subway (1.6 km or 17 minute walk), city bus users (977 meters or 11 minute walk), suburban bus users (790 meters or 9 minute walk), and lastly peripheral bus users (723 meters or 8 minute walk).

The number of transfers involved in completing a public transport trip is negatively and significantly associated with walking distance. When compared to trips that do not involve transferring modes or vehicles, a trip that includes 1 transfer reduces the trip walking distance by 186 meters. Similarly, a trip involving 2 transfers reduces the walking distance by 326 meters and three or more transfers 436 meters, both when compared to a trip involving 0 transfers.

Lastly the model shows a negative and statistically significant association between walking distance and the percent of the population with a university degree and the number of bus stops within the origin CT. Every one percent increase in the population with a university degree or higher residing in the CT of trip origin is associated with a 3.5 meter decrease in walking distance. Moreover, the results indicate that every additional bus stop present in the origin CT of a trip is linked with a decrease in walking distance by 5.7 meters. The intra-class correlation coefficient showed that the origin census tract explains approximately 30% of the total variance in walking distance.

Comparing the findings of this model to a previous study conducted by Wasfi, Ross and El-Geneidy (*14)* using detailed OD survey data collected in the Greater Montreal Region in 2003, daily walking distance varied significantly according to mode of transit. Direct comparisons to the model results of this study cannot be made, as the authors independent variable controlling for mode(s) used was represented as the number of times each mode was used, rather than first mode taken. Nonetheless, both studies observe that commuter train users walk the greatest distance.

Table 3 Results of the multilevel mixed-effect regression modeling total walking distance

	Model 1				Model 2			
Variable	**Coeff.**	**Sig.**	**95% Conf. interval**		**Coeff.**	**Sig.**	**95% Conf. interval**	
Male	10.6		-24	45.2				
Age	-2.5	***	-3.9	-1.1				
Medium income ($20K - 79K)	-18.9		-80.1	42.3				
Low income (< $20K)	16.6		-21.2	54.3				
Headway of first route used (min)	0.1		-1	1.1				
In vehicle distance (km)	14.1	***	9.8	18.5	10.3	***	6.4	14.1
First mode taken *(ref = city bus)*								
Subway	580.6	***	524.7	636.6	580.6	***	531	630.1
Train	1307.5	***	1162.3	1452.7	1351.8	***	1254.7	1448.9
Suburban bus	-186.4	***	-295.8	-77	-109.7	*	-204.9	-14.5
Peripheral bus	-253.8	**	-428	-79.5	-82.0		-216.9	52.8
Commuting flows * First mode: subway					-243.5	***	-321.6	-165.5
Commuting flows * First mode: train					-836.4	***	-1198.3	-474.6
Commuting flows * First mode: suburban bus					150.3	**	56.5	244.2
Commuting flows * First mode: peripheral bus					820.4	***	683	957.9
One transfer	-186.4	***	-228.9	-143.2	-195.4	***	-233.3	-157.5
Two transfers	-325.7	***	-381.2	-270.2	-346.7	***	-394	-299.5
Three or more transfers	-435.9	***	-538.9	-333	-399.7	***	-467.9	-331.4
Percent of population with a university degree	-3.5	*	-6.9	-0.1	-3.1	*	-6	-0.2
Population density (1000 km^2)	-10.9		-22.4	0.7	-14.5	**	-24	-5
Number of bus stops	-5.7	***	-7.6	-3.8	-4.4	***	-5.9	-3
Commuting flows Scenario					365.6	***	316	415.2
Constant	1573	***	1402.5	1743.6	1420.8	***	1288.6	1553.1
AIC	156,718				198,864			
BIC	156,854				199,012			
Observations	9,549				12,220			
Log likelihood	-78,340				-99,412			
Groups	No. Groups	Intraclass correlation				Intraclass correlation		
CT origin	830		0.3		822	0.2		
CT flow					7,377	0.5		

* 95% significance level | ** 99% significance level | *** 99.9% significance level

The results of Model 2, predicting the determinants of total walking when using both the detailed OD and commuting flows scenario, are presented in Table 3. The dummy variable, commuting flows scenario, was used to differentiate walking distance measured under the different scenarios. The coefficient for the commuting flows dummy variable shows that on average the commuting flows method overestimates walking distance by 366 meters, when all other variables in the model are constant. However, the series of interactive terms, multiplying the commuting flows dummy variable by the first mode taken dummy, indicates that the commuting flows scenario underestimates walking for trips that began on a subway or train, whereas this method overestimates walking for trips which began on either a suburban or peripheral bus. However, the interactive terms must be interpreted in combination with the commuting flows dummy and the first mode taken dummy variables. To illustrate the results, we conducted a sensitivity or scenario analysis to estimate walking distances achieved by first public transport mode used. The mean value of all continuous variables was used in our calculations, and each scenario was calculated assuming no transfers occurred. Under the commuting flows scenario, trips beginning on a peripheral bus contributed to walking distances of approximately 2.1 kilometers (23 minutes), which was the highest distance of all modes. Commuting by train is estimated in this scenario to contribute to 1.9 kilometers (20 minutes), while subway trips contributed 1.7 kilometers (18 minutes), city bus trips contributed 1.3 kilometers (15 minutes), and suburban bus trips contributed 1.4 kilometers (15 minutes).

Adjustment table

Comparing estimates of walking distances by mode obtained from Model 1 and 2, it is evident that the commuting flows scenario both over and under-estimates walking distance depending on first mode taken. Accordingly, differences in estimates between the two scenarios were calculated and are presented as adjustment values needed for practitioners and researchers who are estimating walking to and from public transport using the commuting flows scenario. Detailed OD scenario estimates of walking distance obtained from Model 2 closely resembled the estimates of walking obtained from Model 1 (within the 95% confidence interval), and therefore adjustment values between the two scenarios were based on estimates from Model 2.

The adjustment values (presented in Table 4) indicate that the commuting flows scenario most significantly miscalculates walking distance for trips that began on a bus in a peripheral region of Montreal, and most closely estimates trips that began on a subway. For bus trips beginning in a periphery region of Montreal, the commuting flows scenario overestimates walking by 133%. There are two potential explanations for this large discrepancy, is first due to the use of CT centroids as an origin. In peripheral regions the area of a CT is much larger (average area is 15.2 km^2) compared to a dense urban CT (average 1.2 km^2), and the centroid of that CT might be quite distant from where individuals residing in that CT who do use transit actually reside. The second potential explanation is related to the departure time of trips that were routed. We applied a rigid 8 AM departure time for all trips, and in peripheral regions where transit service is infrequent this could likely result in the selection of alternative routes that involve longer walking distances than are observed in reality.

Similarly, the commuting flows scenario overestimates walking distance of trips beginning on a suburban bus by 60%, a city bus by 37% and a subway by 8%. Interestingly, the error in which the commuting flows method overestimates walking distance appears to decrease in regions of higher density. For example, in dense CTs where subway stations are located (with an average area of 0.66 km^2), the commuting flows method is able to estimate walking distance with a relatively high degree of accuracy. Lastly, the commuting flows scenario underestimates average walking distances to commuter rail by 20%, although it is important to note that a very low number of trips (1%) were routed on the commuter train, whereas 6% of trips in the detailed OD data began on commuter train. The willingness of commuters to walk farther distances to use a faster and more comfortable service, such as rail, is not captured in the fastest route algorithm, and accordingly a higher proportion of trips use a bus to access a train station using this routing method.

Table 4 Adjustment table

First mode taken	Detailed OD scenario walking distance (m)	Commuting flows scenario walking distance (m)	Adjustment (m) †
Commuter train	2327.83	1856.98	470.85 (20%)
Subway	1556.60	1678.64	-122.04 (-8%)
City bus	976.01	1341.57	-365.57 (-37%)
Suburban bus	866.30	1382.20	-515.91 (-60%)
Peripheral bus	893.96	2079.96	-1186.00 (-133%)

† These adjustment values are based on the home to work walking distance, and would thus need to be multiplied by 2 if using total daily walking distance.

DISCUSSION AND CONCLUSIONS

The results indicate that walking distances estimated using the commuting flows scenario closely resemble actual walking distances derived from origin-destination microdata. Specifically, we find that estimating walking to public transport using the commuting flows method underestimates average walking distances by only 3%. However, the accuracy of the commuting flows method varies depending on the first mode of public transport used for each trip. Using a scenario analysis, we compare walking estimates for each mode and present adjustment values. These adjustment values must be applied to estimates of walking distance obtained using the commuting flows method to realistically uncover how walking to public transport contributes to daily physical activity goals and to estimate the number of individuals meeting the recommended levels of physical activity through the use of public transport. The adjustment value for commuter train users is 471 meters, indicating that the commuting flows scenario underestimates walking for train users. For the remaining modes of transport, the commuting flows scenario overestimates walking. Therefore, the following adjustment values should be subtracted from walking estimates obtained using this method: 122 meters for subway users, 366 meters for city bus users, 516 meters for suburban bus users and 1186 meters for peripheral bus users.

The accuracy with which the commuting flows method estimated walking distance for subway users, relative to other modes of transport, is likely explained by the density of CTs where a subway station is located. CTs are geographic areas that typically have a population between 2500 and 8000, and therefore the area of the CT is dependent on population density. In densely populated CTs the use of a centroid for the origin and destination will result in walking distances that more accurately resemble actual travel behaviour, compared to low density CTs in suburban and peripheral regions. Larger CTs typically are typically characterised with greater separation of land uses and dwellings are more dispersed. Therefore, the centroid of a CT may be considerably distant from where transit commuters reside.

Walking distances measured in this study echo previous research contributions that have presented evidence of the levels of physical activity that can be attained on a daily basis when commuting by public transport. Consistent with previous literature, commuting by train is associated with the highest levels of walking relative to all other public transport modes (*10; 14*). For effective public health interventions, it is important for municipalities to have a base level understanding of how many residents are commuting by public transport, and the daily physical activity levels that those commuters are attaining from their daily commute. While, detailed travel survey data is a highly effective source of data to measure walking levels, the collection of a representative sample of the population is very costly. Accordingly, the commuting flows method presented in this paper is particularly beneficial for planners and policy makers who wish to better understand regional levels of physical activity and to better understand local determinants of public transport use. One strength of this method lies in the usage of openly available GTFS data. GTFS data is provided by the majority of public transport providers across Canada and other developed countries. This data is regularly updated and reliable. The adjustment values presented from this comparative analysis will enable other municipalities to attain accurate estimates of daily walking to public transport.

Study Limitations

When calculating walking distance for both scenarios, the distance walked while transferring was not measured. Using the example of a transfer from a bus to a subway, accessing the subway platform from street level can take several minutes. Similarly, the access time to a train or subway platform was not measured, but under certain station designs can contribute to additional minutes of physical activity, particularly when individuals choose to take the stairs.

A limitation associated with measuring walking to commuter rail stations is that we do not know which trips an individual drove to a station and parked there, as only main mode is reported in the database. There are over 33,000 park-and-ride spaces across the Montreal commuter train network, and station access by park-and-ride is highly prevalent in suburban and peripheral regions of Montreal, while in higher density areas station access by other modes, walking, bicycle and transit are more common (*29*). Accordingly, walking distances of public transport trips that begin by park-and-ride are overestimated using this method.

ACKNOWLDEGMENTS

The authors wish to thank Mr. Gabriel Sicotte of the ARTM for providing the detailed Montreal OD survey. This work was supported by research grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Social Sciences and Humanities Research Council (SSHRC).

REFERENCES

[1] Canadian Society for Exercise Physiology. *Canadian 24-hour guidelines: An integration of physical activity, sedentary behaviour, and sleep*. http://csepguidelines.ca/. Accessed July 16, 2018.

[2] Public Health Agency of Canada. *Health Status of Canadians 2016: Report of the Chief Public Health Officer* https://www.canada.ca/en/public-health/corporate/publications/chief-public-health-officer-reports-state-public-health-canada/2016-health-status-canadians.html.

[3] Booth, M., A. Bauman, N. Owen, and C. Gore. Physical activity preferences, preferred sources of assistance, and perceived barriers to increased activity among physically inactive Australians. *Preventive Medicine,* Vol. 26, No. 1, 1997, pp. 131-137.

[4] Godin, G., R. Desharnais, P. Valois, L. Lepage, J. Jobin, and R. Bradet. Differences in perceived barriers to exercise between high and low intenders: Observations among different populations. *American Journal of Health Promotion,* Vol. 8, No. 4, 1994, pp. 279-285.

[5] Yoshida, K., K. Allison, and R. Osborn. Social factors influencing perceived barriers to physical exercise among women. *Canadian Journal of Public Health,* Vol. 79, No. 2, 1988, pp. 104-108.

[6] Owen, N., and A. Bauman. The descriptive epidemiology of a sedentary lifestyle in adult Australians. *International Journal of Epidemiology,* Vol. 21, No. 2, 1992, pp. 305-310.

[7] Borodulin, K., N. Sipilä, O. Rahkonen, P. Leino-Arjas, L. Kestilä, P. Jousilahti, and R. Prättälä. Socio-demographic and behavioral variation in barriers to leisure-time physical activity. *Scandinavian Journal of Public Health,* Vol. 44, No. 1, 2016, pp. 62-69.

[8] World Health Organization. *10 facts on physical activity*. http://www.who.int/features/factfiles/physical_activity/en/.

[9] Siegel, P., R. Brackbill, and G. Heath. The epidemiology of walking for exercise: implications for promoting activity among sedentary groups. *American Journal of Public Health,* Vol. 85, No. 5, 1995, pp. 706-710.

[10] Besser, L., and A. Dannenberg. Walking to public transit: Steps to help meet physical activity recommendations. *American Journal of Preventive Medicine,* Vol. 29, No. 4, 2005, pp. 273-280.

[11] Brown, B., and C. Werner. A new rail stop: Tracking moderate physical activity bouts and ridership. *American Journal of Preventive Medicine,* Vol. 33, No. 4, 2007, pp. 306-309.

[12] MacDonald, J., R. Stokes, D. Cohen, A. Kofner, and G. Ridgeway. The effect of light rail transit on body mass index and physical activity. *American Journal of Preventive Medicine,* Vol. 39, No. 2, 2010, pp. 105-112.

[13] Morency, C., M. Trépanier, and M. Demers. Walking to transit: An unexpected source of physical activity. *Transport Policy,* Vol. 18, No. 6, 2011, pp. 800-806.

[14] Wasfi, R., N. Ross, and A. El-Geneidy. Achieving recommended daily physical activity levels through commuting by public transportation: Unpacking individual and contextual influences. *Health & place,* Vol. 23, 2013, pp. 18-25.

[15] Veillette, M., R. Deboosere, R. Wasfi, N. Ross, and A. El-Geneidy. Small Steps, Big Differences: Assessing the Validity of using Home and Work Locations to Estimate Walking Distances to Transit. *Transportation research Record: Journal of the Transportation Research Record,* Vol. In press, 2018.

[16] Statistics Canada. Employed Labour Force 15 Years and Over Having a Usual Place of Work by Income Groups in 2010 - Commuting Flow Census Metropolitan Areas, Tracted Census Agglomerations and Census Tracts.In, 2011.

[17] Dora, C. A different route to health: Implications of transport policies. *BMJ,* Vol. 318, No. 7199, 1999, pp. 1686-1689.

[18] Koplan, J., and W. Dietz. Caloric imbalance and public health policy. *The Journal of the American Medical Association,* Vol. 282, No. 16, 1999, pp. 1579-1581.

[19] Fielding, G., R. Glauthier, and C. Lave. Performance indicators for transit management. *Transportation,* Vol. 7, No. 4, 1978, pp. 365-379.

[20] Zhao, F., L. Chow, M. Li, I. Ubaka, and A. Gan. Forecasting transit walk accessibility: Regression model alternative to buffer method. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1835, 2003, pp. 34-41.

[21] Murray, A., and X. Wu. Accessibility tradeoffs in public transit planning. *Journal of Geographical Systems,* Vol. 5, No. 1, 2003, pp. 93-107.

[22] Daniels, R., and C. Mulley. Explaining walking distance to public transport: The dominance of public transport supply. *Journal of Transport and Land Use,* Vol. 6, No. 2, 2013, pp. 5-20.

[23] El-Geneidy, A., M. Grimsrud, R. Wasfi, P. Tétreault, and J. Surprenant-Legault. New evidence on walking distances to transit stops: Identifying redundancies and gaps using variable service areas. *Transportation,* Vol. 41, No. 1, 2014, pp. 193-210.

[24] Weinstein Agrawal, A., M. Schlossberg, and K. Irvin. How far, by which route and why? A spatial analysis of pedestrian preference. *Journal of urban design,* Vol. 13, No. 1, 2008, pp. 81-98.

[25] O'Sullivan, S., and J. Morrall. Walking distances to and from light-rail transit stations. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1538, 1996, pp. 19-26.

[26] World Health Organization. Global reccomendations on physical activity for health.In, World Health Organization, Switzerland, 2010.

[27] Agence Métropolitaine de Transport. Enquête origine-destination 2013.In, Montreal, Quebec, 2013.

[28] Farber, S., M. Morang, and M. Widener. Temporal variability in transit-based accessibility to supermarkets. *Applied Geography,* Vol. 53, 2014, pp. 149-159.

[29] Steer Davies Gleave. Transport Collectif Train de l'Ouest de Montreal.In, Montreal, 2016.