JTLU

Does travel behavior matter in defining urban form? A quantitative analysis characterizing distinct areas within a region

Cynthia Jacques [a]

Novatech Engineering Consultants Ltd.

Ahmed M. El-Geneidy [b]

McGill University

Abstract: Research that attempts to characterize urban form is confronted with two key issues: criticism of the use of aggregate units of analysis, such as census tracts, and a general lack of consideration of variables related to elements other than the built environment, such as residents' behavior. This methodological study explores the impact of travel behavior variables in the quantitative characterization of urban form at the census tract level for the Montreal region. Two separate factor-cluster analyses are performed: the first includes built-environment variables commonly used to typify areas within a region, and a second includes additional travel behavior variables. The results of both models are compared to satellite images to determine which analysis more accurately represents the reality on the ground. The results provide empirical evidence that travel behavior variables, in addition to built form, provide a more accurate representation of urban form at the census tract level. These variables refine the model output by moderating the effect of features that generally led to misleading results. This effect is particularly evident in areas represented by large census tracts. These results suggest that considering both built environment and behavioral characteristics in an analysis of urban form yields more precise results at the (aggregate) census tract level. The findings from this study could be helpful for engineers and planners when conducting property value studies, urban investment analysis, and policy intervention prioritization and when expanding the well-known land use classification of urban and rural categories.

1   Introduction

Cities and regions are inherently complex and dynamic systems shaped by a multitude of factors and constantly changing over time. A growing body of literature seeks to develop methodologies for characterizing urban form to identify distinct areas and different neighborhood types within a city; however, this research is confronted with two key issues. For one, the appropriate unit or scale at which to measure urban form to characterize distinct areas has been debated, with some scholars criticizing the use of aggregate or pre-defined boundaries, such as census tracts or traffic analysis zones (TAZs), as being too large and yielding misleading classification. Second, the majority of existing studies that attempt to characterize urban form focus only on elements of the built environment, ignoring other elements that might define an area. But because (stationary) elements of the built environment are not the only aspects that characterize a neighborhood or region, Kevin Lynch claims that "[m]oving elements of a city, and in particular the people and their activities, are as important as the stationary parts" (1960, p. 2). Examples of studies incorporating behavioral considerations of local residents to characterize areas within a region are limited to date.

This study performs two separate factor-cluster analyses to examine the impact of including travel behavior considerations in the methodology for characterizing distinct area types at the census tract level. The first analysis considers variables related

to density, land use mix, street design, accessibility, housing, and demographics. The second analysis includes all of the same variables as the first, as well as additional variables related to travel behavior (e.g., mode share, travel distance, vehicle ownership). The purpose is to address the following research question: Does the inclusion of travel behavior considerations provide a more accurate characterization of distinct areas within a region at the census tract level than the consideration of built environment measures alone? An accurate neighborhood classification could have various uses by engineers and planners. For example, it could be included in predictions of property values. Planners and engineers strive to implement different policies and investments in specific neighborhoods; having a better classification could help in directing these investments or policies, such as prioritizing areas where to implement water service upgrades, public transit improvements, and land use zoning changes. In addition, such classification could add new layers to the widely used land use classification of "urban and rural." Such an addition would be helpful for neighborhood design interventions, since urban designers are challenged to identify areas where interventions are needed in a region.

The Montréal Metropolitan Region makes for an interesting case study for this analysis as it is among the largest, most diverse cities in Canada. The Montréal region is made up of 82 municipalities on the Island of Montréal itself and the surrounding shores, and it is served by a robust bus and metro system as well as a commuter train system that extends

[a] c.jacques@novatech-eng.com [b] ahmed.elgeneidy@mcgill.ca

into the more remote areas of the region. Moreover, in recent years there has been increasing development off-island on the north and south shores, particularly along major expressways. The authors' familiarity with the region also facilitates the assessment of model performance in characterizing distinct area types within the region.

To begin, this study provides a review of the existing literature related to the characterization of distinct urban areas or neighborhoods, as well as the relationship between urban form and travel behavior. Next, the methodology and data used in this analysis are described, followed by a presentation of the results of two separate factor-cluster analyses (with and without travel behavior measures). This is followed by a discussion of how the results of these two models differ and what impact the inclusion of travel behavior might have on the classification of distinct areas within a region in general and at the aggregate census tract level. The paper concludes by noting the implications that these results have for urban form and characterization research.

2   Literature Review

The factors that shape and define cities are numerous and complex. Song & Knaap (2004) note that neighborhoods are often referred to as either "traditional" or "suburban sprawl"; however, this simplistic approach fails to capture the great variation in physical form of metropolitan areas. Bagley et al. (2002) assert that "restricting the designation of an entire neighborhood to one of two discrete types either results in discarding considerable data (for "hybrid" neighborhoods) or distorting the subsequent analysis (through misclassification)" (M. Bagley et al. 2002, p. 690). Better characterization of distinct areas within a region is important for informing and directing land use and transportation policy, allowing for effective policy evaluation, and facilitating our understanding of various ubiquitous yet ambiguous terms, such as "sprawl." "suburban," and "mixed use" (Manaugh, Miranda-Moreno, and El-Geneidy 2010; Song & Knaap 2007). As such, several researchers have explored more comprehensive approaches to characterizing distinct areas within a region.

Talen (2002) describes the "transect" approach, which defines a rural-to-urban continuum that is divided into six distinct area types with varying degrees of development intensity: rural preserve, rural reserve, sub-urban, general urban, urban center, and urban core. However, no quantitative method for classifying areas as these six distinct transects has been developed to date (Song & Knaap 2007). Other researchers have identified different area types within a region using qualitative methods such as maps, images, interviews, and site visits (for

example, Southworth & Owens 1993; Wheeler 2003). There is, however, a growing body of literature that seeks to develop quantitative methods for classifying distinct areas within a region.

Handy (1996) used qualitative (maps, photos, site visits) and quantitative (analysis of geographic data) methods to identify differences between three neighborhood types in Austin, Texas: traditional, early-modern, and late-modern neighborhoods. In their examination of the effects of neighborhood characteristics and other factors on commuting distances, Manaugh et al. (2010) identified different neighborhood typologies for the Montreal region through a factor-cluster analysis, using urban form, sociodemographic, and accessibility variables. The outcome of the analysis is the classification of typologies at both home and job locations. Bagley et al. (2002) performed a factor analysis on 18 variables (including urban form, socio-demographic, and attitudinal variables) for five San Francisco Bay Area neighborhoods. They identified two factors (traditional and suburban) and found that study neighborhoods scored highly on both factors, suggesting that within a single neighborhood there can be characteristics that are more urban in nature and other elements that are more suburban. Song and Knapp (2007) developed a quantitative methodology for characterizing neighborhood types using a factor-cluster analysis on numerous variables that define the built environment. The authors applied this methodology to the neighborhoods of new single-family homes within the Portland Metropolitan Region, identifying six distinct neighborhood types: Sporadic Rural Developments, Bundled Rural Developments, Outer Ring Suburbs, Downtown, Inner and Middle Ring Suburban Redevelopments/Infill, Composite Greenfields, and Partially Clustered Greenfields.

Most studies that include transportation considerations in a model for characterizing distinct area types within a region use variables related to the provision of transportation infrastructure and service, such as the supply of bike lanes and sidewalks and the number of bus stops, among other measures (for example, Manaugh et al. 2010; Song & Knaap 2007). But these measures are still related to the built environment. Bagley et al. (2002) went into more depth by including attitudinal and travel-related data collected through mail-out surveys and travel diaries completed by residents of the examined neighborhoods, capturing residents' perception of whether a neighborhood is pleasant for walking or cycling, whether transit service is good, whether enough parking is provided, and so on. Rodriguez et al. (2006) included census commuting mode share data to capture travel behavior (in addition to built environment measures) in the calculation of a built environment index for walkability, using three different methods. While the cluster

analysis performed in their study identified three distinct areas (urban, suburban and exurban), the main purpose was to compare the results of the different analytical methods rather than to identify area types.

Examples of studies that include travel behavior in a model for characterizing distinct area types within a region are quite limited to date, despite much literature on the relationship between travel behavior and the built environment (R. Cervero & Kockelman 1997; R. Cervero & Radisch 1996; Ewing & Cervero 2001; Frank & Pivo 1994; Næss 2011; Niemeier, Bai, & Handy 2011, among others). Song & Knaap (2007) point out that the main purpose of quantitative studies trying to classify neighborhoods is motivated by the desire to utilize these classifications in travel behavioral research or physical activity. Thus, travel behavior is usually considered after the distinct areas have been identified, to examine how travel behavior differs amongst these areas.

There is, however, a growing body of literature suggesting that built form factors may have little influence on travel behavior relative to other considerations such as residential location choices and individual attitudes (M. Bagley & Mokhtarian 2002; Boarnet & Crane 2001; Cao, Handy, & Mokhtarian 2006; Crane & Crepeau 1998; Kitamura, Mokhtarian, & Laidet 1997). The theory of residential self-selection suggests that an individual might choose to live in a certain neighborhood because it allows him or her to travel in a certain way (e.g., a neighborhood that is walkable). In such a case, the influence that the built environment has on travel behavior is secondary to the influence that it has on the individual's choice to live there. Nevertheless, some researchers have found that the built environment has an impact on travel behavior even after accounting for attitudes and lifestyle considerations (Handy, Cao, and Mokhtarian 2007; Schwanen and Mokhtarian 2005).

The studies discussed above indicate a growing rationale that built environment considerations alone are not enough to fully understand and identify distinct area types within a city or region. Emerging research hypothesizes that "the 'character' of an urban area is defined not just by the types of places found there, but also by the people that choose to make that area part of their daily life" (Cranshaw, Schwartz, Hong, & Sadeh 2012, p. 1). Using location-based social media data, the authors developed an algorithm that maps distinct geographic areas of the city based on the behavior of residents. The resulting clusters are called "Livehoods," which reflect the dynamic nature of activity patterns of city residents.

In addition to the lack of attention to non-built-environment variables in defining urban form, several researchers also note issues in choosing the appropriate scale at which

to measure variables to characterize distinct areas (Manaugh et al. 2010; Moudon, Hess, Snyder, and Stanilov 1997; Song and Knaap 2007; Srinivasan 2002; Talen 2003). Among these researchers, there is concern about characterizing urban form using aggregate units, such as census tracts, TAZs, or other predefined neighborhoods. Song and Knaap (2007) note that the size of the unit of analysis can lead to a misleading measure, since some of these areas are too large, ignoring the fact that land use characteristics and patterns do vary across different sections of the predefined neighborhood. This issue is particularly concerning, since urban planners generally "[rely] principally on just two sources of data about urban form, both of which are aggregate: land use zone maps and census geography statistics" (Talen 2003, p. 198). Thus, there is also a need to explore ways to improve the characterization of urban form at aggregate levels, to enhance research that is limited to such units of analysis or at least understand the bias imposed by using only the built form characteristics in such aggregate level of analysis.

3 Methodology and Data

Drawing from the methodology used in similar studies (Manaugh et al. 2010; Song and Knaap 2007), this study employs a factor-cluster analysis on numerous urban form measures to characterize different areas within the Montreal Metropolitan Region. Factor analysis is a commonly used data reduction technique; it reduces a dataset of many correlated variables to a smaller, more manageable set of factors that correspond to a significant portion of the variability of the full dataset (Rogerson 2010). Cluster analysis is then performed on the resulting factor scores to group together census tracts that display similar characteristics. An advantage of cluster analysis is that it identifies clusters regardless of spatial location (Song and Knaap 2007).

To test the effect that the inclusion of travel behavior has on the characterization of areas within a region, two separate factor-cluster analyses are performed: the first including variables commonly used to typify areas within a region, and a second analysis that includes the same variables as the first analysis as well as additional travel behavior variables. Census tracts were chosen as the unit of analysis in both models so that characterization improvements, if found, can inform discussion on the use of behavior as a means of improving analysis at this imperfect but important level of aggregation. A total of 872 census tracts represent the Montreal region; 18 were excluded from this analysis as there were no data available for most of the variables (e.g., census tracts for Parc Lafontaine, Parc Angrignon, and Mohawk Nation Reserve, among others).

The variables included in this analysis were derived from existing studies that attempt to characterize neighborhood or urban area types discussed in the previous section. Variables most commonly used to typify distinct areas within a region relate to density, land use, street network design, accessibility, housing, and demographics. A total of 35 such variables were generated for each census tract in the Montreal region using various applications of Geographic Information Systems (GIS) software. An additional 14 variables related to travel behavior were also generated for each census tract observed in this analysis and were included in the second factor-cluster analysis. These variables were derived from the 2008 Origin-Destination (OD) survey for the Montreal Metropolitan Region and include mode share and travel distances for five different modes (private vehicle, walking, cycling, transit, and park-and-ride). Travel distances were derived using GIS software, with trip distances measured along the street network. Table 1 lists each variable included in these analyses, indicating the unit of measure, data source, and summary statistics for the census tracts.

Satellite images were used to determine the number of clusters that most accurately represents the spatial distribution of the outcomes of each analysis conducted in this study as well as to compare the results of the analysis with and without travel behavior considerations. The authors' knowledge of the study area also helped in this regard.

4 Analysis without travel behavior

4.1 Factor analysis

The first factor analysis was performed using all of the variables listed in Table 1 except for the travel behavior variables. With the eigenvalue set equal to one and principal component extraction and varimax rotation applied, the factor analysis extracted seven factors characterizing physical urban form. Of the initial 35 variables included in the analysis, 30 variables were found to significantly explain the variation in the model (Government and Institutional, Commercial, Total Number of Intersections, Percent of Small Blocks, and Year of Construction 1981–1990 were not significant). The factor analysis was repeated several times with the number of factors set to six, five, four and three factors, but the seven-factor setting was found to yield the best results. The seven extracted factors reproduce about 82 percent of the total variation among the 30 variables. The seven factors resulting from this first factor analysis are described below.

Accessible Density: Factor 1 represents accessibility to jobs, stores, food, and restaurants by car and transit, as well as high population density and pre-1946 housing, with negative loadings for single-detached housing and distance to the central business district (CBD).

Large Wealthy Households: Factor 2 has positive loadings for the average number of rooms and bedrooms, the average number of people per household, owner-occupied dwellings, and the median household income. It has negative loadings for apartments and rented dwellings.

Large Open Areas: Factor 3 is characterized by large blocks and cul-de-sacs surrounded by open space and few residential uses.

Recent Housing: Factor 4 represents housing constructed after 1990, with positive loadings for housing constructed from 1991 onward.

Post-War Housing: Factor 5 represents post-war housing, with a positive loading for dwellings construction between 1946 and 1960.

Industrial Lands: Factor 6 captures industrial and resource land uses.

Park Lands: Factor 7 is characterized by a positive factor loading for parklands.
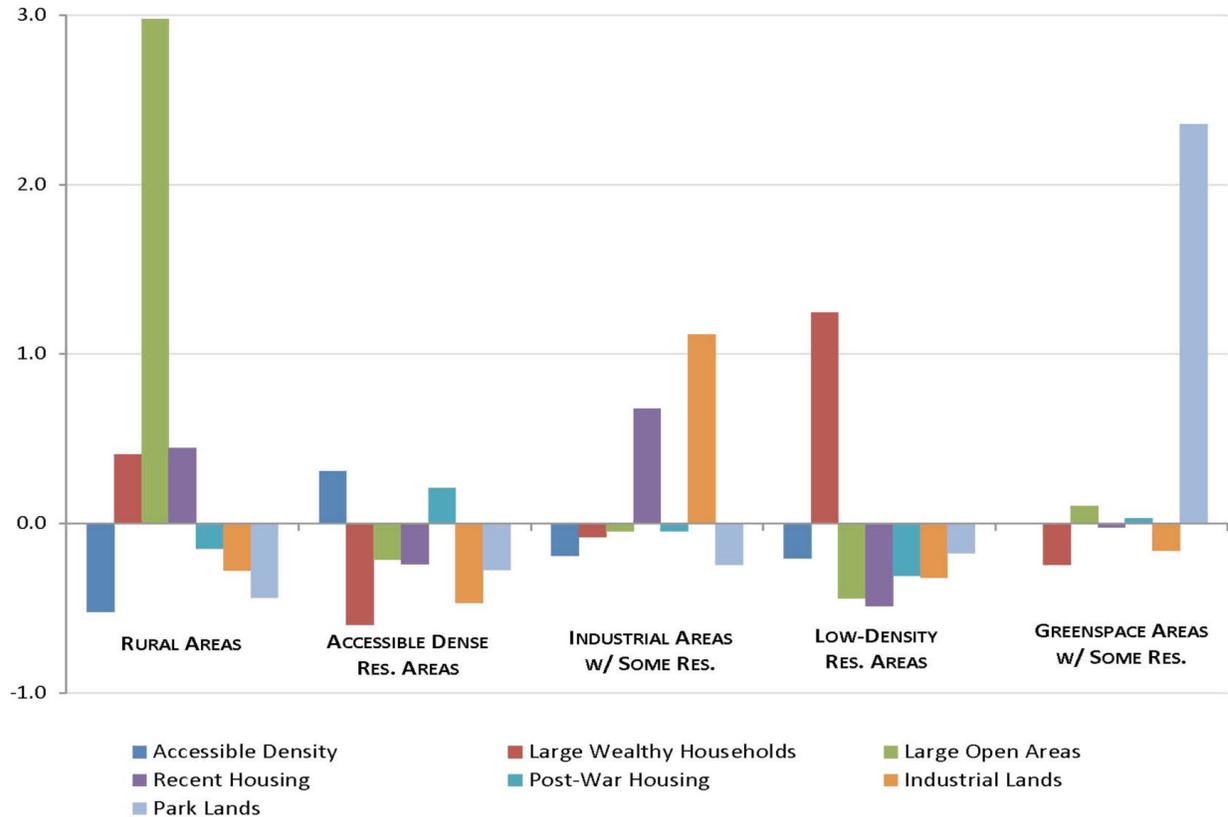
4.2 Cluster analysis

Using the seven factors from the first factor analysis (which excluded transportation variables), a k-means cluster analysis was performed to identify groups of census tracts with similar characteristics, regardless of their location within the region. The resulting clusters identify distinct area types within the region such that the similarities for instances of the same area type are maximized, while the similarity between the different area types is minimized (Manaugh et al. 2010). The optimal number of cluster groupings was found to be five. A larger number of clusters resulted in too few census tracts falling into a cluster by themselves in a way that did not make logical sense, while a smaller number of clusters resulted in clusters that did not adequately distinguish areas of known distinct urban form. The centroid values are presented in Figure 1. The name and a brief description of each cluster resulting from this analysis are provided below.

Rural Areas: Cluster 1 has a strong positive input for the large open areas factor (which includes agricultural land and natural features such as forests) that clearly dominates all other factors. It also has positive inputs for large wealthy households and recent housing and negative inputs for all other factors, including accessible density. This cluster accounts for 6 percent of census tracts.

Table 1: Unit, source, and summary statistics for all variables.

Variable	Unit	Source	Min	Max	Mean	Std. Dev.
Density						
Population Density*	Persons/sq. km	2006 Census	272.56	45344.71	6662.41	5233.28
Average Number of People per Household	Count	2006 Census	2	3.8	2.86	0.25
Average Number of Rooms	Count	2006 Census	2.8	10.1	5.56	1.26
Average Number of Bedrooms	Count	2006 Census	0.9	4.2	2.34	0.57
Single Detached Homes	Percent of CT Dwellings	2006 Census	0	100	30.92	32.66
Apartments	Percent of CT Dwellings	2006 Census	0	100	60.71	35.28
Land Use						
Residential	Percent of CT	DMTI	0.23	99.74	56.91	24.33
Commercial	Percent of CT	DMTI	0	51.05	2.92	5.16
Industrial and Resource	Percent of CT	DMTI	0	95.39	12.21	15.98
Government and Institutional	Percent of CT	DMTI	0	78.62	5.55	8.78
Park	Percent of CT	DMTI	0	64.51	5.61	8.95
Open Space	Percent of CT	DMTI	0	98.86	16.24	24.77
Street/Block Design **						
Average Block Length	Metres	DMTI	67.47	1196.07	148.85	79.28
Average Block Size	Square kilometres	DMTI	0.004	2.49	0.06	0.16
Percentage of Small Blocks (<0.016 sq. km or smaller)	Percent	DMTI	0	100	52.97	19.28
Number of Intersections	Count	DMTI	3	697	84.02	74.53
Number of Cul-de-Sacs	Count	DMTI	0	321	9.06	18.19
Accessibility						
Distance to the CBD	Average	DMTI	299.5	71004.28	16243.64	12026.61
Access to Jobs by Car (gravity-based)	Average	El-Geneidy et al. 2011	1291.85	103270.66	33364.83	25260.29
Access to Jobs by Transit (gravity-based)	Average	El-Geneidy et al. 2011	228.32	56663.46	12845.33	11210.1
Access to Big Box Stores by Car (30 minutes)	Average	El-Geneidy et al. 2011	4.5	164	86.25	41.75
Access to Food Stores by Car (30 minutes)	Average	El-Geneidy et al. 2011	53.44	3121.75	1477.58	1034.84
Access to Restaurants by Car (30 minutes)	Average	El-Geneidy et al. 2011	89.33	5552.88	2522.63	1853.3
Access to Food Stores by Transit (45 minutes)	Average	El-Geneidy et al. 2011	2.67	1866.8	650.7	566
Access to Restaurants by Transit (45 minutes)	Average	El-Geneidy et al. 2011(El-Geneidy, Cerdá, Fischler, & Luka 2011)	1	3553.2	1323.45	1177.27
Housing/Demographics						
Owner-Occupied Dwellings	Percent	2006 Census	0	100	52.34	27.67
Rented Dwellings	Percent	2006 Census	0	100	47.66	27.67
Median Household Income	CAD	2006 Census	13024	262458	51173.49	22271.9
Year of Construction Pre-1946	Percent	2006 Census	0	86.76	15.74	20.15
Year of Construction 1946-60	Percent	2006 Census	0	81.56	19.6	16.55
Year of Construction 1961-70	Percent	2006 Census	0	83.13	17.83	13.25
Year of Construction 1971-80	Percent	2006 Census	0	73.8	16.24	12.81
Year of Construction 1981-90	Percent	2006 Census	0	77.93	15.35	13.34
Year of Construction 1991-2000	Percent	2006 Census	0	69.63	9.01	10.96
Year of Construction 2001-2006	Percent	2006 Census	0	56.49	6.04	8.71
Travel Behavior***						
Walking Mode Share	Percent	AMT 2008	0	56.41	13.69	10.23
Cycling Mode Share	Percent	AMT 2008	0	25.00	1.88	2.97
Transit Mode Share	Percent	AMT 2008	0	60.00	23.53	10.50
Park-and-Ride Mode Share	Percent	AMT 2008	0	11.76	1.52	1.67
Private Vehicle Mode Share	Percent	AMT 2008	6.67	95.45	59.37	17.90
Walking Travel Distance	Average	AMT 2008	0	10335.32	1238.35	885.27
Cycling Travel Distance	Average	AMT 2008	0	71013.23	3137.58	5343.4
Transit Travel Distance	Average	AMT 2008	0	29385.95	8694.46	3835.42
Transit Travel Distance Standard Deviation	Average	AMT 2008	0	25089.89	6117.43	3552.37
Park-and-Ride Travel Distance****	Average	AMT 2008	0	57051.53	12823.7	12825.73
Private Vehicle Travel Distance****	Average	AMT 2008	1370.33	24719.8	9466.02	3678.23
Private Vehicle Travel Distance Standard Deviation	Average	AMT 2008	0	30278.4	9366.32	3722.04
Number of Vehicles per Household	Average	AMT 2008	0.22	3.42	1.31	0.49
Number of Bus Stops in CT	Count	AMT 2008 (AMT 2008)	0	9148	331.43	718.94

Figure 1: Cluster centroid values from the analysis excluding travel behavior variables.

Accessible Dense Residential Areas: Cluster 2 has the strongest positive input for accessible density and has negative inputs for large wealthy households, open space, recent housing, industry, and parks. This grouping accounts for 39 percent of all census tracts included in the analysis.

Industrial Areas with Some Residential: Cluster 3 is characterized by a strong positive input for industrial lands, as well as some recent housing (post-1990), with negative inputs for all other factors. This cluster represents 25 percent of census tracts.

Low-Density Residential Areas: Cluster 4 is dominated by a strong positive input for large wealthy households and has negative inputs for all other factors, including accessible density. This group accounts for 20 percent of census tracts.

Greenspace Areas with Some Residential: Cluster 5 is dominated by the factor for parklands (which includes golf courses) and has relatively little input for all other factors. This cluster accounts for 10 percent of census tracts in the analysis.
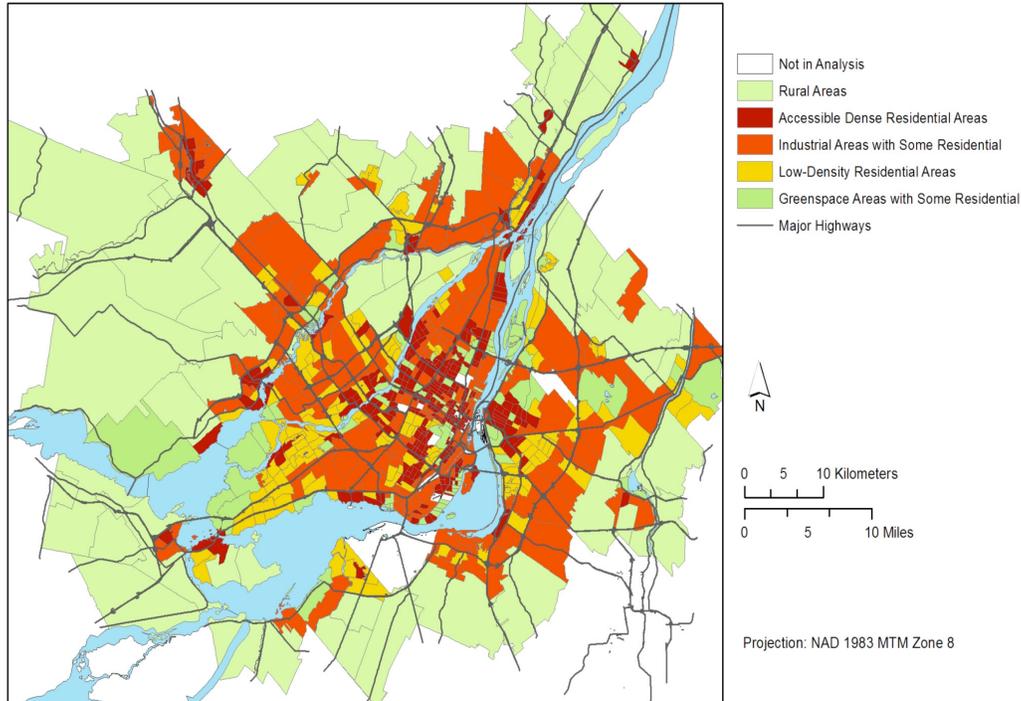
The spatial distribution of the clusters resulting from the factor-cluster analysis excluding travel behavior variables is presented in Figure 2.

5 Analysis with travel behavior

5.1 Factor analysis

The second factor analysis was performed using all 30 of the variables that were found to be significant in the first factor analysis, as well as the additional 14 travel behavior variables listed in Table 1. With the eigenvalue set equal to one and principal component extraction and varimax rotation applied, the factor analysis on these 44 variables extracted eight factors characterizing urban form. Of the initial 44 variables included in the analysis, 42 variables were found to significantly explain the variation in the model (Park-and-Ride Mode Share and Walking Average Travel Distance were not significant). The factor analysis was repeated several times with alternate factor counts selected, but eight-factor setting produced the best results. The eight extracted factors reproduce about 77% of the total variation among the 42 variables. The eight factors resulting from this second factor analysis are described below.

Accessible Walkable Density: Factor 1 represents accessibility to jobs, stores, food, and restaurants by car and transit, as well as high population density, apartments, pre-1946 housing, transit

Figure 2: Spatial distribution of area-type clusters resulting from the analysis, excluding travel behavior variables.

mode share, and walking mode share. This factor has negative loadings for single-detached housing, distance to the CBD, private vehicle mode share, and travel distance variables, as well as for the average number of vehicles per household.

Large Wealthy Households: Factor 2 has positive loadings for the average number of rooms and bedrooms, the average number of people per household, owner-occupied dwellings, and the median household income. It has a negative loading for rented dwellings.

Large Open Areas: Factor 3 is characterized by large blocks, cul-de-sacs, open space, and few residential uses.

Recent Housing: Factor 4 represents housing constructed after 1990, with positive loadings for housing constructed from 1991 onward.

1970s Bus Stops: Factor 5 has a positive factor loading for number of bus stops and for housing constructed between 1971-1980, and a negative loading for housing from 1946 to 1960.

*Cycling:* Factor 6 has positive factor loadings for cycling mode share and cycling travel distance.

No Industry: Factor 7 has a negative factor loading for industrial and resource land uses.

Park Lands: Factor 8 is characterized by a positive factor loading for parklands.

5.2 Cluster analysis

Using the eight factors derived from the factor analysis including transportation variables, a k-means cluster analysis was performed to identify groups of census tracts with similar characteristics. The optimal number of cluster groupings extracted from the eight factors was found to be six. Although the use of six clusters resulted in one cluster that was relatively small (only five census tracts), the overall output from this number of clusters was the most representative of the reality on the ground, and the five census tracts in a single cluster made logical sense. A smaller number of clusters resulted in clusters that did not adequately distinguish areas of known distinct urban form. The centroid values are presented in Figure 3. The name and a brief description of each cluster resulting from this analysis are provided below.

Accessible Dense Residential Areas: Cluster 1 is characterized by relatively strong positive inputs for the accessible walkable density and no industry factors, and a relatively strong negative input for large wealthy households, as well as for large open areas and parks. Therefore these areas are likely the most dense, walkable areas in the region. This cluster accounts for 39% of census tracts.
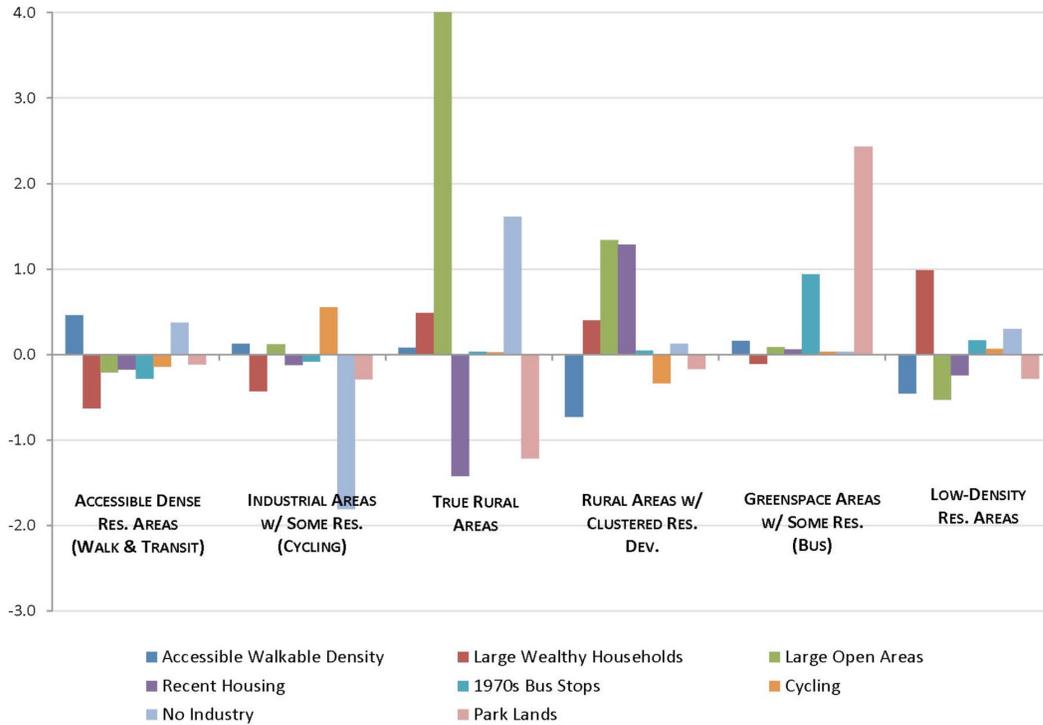
Figure 3: Cluster centroid values from the analysis including travel behavior variables.

Industrial Areas with Some Residential (Cycling): Cluster 2 is dominated by a strong negative input for no industry, indicating that industry is present in these areas. This cluster also has a notable positive input for the cycling factor (cycling mode share and cycling travel distance), indicating that individuals who live in these areas tend to cycle for many trips and may cycle longer distances. This grouping represents 14% of census tracts.

True Rural Areas: Cluster 3 is dominated by a strong positive input for large open areas (which includes agricultural and natural lands such as forests), and a strong negative input for recent housing and parklands. It also has a positive input for no industry, and relatively small positive inputs for accessibility and large wealthy households, but these are negligible compared to the input for large open spaces. Therefore, this cluster represents true rural areas with relatively little to no recent housing. This group accounts for only 0.6% of census tracts.

Rural Areas with Clustered Residential Development: Cluster 4 has strong positive inputs for large open areas and recent housing, a positive input for large wealthy households, and a relatively strong negative input for accessible walkable density. Therefore, this cluster identifies recent residential developments abutting rural lands. This cluster accounts for 12% of census tracts.

Greenspace Areas with Some Residential (Bus-Served): Cluster 5 has a strong positive input for parks areas, as well as a positive input for 1970s housing developments and number of bus stops. This grouping represents 7.7% of census tracts.

Low-Density Residential Areas: Cluster 6 has a relatively strong input for large wealthy households, and negative inputs for accessible walkable density, large open areas and park lands, as well as a small negative input for recent housing (post 1990). This group accounts for 26.7% of the census tracts in the analysis.

The spatial distribution of the clusters resulting from the factor-cluster analysis including travel behavior variables is presented in Figure 4.

6 Comparing the Two Models

In comparing the nature of the clusters resulting from each analysis, there does not seem to be a very large difference between the two models. Although the types of areas identified in both analyses are quite similar, the degree of difference between the two models becomes immediately evident when examining the maps of these clusters' spatial distribution throughout the Montreal region (Figure 2 without travel behavior and Figure 4 with travel behavior). In addition to the identification of a sixth
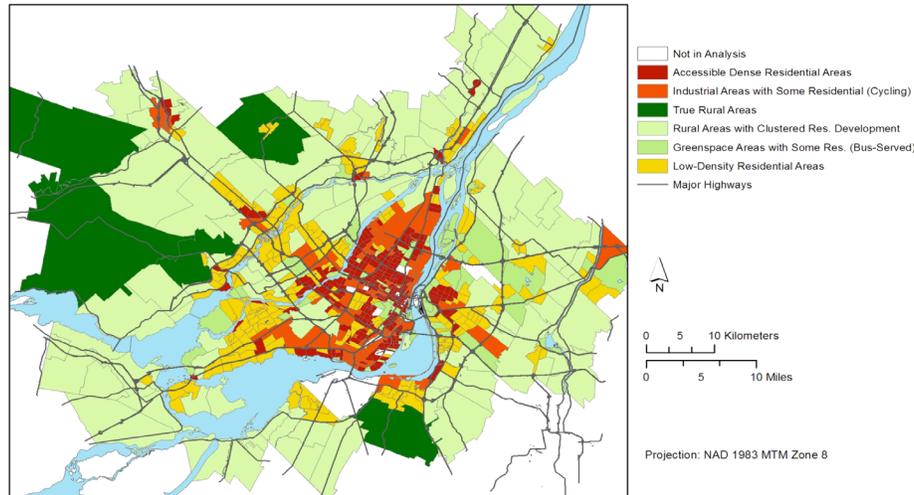
Figure 4: Spatial distribution of area-type clusters resulting from the analysis including travel behavior variables.

cluster in the model with travel behavior, the overall distribution of the clusters is very different.

The main difference in the types of clusters that were generated from these analyses is that the model with travel behavior yielded two distinct clusters representing the rural areas defined in the first analysis. These two new categories (True Rural Areas and Rural Areas with Clustered Residential Development) provide a clearer distinction between the rural areas in the Montreal region that remain virtually untouched by new residential development (True Rural Areas), and the rural areas that are abutted by recent (i.e., since the 1990s) residential development clusters. This second rural group represents areas that are already, or are at the most risk of, being taken up by development. It should be noted that six clusters were also tested for the model without travel behavior, but the resulting clusters were not as accurate to the reality on the ground as were the results with only five clusters.

Another notable difference between the two models for characterizing distinct types of urban form within the Montreal region is the extent of the cluster representing Industrial Areas with Some Residential. Although this area type is present in both models, the extent of this area type expands far beyond the Island of Montreal itself, roughly following the lines of the major highways, in the model excluding travel behavior (Figure 2). Although the industrial area type is still present in the model with travel behavior (Figure 4), it is far more concentrated and more accurately represents the extent of industrial areas in the city (e.g., parts of Anjou, Montreal-Est, Dorval, Lachine, Lasalle, and the Sud-Ouest).

The additional areas classified as Industrial Areas with Some Residential in the model without travel behavior have mostly been replaced by Rural Areas with Clustered Residential Development in the model with travel behavior. In examining the factor inputs for the clusters with travel behavior (Figure 3), the inclusion of travel behavior has resulted in the Industrial Areas with Some Residential having slightly positive inputs for accessible walkable density (walking and transit) and cycling. In contrast, Rural Areas with Clustered Residential Development have negative inputs for accessible walkability and cycling. This is in contrast to the model without travel behavior (Figure 1), where both the Rural Areas and the Industrial Areas clusters had negative inputs for accessibility. It is logical that walking, cycling, and transit use are more likely to occur at higher rates closer to the urban core; thus it makes sense that in the model with travel behavior, the Industrial Areas with Some Residential cluster is more limited to areas on or just slightly off the island. It appears as though the effect of industrial uses is being moderated by the travel behavior information, which is picking up more on the residential aspect of the cluster (i.e., the area in the census tract is no longer being identified as industrial if only a small portion of the cluster is industrial; the travel behavior information moderates the impact of these features to give a better overall representation of the census tract).

It is interesting to take a closer look and to compare the results to the reality on the ground. Using satellite images and the authors' knowledge of the region, it was found that the model including travel behavior resulted in a distribution of typologies that are far more accurate than the model without travel behavior. A couple of examples are provided herein.

Figure 5 provides a detailed look at a small portion of the northwestern part of Laval and the southwestern part of the North Shore (around Saint-Eustache), comparing the results of the two models with a satellite image of the area to determine which model more accurately represents the different area types within this "snapshot." The resulting area characterization is quite different between the two models. The distribution of

Low-Density Residential Areas is much more extensive in the model with travel behavior, while the extent of the Accessible Dense Residential Areas is much more restricted, which more accurately characterizes this rather suburban area. The dark accessible dense strip just south of the highway connection at the north of the site in the model with travel behavior is picking up on a commercial strip with nearby residential, thus emphasizing the increased accessibility of this small strip in terms of access to shops (residents living nearby could easily walk to a commercial area). The Accessible Dense Residential Area at the southern tip of the North Shore is picking up on the Grand-Moulin commuter train station that offers accessibility by transit and is surrounded by denser (although still mostly single-family) residential development on narrower streets. Thus, it would appear that the inclusion of travel behavior has moderated the impact of these commercial and rail station elements, lessening the extent of areas represented by the accessible dense cluster and allowing for a more accurate representation of the surrounding residential areas as low-density.

As noted in the general comparison of the two models, the east, north, and southwest portions of this snapshot that were classified as Industrial Areas with Some Residential in the model without travel behavior become classified as Rural Areas with Clustered Residential Development in the model with travel behavior. This new classification captures areas with clustered residential subdivisions abutting agricultural and forested land (visible along the eastern edge of the satellite image). The presence of some car dealerships directly adjacent to the highway in the north of this snapshot was likely exaggerating the extent of industrial and resource uses in the model without travel behavior. Finally, the northwestern portion of the snapshot becomes distinguished as True Rural in the model with travel behavior, more accurately capturing the most rural area in this snapshot.
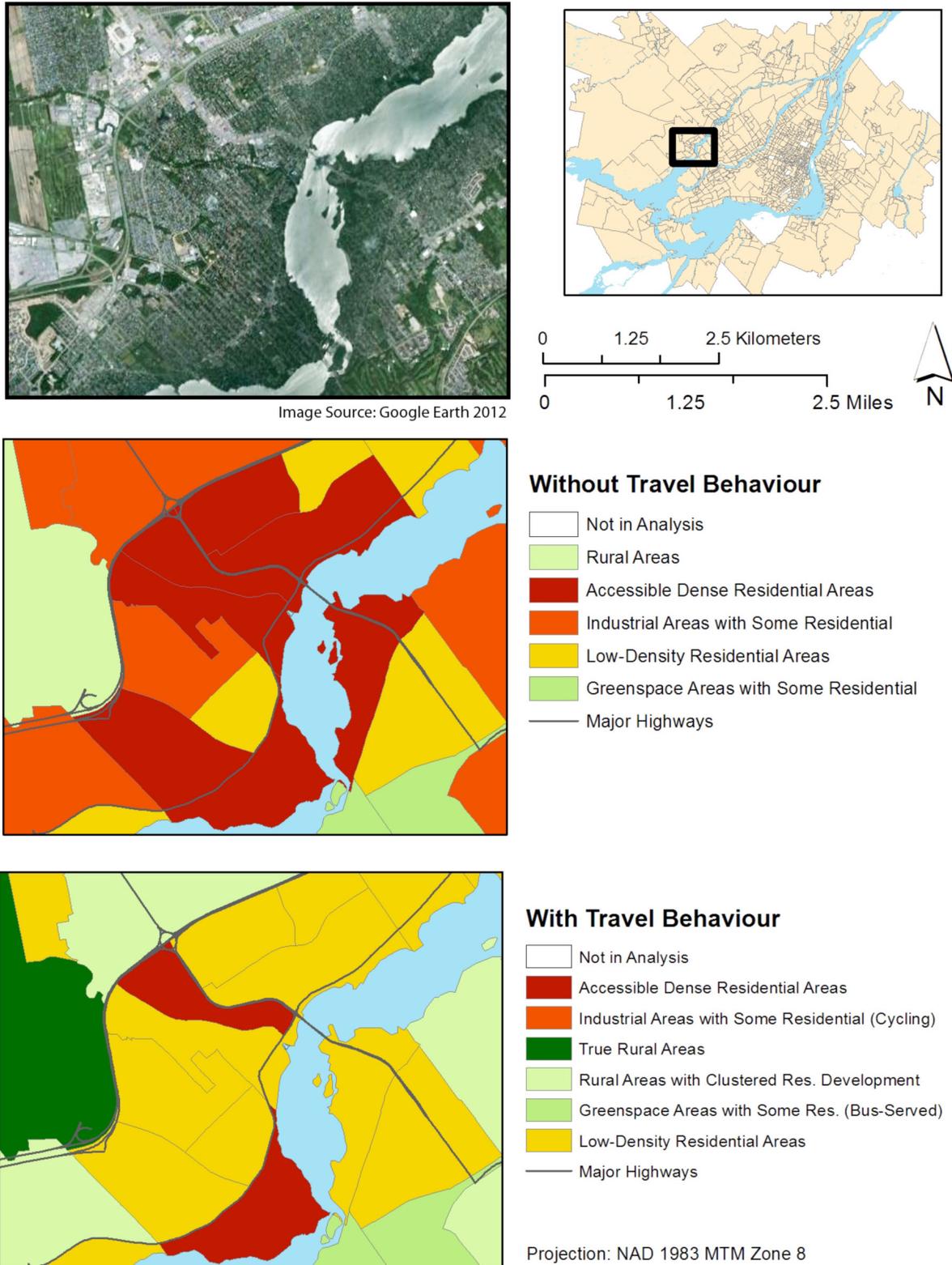
Figure 6 provides a detailed snapshot of a small area along the South Shore (around Longueuil), comparing the results of each model to a satellite image of the area. In both models, there is a distinct area typified as Accessible Dense Residential Areas along the western portion of the snapshot. This area, encompassing the Université de Sherbrooke campus and its surrounds, has a more grid-like street network, a combination of plexes, small apartment buildings and single-family homes on smaller lots, and a shopping mall complex, as well as good access to transit (e.g., buses and a subway station). An area with similar housing types is captured just southwest of the highway running east-west in the snapshot. The extent of the Accessible Dense Residential Areas in the model with travel behavior is more constrained than in the model without travel behavior, which is more representative of the actual area.

Another notable difference between the results for the two models in this snapshot is the extent of the Industrial Areas with Some Residential. In the model without travel behavior, the extent of this typology is more expansive, while in the model with travel behavior it is limited to just north of the highway running east-west. In examining the satellite image, it was found that industrial uses are concentrated more or less along the northern edge of the highway. In addition, the large census tract in the northeast of the snapshot, directly adjacent to the east-west highway, contains the Saint-Hubert Airport. While the airport comprises only a portion of the census tract along the highway, this land use appears to dominate the classification of the census tract in the model without travel behavior. In the model with travel behavior, however, the effect of the airport is subdued somewhat to instead reflect the distinct agricultural uses just north of the airport (and some clustered houses). The same is true for the area just south of the highway intersection along the eastern portion of the snapshot, where the majority of the industrial uses relate to farming and there are residential divisions bordering fields and forested areas.

From the comparison of the two models with satellite images of the region, it is clear that the inclusion of travel behavior considerations improves the characterization of urban form at the census tract level. The main effect that the behavioral variables seems to be having on area classification is moderating the impact of specialized land uses (i.e., the presence of an airport or industry), allowing for a more accurate depiction of the census tract overall. This impact is particularly evident in areas represented by larger census tracts, which are generally the areas criticized as being the least accurately represented by urban form measures derived at the census tract or other aggregate levels.

7   Conclusions and implications

While much of the data on urban form generally used by planners are readily available at the census tract (or some other aggregate) level, there has been much criticism over the accuracy of analyses conducted at this unit of analysis. Moreover, there has been a general lack of consideration of non-built-environment factors in existing quantitative measures of urban form, despite much research that affirms a relationship between built form and (travel) behavior. This methodological study explores the impact of the inclusion of travel behavior in a quantitative characterization of urban form at the census tract level to test whether the inclusion of these variables can improve model outputs. The results provide empirical evidence that the inclusion of travel behavior considerations, in addition to built form variables, provides a more accurate representation of the actual

Image Source: Google Earth 2012

Without Travel Behaviour

- Not in Analysis
- Rural Areas
- Accessible Dense Residential Areas
- Industrial Areas with Some Residential
- Low-Density Residential Areas
- Greenspace Areas with Some Residential
- Major Highways

With Travel Behaviour

- Not in Analysis
- Accessible Dense Residential Areas
- Industrial Areas with Some Residential (Cycling)
- True Rural Areas
- Rural Areas with Clustered Res. Development
- Greenspace Areas with Some Res. (Bus-Served)
- Low-Density Residential Areas
- Major Highways

Projection: NAD 1983 MTM Zone 8

Figure 5: Detailed view of the results from both analyses (with and without travel behavior) and a satellite image for the northwestern portion of Laval and the southwestern portion of the North Shore around Saint-Eustache.

Image Source: Google Earth 2012
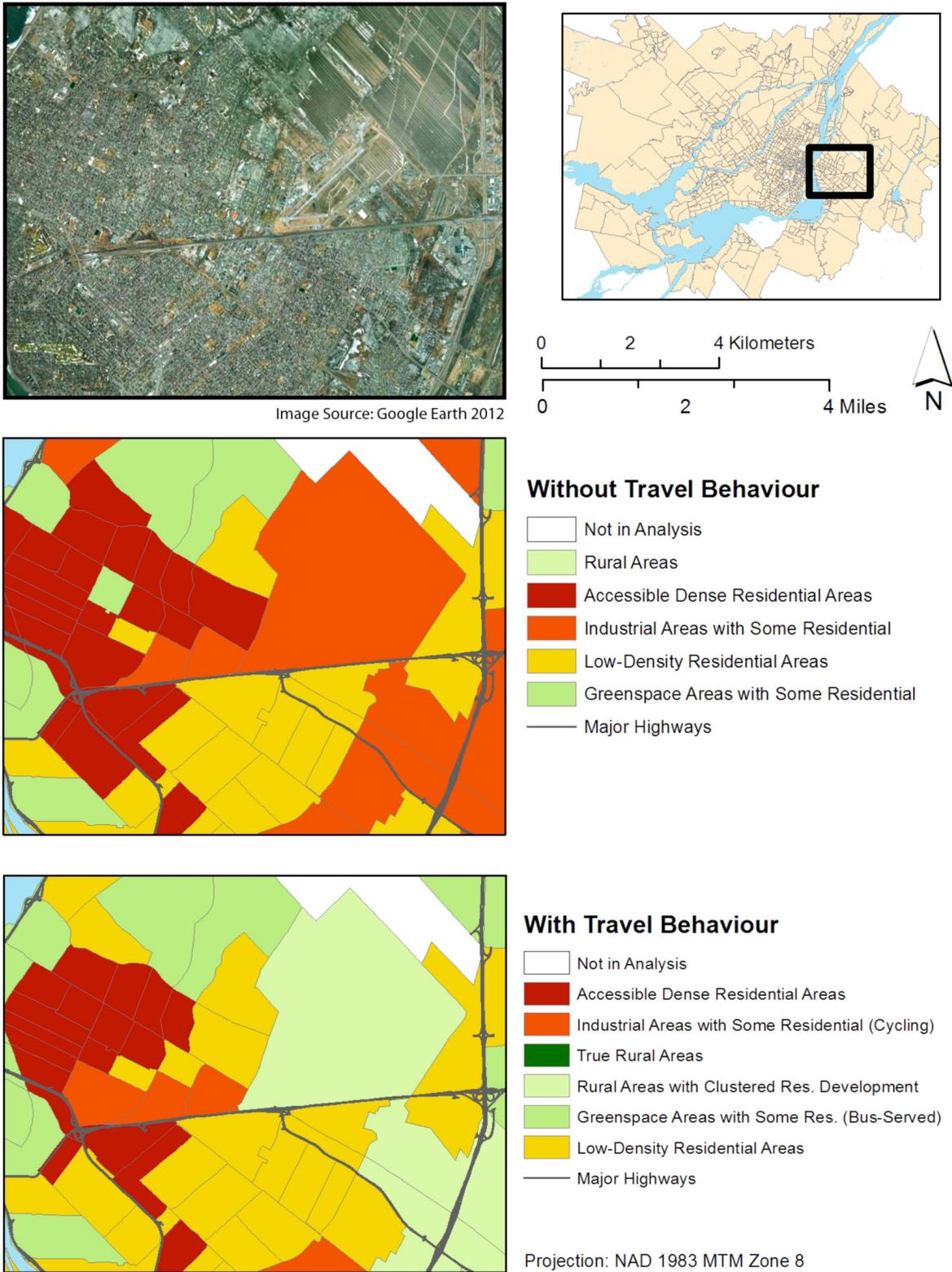
Figure 6: Detailed view of the results from both analyses (with and without travel behavior) and a satellite image for a portion of the South Shore around Longueuil.

urban form on the ground at the census tract level. The inclusion of travel behavior appears to refine the output of the model for characterizing urban form typologies, primarily in areas represented by large census tracts, by moderating the effects of features that normally over-influence the characterization of these census tracts, resulting in misleading model outputs. Considering both built environment and behavioral characteristics together in an analysis of urban form may help to yield more accurate results, even at the aggregate census tract level.

It is important to note that the produced urban form classification cannot be used in travel behavior research, since doing so will impose an endogeneity problem. Yet the new urban form classification can be used in other studies such as land valuing, land taxation, prioritizing investment, and policy interventions. The study also adds a new form of classification going beyond the traditional "urban and rural" classification.

The findings of this study have significant research implications, particularly in cases where data availability limits analysis to aggregate or pre-defined boundaries. A limitation of this study, however, is that the degree to which the inclusion of behavioral considerations improves model outputs at the census tract level cannot be fully assessed without comparing these results to those generated by applying this methodology at a more disaggregate unit of analysis. Thus, it would be valuable to conduct an assessment comparing the model outputs at both the census tract level and, for example, the property level (as in Manaugh et al. 2010 and Song and Knaap 2007) to gain a better understanding of how much these variables actually improve aggregate-level analyses. It would also be interesting to conduct this study for other regions, comparing the results of analyses with and without travel behavior at the census tract level, to see if similar results are obtained.

8   Acknowledgements

References

AMT (Metropolitan Transportation Agency) 2008. Fichier de déplacements des personnes dans la région de Montréal: Enquête origine-destination 2008. In A. m. d. transport (Ed.). Montreal, QC.

Bagley, M., and P. Mokhtarian. 2002. The impact of residential neighborhood type on travel behavior: A structural equations modeling approach. *The Annals of Regional Science* 36: 279–297. doi: 10.1007/s001680200083.

Bagley, M., P. Mokhtarian, and R. Kitamura. 2002. A methodology for the disaggregate, multidimensional measurement of residential neighbourhood type. *Urban Studies* 39(4): 689-704. doi:10.1080/00420980220119525.

Boarnet, M., and R. Crane. 2001. The influence of land use on travel behavior: Specification and estimation strategies. *Transportation Research Part A* 35: 823-845. doi: 10.1016/S0965-8564(00)00019-7.

Cao, X., S. Handy, and P. Mokhtarian. 2006. The influences of the built environment and residential self-selection on pedestrian behavior: Evidence from Austin, TX. *Transportation* 33: 1–20. doi: 10.1007/s11116-005-7027-2.

Cervero, R., and K. Kockelman. 1997. Travel demand and the 3D's: Density, diversity, and design. *Transportation Research Part D* 2(3): 199–219. doi: 10.1016/S1361-9209(97)00009-6.

Cervero, R., and C. Radisch. 1996. Travel choices in pedestrian versus automobile oriented neighborhoods. *Transport Policy* 3(3): 127–141. doi: 10.1016/0967-070X(96)00016-9.

Crane, R., and R. Crepeau. (1998). Does neighbourhood design influence travel? A behavioral analysis of travel diary and GIS data. *Transportation Research Part D* 3(4): 225–238. doi: 10.1016/S1361-9209(98)00001-7.

Cranshaw, J., R. Schwartz, J. Hong, and N. Sadeh. 2012. The Livehoods project: Utilizing social media to understand the dynamics of a city. Paper presented at the 6th International AAAI Conference on Weblogs and Social Media, Dublin, Ireland.

El-Geneidy, A., A. Cerdá, R. Fischler, and N. Luka. 2011. The use of accessibility measures to evaluate the impacts of transportation plans: An application in Montréal, Québec. *Canadian Journal of Urban Research: Canadian Planning and Policy* (supplement) 20(1): 81–104.

Ewing, R., and R. Cervero. 2001. Travel and the built environment: A synthesis. *Transportation Research Record* 1780: 87–114. doi: 10.3141/1780-10.

Frank, L., and G. Pivo. 1994. Impacts of mixed use and density on utilization of three modes of travel: Single-occupant vehicle, transit, and walking. *Transportation Research Record*

1466: 44–52.

Handy, S. 1996. Urban form and pedestrian choices: Study of Austin neighborhoods. *Transportation Research Record* 1552: 135–144. doi: 10.3141/1552-19.

Handy, S., X. Cao, and P. Mokhtarian. 2007. Self-selection in the relationship between the built environment and walking: Empirical evidence from Northern California. *Journal of the American Planning Association* 72(1): 55–74. doi: 10.1080/01944360608976724.

Kitamura, R., P. Mokhtarian, and L. Laidet. 1997. A microanalysis of land use and travel in five neighborhoods in the San Francisco Bay Area. *Transportation* 24: 125–158. doi: 10.1023/A:1017959825565.

Lynch, K. 1960. *The Image of the City.* Cambridge, Mass.: MIT Press.

Manaugh, K., L. Miranda-Moreno, and A. El-Geneidy. 2010. The effect of neighbourhood characteristics, accessibility, home–work location, and demographics on commuting distances. *Transportation* 37: 627–646. doi: 10.1007/s11116-010-9275-z.

Moudon, A., P. Hess, M. Snyder, and K. Stanilov. 1997. Effects of site design on pedestrian travel in mixed-use, medium-density environments. *Transportation Research Record* 1578: 48–55.

Næss, P. 2011. "New urbanism" or metropolitan-level centralization? A comparison of the influences of metropolitan-level and neighborhood-level urban form characteristics on travel behavior. *Journal of Transport and Land Use* 4(1): 25–44. doi: 10.5198/jtlu.v4i1.170.

Niemeier, D., S. Bai, and S. Handy. 2011. The impact of residential growth patterns on vehicle travel and pollutant emissions. *Journal of Transport and Land Use* 4(3): 65–80. doi: 10.5198/jtlu.v4i3.226.

Rodriguez, D., H. Young, and R. Schneider. 2006. An easy to compute index for identifying built environments that support walking. Paper presented at the Transportation Research Board Annual Meeting, Washington, DC.

Rogerson, P. 2010. *Statistical Methods for Geography: A Student's Guide (3 ed.).* Thousand Oaks, CA: Sage Publications Ltd.

Schwanen, T., and P. Mokhtarian. 2005. What affects commute mode choice: Neighborhood physical structure or preferences toward neighborhoods? *Journal of Transport Geography* 13: 83–99. doi: 10.1016/j.jtrangeo.2004.11.001.

Song, Y., and G. Knaap. 2004. *Internally Connected, No Commercial, With a Touch of Open Space: The Neighborhoods of New Homes in the Portland Metropolitan Area.* College Park, MD: National Center for Smart Growth Research & Education, University of Maryland.

Song, Y., and G. Knaap. 2007. Quantitative classification of neighbourhoods: The neighbourhoods of new single-family homes in the Portland Metropolitan Area. *Journal of Urban Design* 12(1): 1–24. doi: 10.1080/13574800601072640.

Southworth, M., and P. Owens. 1993. The evolving metropolis. *Journal of the American Planning Association* 59(3): 271–288. doi: 10.1080/01944369308975880.

Srinivasan, S. 2002. Quantifying spatial characteristics of cities. *Urban Studies* 39(11): 2005–2028. doi: 10.1080/0042098022000011335.

Talen, E. 2002. Help for urban planning: The transect strategy. *Journal of Urban Design* 7(3): 293–312. doi: 10.1080/1357480022000039349.

Talen, E. 2003. Measuring urbanism: Issues in smart growth research. *Journal of Urban Design* 8(3): 195–215. doi: 10.1080/1357480032000155141.

Wheeler, S. 2003. The evolution of urban form in Portland and Toronto: Implications for sustainability planning. *Local Environment* 8(3): 317–336. doi: 10.1080/13549830306656.