The Canadian Geographer
Le Géographe canadien
Canadian Association of Geographers
Association Canadienne des Géographes

Comparison of apportionment methods for assigning trip data to rezoned traffic analysis zones: A case study of Toronto, Canada

Hong Yao
Department of Geography and Planning, Queen's University

DongMei Chen
Department of Geography and Planning, Queen's University

Key Messages

- Six apportionment methods for assigning trip data to rezoned traffic analysis zones were examined using the trip data of Toronto from 2001 and 2016.
- The apportionment results from the six methods were compared using hot-spot analysis and the Wilcoxon signed-rank test.
- The results of comparisons provide references for the choices of methods in apportioning trip data.

Due to regional development impacts, existing traffic analysis zones may be rezoned. Apportioning data to new traffic analysis zones is essential to ensure data analysis consistency and comparability. The traditional method uses area to apportion data. This study introduces six methods in apportioning trip data, including the traditional method (M1); the residential land use counted method (M2); the population counted at the dissemination area (DA) level method (M3); the integrated method using both residential land use and the population in DAs (M4); the population counted at the dissemination block (DB) level method (M5); and the integrated method using both residential land use and the population in DBs (M6). These methods are demonstrated in the case of Toronto, Canada using trip data from 2001 and 2016. Results from the six methods are compared and analyzed using mapping, Getis-Ord Gi hot-spot analysis, and the Wilcoxon signed-rank test. Our findings show that the traditional method and the population counted in DAs method are significantly different ($p < 0.05$) from the residential land use counted method and the integrated method using both residential land use and the population in DAs and the integrated method using both residential land use and the population in DBs. These results provide references for selecting appropriate apportionment methods, which is the basis for transportation planning and policymaking.*

Keywords: méthodes de répartition, comparaison, zones d'analyse du trafic, données liées aux déplacements, Toronto

Comparaison des méthodes de répartition des données liées aux déplacements au sein des zones d'analyse du trafic : une étude de cas de la région de Toronto, Canada

Considérant le développement urbain à l'échelle régionale, les zones d'analyse du trafic devraient être reconfigurées dans la région de Toronto. Pour ce faire, il est essentiel d'associer les données aux nouvelles zones d'analyse du trafic afin d'assurer la cohérence et la comparabilité des données. La méthode traditionnelle utilise la zone pour répartir ou attribuer les données. La présente étude compare six méthodes d'attribution des

Correspondence to / Adresse de correspondance: DongMei Chen, Department of Geography and Planning, Queen's University, Mackintosh-Corry Hall, 68 University Avenue, Kingston, ON K7L 3N6. Email/Courriel: chendm@queensu.ca

DOI: 10.1111/cag.12675

données liées aux déplacements, c'est-à-dire la méthode traditionnelle (M1), la méthode de dénombrement de l'usage des terrains résidentiels (M2), la méthode fondée sur le niveau du dénombrement de la population dans l'aire de diffusion (AD) (M3), la méthode intégrée utilisant l'usage des terrains résidentiels et la population dans les AD (M4), la méthode fondée sur le niveau du dénombrement de la population dans l'îlot de diffusion (ID) (M5) et, finalement, la méthode intégrée utilisant l'usage des terrains résidentiels et la population dans les ID (M6). Ces diverses méthodes sont expérimentées dans la région de Toronto, au Canada, en utilisant les données liées aux déplacements de 2001 à 2016. Les résultats des traitements des six méthodes sont comparés et analysés en utilisant la cartographie, l'analyse des points chauds Getis-Ord Gi et le test Wilcoxon signed-rank. Nos conclusions montrent que la méthode traditionnelle et le dénombrement de la population dans la méthode des AD sont sensiblement différents des autres méthodes. Ces résultats fournissent des informations sur le choix des méthodes appropriées de répartition des données sur le trafic, ce qui constitue un élément important pour l'élaboration de politiques et la planification en matière de transport.*

Mots clés : apportionment methods, comparison, rezoned traffic analysis zones, trip data, Toronto

Introduction

Traffic analysis zones (TAZs) are defined as geographic areas dividing the planning region into spatially homogeneous areas with respect to socio-economic characteristics and land use (Cambridge Systematics and AECOM Consult 2007; Martinez et al. 2009). The TAZ is a commonly used geography unit in travel demand modelling for transportation planning with aggregated individual trips based on origin and destination zones. This type of areal (zonal) data is useful for providing an overview of trip characteristics and for suggesting where certain transportation policies might best be deployed (Martinez et al. 2009; Mustafa and Zhong 2014; Altan and Ayozen 2018).

In Canada, TAZs are usually constructed based on census block boundaries with associated demographic-social-economic data. The delineation of TAZ boundaries requires consideration of several criteria. These include homogeneity (single predominant land use and homogeneous socio-economic characteristics); spatial contiguity (units composing a zone should be adjacent to each other); spatial compactness; compatibility with existing linear features (e.g., rivers, bridges, railroads, and roads); and compatibility with census geography (You et al. 1998; Miller and Shaw 2001; Cambridge Systematics and AECOM Consult 2007).

However, regional developments, future planned transportation corridors, new political boundaries, and increased population and trips significantly impact the determination of how to delineate TAZ boundaries (Cambridge Systematics and AECOM Consult 2007). The existing TAZs may be rezoned to reflect these changes. To ensure data analysis consistency and comparability across different periods, apportioning the respective data to new TAZs is essential. It is especially imperative for transportation planning and comparisons related to time series analysis.

Scaling and aggregating spatial data through rezoning are common spatial issues under the modifiable area unit problem (MAUP) (Openshaw and Taylor 1982; Wong and Amrhein 1996). The modifiable area units and boundary issues should be given specific attention during the specification of a TAZ (Viegas et al. 2009). Several previous research studies have highlighted the importance of aggregation approaches to TAZs in transportation-planning analysis (Ding 1994, 1998; Alvanides et al. 2000; Chang et al. 2002; Viegas et al. 2009; Altan and Ayozen 2018). Through a GIS-based human-interactive TAZ design algorithm that generates TAZ alternatives, Ding (1998) demonstrated the significant impact of spatial data aggregation on the outcomes of transportation-planning models. In another study, Chang et al. (2002) illustrated the effects of TAZs, centroids, and network details on statewide traffic demand modelling in Idaho. Viegas et al. (2009) analyzed and measured the MAUP effects on the TAZ delineation and transportation demand models and revealed the information loss in the traffic assignment step of transportation planning models. While these studies have emphasized the impact of the MAUP on TAZ delineation and what approaches can better establish TAZs, little attention has been given to apportioning the respective data once rezoning occurs.

A traditional method in apportioning data is based on area alone (Beale 2012; Harder et al. 2013). For instance, a census block is split by a buffer in an overlay operation, resulting in two output features. The attribute values are divided according to the area and can be implemented with a tool named Make Feature Layer in ArcGIS. In addition to the traditional method, some new apportionment methods have been developed. Beale (2012) conducted a study describing how to apportion the population between areas. In this study, two methods of mask area weighting and filtered area weighting were put forward and were deemed improvements on the traditional method. The mask area weighting method used ancillary data as a mask to eliminate the areas deemed to be uninhabitable, e.g., water areas. The filtered area weighting method apportioned data to a finer unit of analysis—e.g., postcode data, housing centroids, etc.—using the ancillary dataset. The filtered area weighting method was echoed by ArcGIS REST API (Esri 2020b) in its dasymetric apportionment technique. A study by Cambridge Systematics and AECOM Consult (2007) utilized census data at the block group level to redistribute the original zone's household data to new zones. They pointed out that local knowledge of land use can assist in deciding how the socio-economic data needs to be distributed by new zone structures. Gorr and Kurland (2010) presented a method to apportion census data to administrative areas using block centroid population, assuming the populations are evenly distributed across the census tract. In another study exploring how to enhance positional accuracy to reduce bias in health-related studies, Rosu and Chen (2016) apportioned the dissemination block population into the postal code boundary using land use and spatial distribution of populations within postal code boundaries.

Although increased attention is being given to defining TAZs, the trip data apportionment remains very important for transportation planning. Incorrect information on trip numbers generated in the trip generation model will lead to sizable errors in transportation planning modelling. Trip generation is the first sub-model of the conventional four-step model sequence. The trip distribution model, which is concerned with the question "where do the trips go" predicts how many of the trips originating in zone i will terminate in each of zones 1, 2, … j (Hanson and Giuliano 2004). The input data for the trip distribution model is trip numbers. Similarly, the subsequent mode split and trip assignment depend on the trip data. Additionally, more and more studies integrate transportation planning into urban smart growth strategies (Behan et al. 2008). Trip data are the primary information in establishing smart growth strategies. Further, it is crucial that accurate information is being obtained when examining the change of trip numbers over time, as incorrect information will lead to inappropriate urban planning and policies.

The aim of this study is to examine different methods of apportioning trip data used in travel demand studies and compare their differences in apportionment results from different methods. In this study, six scenarios are designed to demonstrate the development of apportionment methods upon data availability. The six methods are as follows: (1) Method 1 (M1) is the traditional method; (2) Method 2 (M2) is the residential land use counted method; (3) Method 3 (M3) is the population counted at the dissemination area (DA) level method; (4) Method 4 (M4) is the integrated method using both residential land use and the population in the DAs; (5) Method 5 (M5) is the population counted at the dissemination block (DB) level method; and (6) Method 6 (M6) is the integrated method using both residential land use and the population in DBs. The City of Toronto was chosen as a case study to demonstrate the development of apportionment methods. Six models for the six methods using ArcGIS ModelBuilder were established for running the geoprocessing operations of apportionment. The apportionment results between the six methods were compared and analyzed using mapping, Getis-Ord G* hot-spot analysis, and the Wilcoxon signed-rank test. The uses of the Wilcoxon signed-rank test were to determine whether there are significant differences in the apportionment results between the six methods.

Study area and data

The City of Toronto was chosen due to its high number of TAZ boundary changes in the past years. In the past 25 years, Toronto has undergone changes in TAZ boundaries, resulting in increases in the number of TAZs—from 460 TAZs in 1991 to 463 TAZs in 1996, to 481 TAZs in 2001, and to 625 TAZs in 2006 through 2016 (DMG 2006). In this study, the 2001 and 2016 TAZ boundaries and the 2001 personal trip numbers were taken as an example to illustrate how trip data are apportioned

Figure 1
Study area with 2016 Traffic Analysis Zones (TAZs).

from the earlier TAZ boundaries (2001) to the rezoned TAZ boundaries (2016). Figure 1 depicts the study area for our analysis. The City of Toronto is located at the northwest shore of Lake Ontario, covering an area of 630 square kilometres. The total population was 2.48 million in 2001 and increased to 2.73 million in 2016 (City of Toronto 2017). It is the largest city in Canada and the capital of the province of Ontario.

Figure 2a shows the distribution map of trip numbers of TAZs in 2001. The enormous discrepancies in the distribution of trip numbers mainly occur in peripheral areas. These include West Humber-Clairville, Islington-City Center West, Mimico in western areas (Etobicoke), Humber Summit, Milliken, and York University Heights in the northern areas (North Toronto). In contrast, tiny discrepancies occur in the east of Toronto along the periphery, downtown areas, and neighbourhoods in inner areas, including Woodbine-Lumsden, Flemingdon Park, Runnymede-Bloor West Village, Annex, High Park North, and Humber Heights-Westmount. The other neighbourhoods are surrounded by a mix of high and low trip numbers. The spatial distribution of people and land use within an urban area has an important impact on travel because travel is a derived demand that arises when people are spatially separated from the location at which they wish to be (Hanson and Giuliano 2004). Figures 2b and 2c display the maps of the distribution of the population and land use, respectively. Comparing with the map of trip numbers, the distributions of residential land use and population are similar to the distribution of trip numbers, showing a high spatial coincidence.

The 2001 and 2016 TAZ boundaries and the 2001 trip numbers were drawn from the Transportation Tomorrow Survey (TTS) of the Data Management Group (DMG 2001, 2006). TTS is a comprehensive travel survey covering the Greater Toronto region every five years, starting in 1986. It is conducted and maintained by the Data Management Group of the Joint Program in Transportation and Civil Engineering Department of the University of Toronto. The TTS database consists of telephone interviews with Toronto residents, using a consistent survey instrument, and set of procedures and questions. It provides statistically reliable travel behaviour data for each survey year (Miller and Soberman 2003). The survey has been undertaken

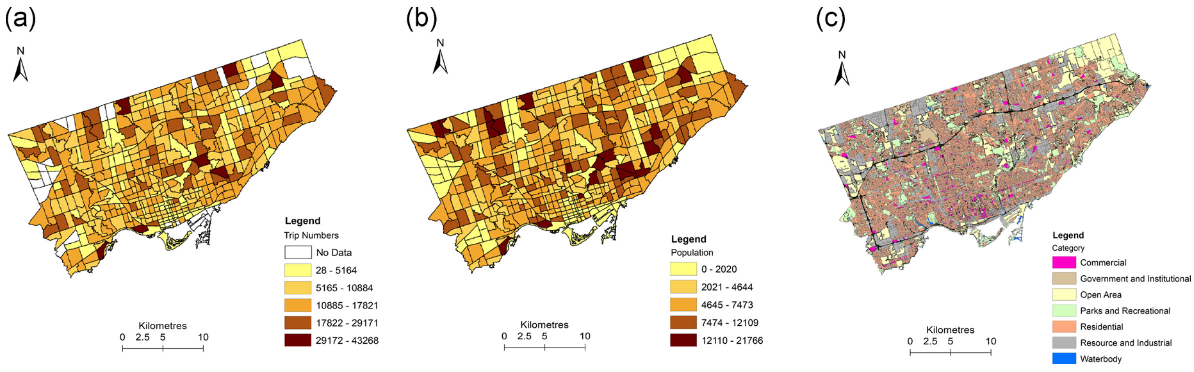

Figure 2
The distribution of trip numbers, population, and land use in 2001 in Toronto: (a) trip numbers; (b) population; and (c) land use.

for the years 1986, 1991, 1996, 2001, 2006, 2011, and 2016 with sample sizes being a 4.2%, 1.4%, 5.0%, 5.5%, 5.2%, 5.1%, and 5.0% random sample of households in the survey area, respectively (DMG 2016). The TTS database provides four sets of attributes: household, person, trip, and transit trip. These data are reported at the TAZ level.

The 2001 Digital Boundary Files at the dissemination area (DA) level and the dissemination block (DB) level were downloaded from the Statistics Canada website (Statistics Canada 2001a, 2001b). The DA is a small, relatively stable geographic unit composed of one or more blocks with a population of 400 to 700 persons. The City of Toronto was divided into 3954 DAs in 2001 (Statistics Canada 2001d). A DB is an area bounded on all sides by roads or boundaries of standard geographic areas. The DB is the geographic area for which population and dwelling counts are disseminated (Statistics Canada 2001c). In 2001, the term “block” was used; the fuller term “dissemination block” came into use in 2006 (Statistics Canada 2006). The study area was divided into 12,593 block boundaries in 2001 (Statistics Canada 2001c). The projection for the two boundaries was standardized to NAD UTM ZONE 17.

The DA level populations of 2001 were retrieved and downloaded from the Canadian Census Analyser at CHASS, University of Toronto (CHASS Data Centre 2001). The DB level populations of 2001 were obtained using the PCensus 10 for MapPoint software from Queen's University. The 2001 land use boundaries used in the study were obtained from DMTI Spatial Inc. downloaded from the website of the Scholars GeoPortal of the Ontario Council of University Libraries (DMTI Spatial Inc. 2001a). The categories of land use are divided into commercial, governmental, institutional, open area, parks and recreational, residential, resources and industrial, and waterbody (DMTI Spatial Inc. 2001b).

Methods

Based on the review of the traditional method and the newly developed methods delineated in the Introduction section, six methods were introduced for apportioning the trip data to new TAZs based on the currently available data in this study. The development of six apportionment methods was designed to correspond with six scenarios of different data availability, as illustrated in Table 1.

M1 for scenario 1: The traditional method

This method assigns the trip numbers within the source TAZs (2001) to the target TAZs (2016) according to the geometry area. The percentage of each source TAZ area that falls inside the target TAZ is calculated, which is called the weight. The weight is used to allocate the trip numbers within the source TAZs to the target TAZs.

The first step involves overlaying the source TAZs with the target TAZs. Next, the weight for each overlaid boundary is calculated by dividing the area of overlaid boundary by the area of source TAZ. Once the weight for each overlaid boundary is calculated, the trip numbers within each overlaid boundary

Table 1
Six scenarios with corresponding data used.

Scenario	2001 TAZs	2016 TAZs	2001 trip numbers	2001 residential land use	2001 DA boundaries	Population at DA level	2001 DB boundaries	Population at DB level
Scenario 1	x	x	x					
Scenario 2	x	x	x	x				
Scenario 3	x	x	x		x	x		
Scenario 4	x	x	x	x	x	x		
Scenario 5	x	x	x				x	x
Scenario 6	x	x	x	x			x	x

can be obtained by multiplying the trip numbers within the source TAZ by the weights. Finally, the trip numbers are summarized by the target TAZs. Figure 3a displays a visual demonstration of the process with the equations of apportionments.

M2 for scenario 2: The residential land use counted method

This method makes use of residential land use boundaries to calculate apportionment weight. First, the residential land use boundaries are

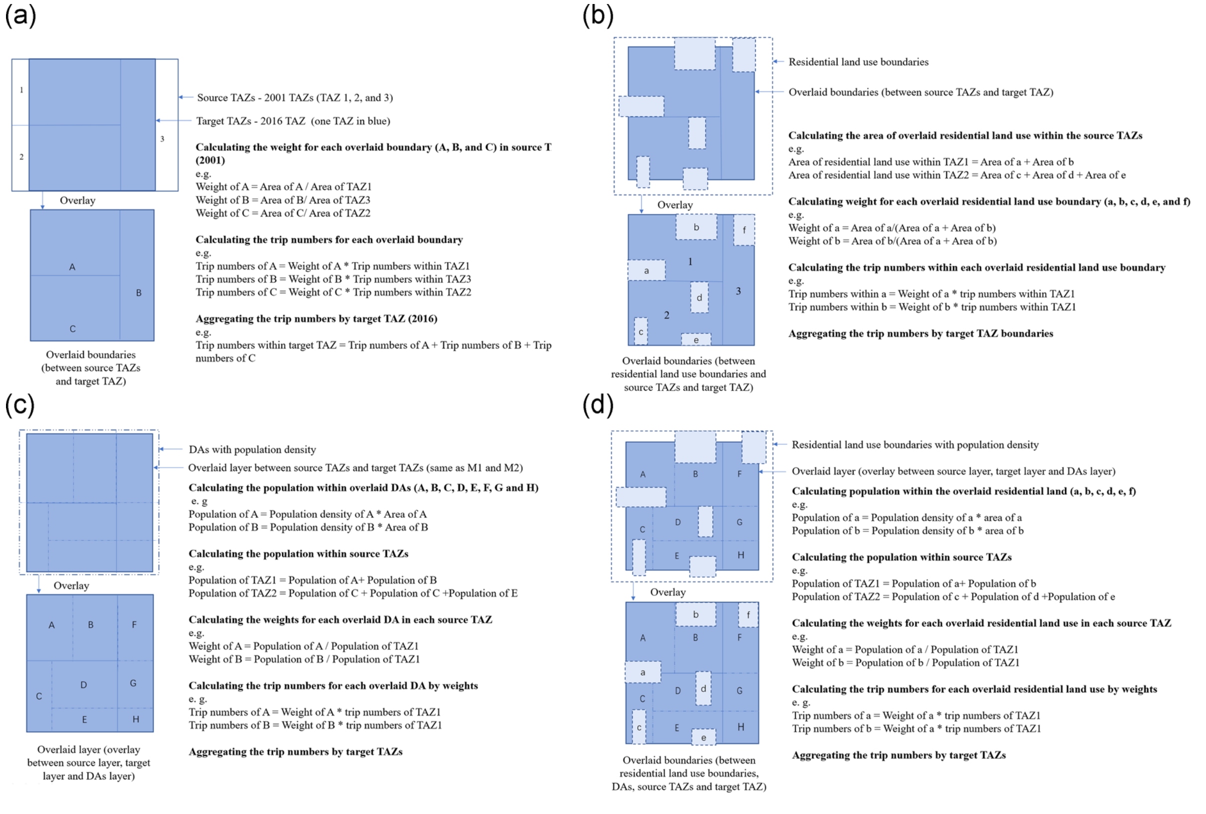

Figure 3
Methods for apportioning the number of trips: (a) traditional method; (b) residential land use counted method; (c) population counted at DA level method; and (d) integrated method using both residential land use and population in the DAs.

overlaid by the target TAZs and the source TAZs. Then, the ratio of areas of overlaid residential land use boundaries within the source TAZs is calculated. The ratio is the weight used to apportion data in the following step. The trip numbers within each overlaid residential land use boundary are calculated by multiplying the weight by the trip numbers within each source TAZ. Finally, the trip numbers are summarized by the target TAZs. Figure 3b displays a visual demonstration of the process for M2.

M3 for scenario 3: The population counted at the DA level method

Being different from M2, M3 considers population distribution using population data at the level of DAs within source TAZ boundaries. The first step involves the calculation of population density within the DAs. Once the population density is calculated for all the DAs, the DAs interact with the layers overlaying the source TAZs with the target TAZs. Next, the population is calculated for each overlaid DA by multiplying their areas by the population density of DA. Based on the calculated population within the overlaid DAs, the total population of the source TAZs can be obtained by summing the population of overlaid DAs that falls inside the source TAZs. The next step is to calculate the fraction of the population of overlaid DAs in each source TAZ. This fraction is called the weight. Following this, using the weight calculates the trip numbers for each overlaid DA. Finally, the trip numbers by target TAZ boundaries are aggregated. The visual demonstration of the apportionment process is shown in Figure 3c.

M4 for scenario 4: The integrated method using both residential land use and the population in the DAs

This method considers the residential land use and spatial distribution of populations at the level of the DAs. Firstly, the residential land use layer intersects with the overlaid layer, an overlaid layer between the source layer, target layer, and DAs. Secondly, the population density within each overlaid residential land use boundary is calculated by dividing the population within each DA by the area of the overlaid residential land use. Next, to calculate the total population within each source TAZ, the population within each overlaid residential land use is calculated by multiplying its area by the corresponding population density. After obtaining the population within each overlaid residential land use, the population within each source TAZ can be calculated by summing the population of each overlaid residential land use in each source TAZ. The weight of each overlaid residential land use is calculated by dividing the population of overlaid residential land use by the population within each source TAZ. The trip numbers within each overlaid residential land use can be calculated by multiplying the weight by the trip numbers within each source TAZ. Finally, the trip numbers are summed by the target TAZs. Figure 4d shows the apportionment processes for M4.

The remaining two methods are M5 and M6, corresponding to scenarios 5 and 6, respectively. M5 has the same apportionment method and procedures as M3, except for replacing the population data at the DA level with the population data at the DB level. Similarly, M6 is similar to M4 except for replacing the census data at the DA level with the census data at the DB level. Further, six models for the six methods have been developed using ArcGIS ModelBuilder to encapsulate the entire steps described in the six methods. A flow chart for M4 shown in Figure 4 is presented as an example for illustration purposes.

The apportionment results from the six methods were mapped and the differences in the apportionment results between the methods were summarized in tabulation using the number of the zones included in each range of values.

The Getis-Ord Gi* hot-spot analysis in ArcGIS was conducted to analyze the apportioned trip patterns among TAZs. The Getis-Ord Gi* statistic identifies spatial clusters of high values (hot spots) and spatial clusters of low values (cold spots) (Getis and Ord 2010). The resultant Z-score identified the TAZs having the high or low values of cluster spatially. A positive and large Z-score indicates an intense clustering of high values (hot spot), while a negative and small Z-score signifies an intense clustering of low values (cold spot) (Esri 2020a). The hot-spot analysis tool in ArcGIS calculates the Getis-Ord Gi* statistic for each TAZ. In this study, the delineations of hot spots of high trip numbers and cold spots of low trip numbers in TAZs were compared to see whether the trip patterns varied significantly across the six apportionment methods by using the fixed distance band setting in ArcGIS.

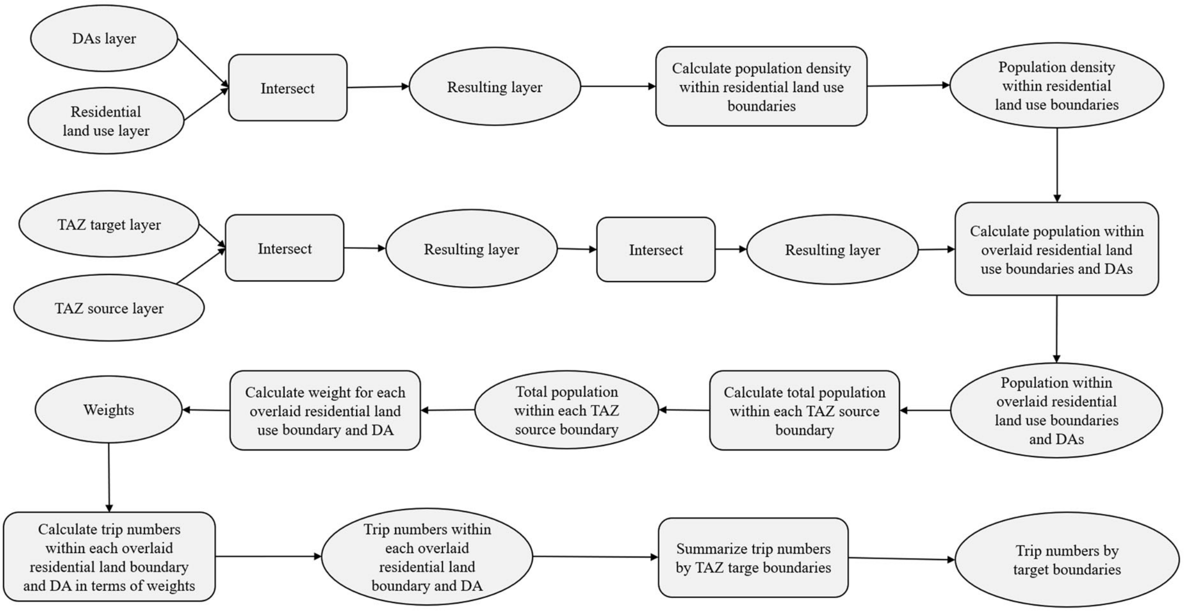

Figure 4
Flow chart for Method 4.

The Wilcoxon signed-rank test was also conducted to determine whether there was a significant difference in the distribution of trip numbers between every two methods. The Wilcoxon signed-rank is similar to the paired t-test and is used to compare two related samples and determine whether the two groups′ mean ranks are different. It is a nonparametric or distribution-free test (Burt and Barber 1996). Based on these characteristics, the Wilcoxon signed-rank test for paired samples is an appropriate test to determine whether the two groups′ distribution was significantly different.

Results and discussion

The trip numbers of people using the different apportionment methods were mapped in Figure 5a. Table 2 summarizes the numbers of zones by a range of trip numbers for the six apportionment methods.

From Table 2, most of the trip numbers apportioned in the six methods are in the range of 0 to 4,000, 4,001 to 9000, and 9,001 to 14,000. In the range of 0 to 4,000, the larger differences in the number of zones are between M1 and M6, M2 and M3, M3 and M4, M3 and M6, and M5 and M6 with 20 to 29 zones. In the range of 4,001 to 9000, the differences occur between M1 and M2 and M1 and M4 with 20 and 30 zones, respectively. The differences between the apportionment results for the other three ranges are not significant.

Figure 5b displays Getis-Ord Gi* values of trip numbers for the six methods. The dark blue areas represent the clustering of low values, and the red areas represent the clustering of high values. The significant differentiation in the high-high hot spots for the six methods mainly occurs in the inner-city area, while low-low cold spots are in the downtown area. The other cold spots difference occurs in Etobicoke, and hot spots in North Toronto. Table 3 presents the number of TAZs for the two areas with significant differentiation in the hot spots and cold spots for the six methods, i.e., the downtown area and inner city.

The number of zones is used to illustrate the variability of hot and cold spots between the six methods. The significant differences in the number of zones in the hot-spot area are between M1and M2 (24 zones), M1 and M4 (27 zones), M1 and M6 (28 zones), M3 and M4 (17 zones), M3 and M6 (18 zones), and M5 and M6 (16 zones), indicating

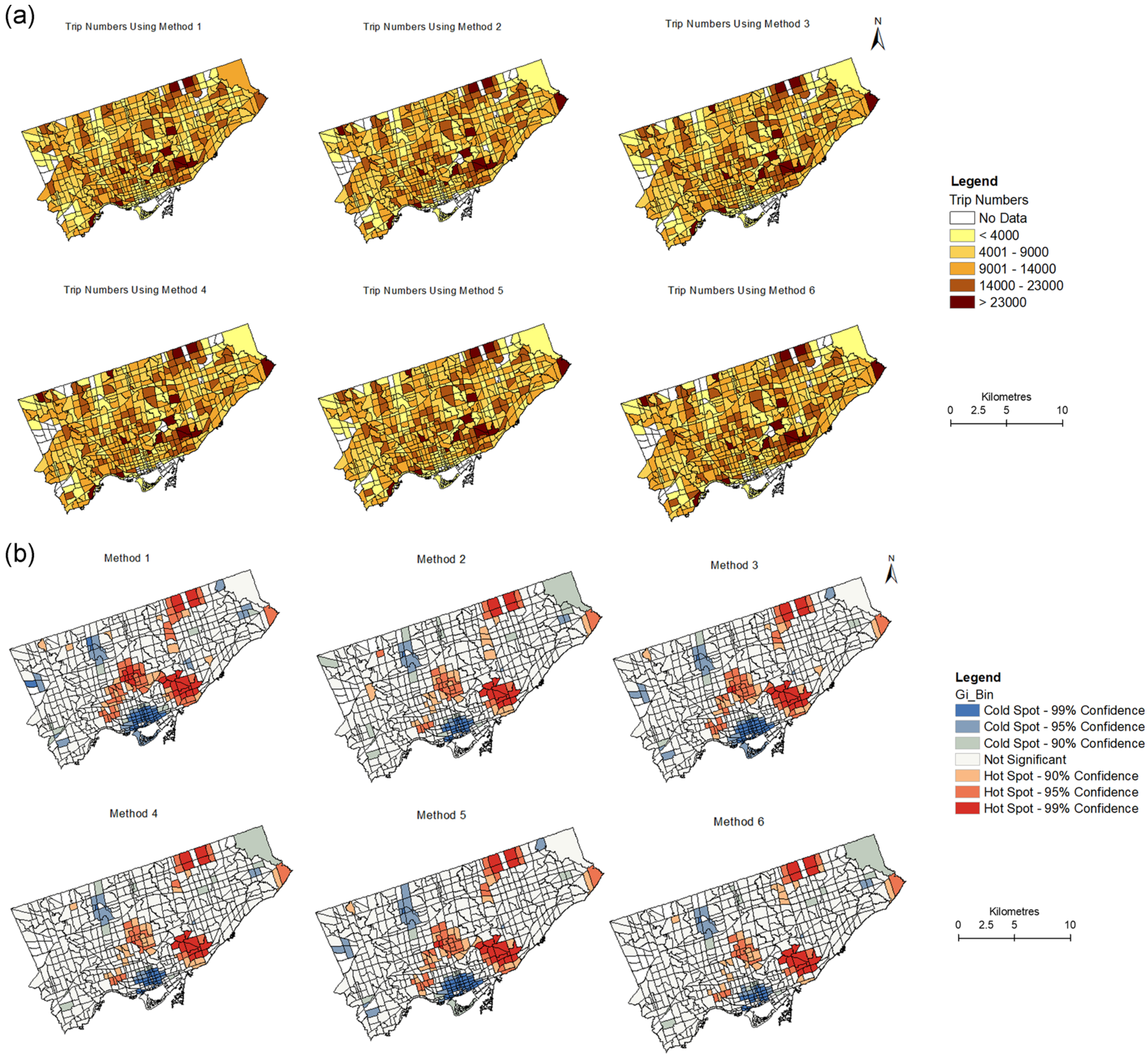

Figure 5
Spatial distribution of trip numbers after using different apportionment methods: (a) trip numbers within target TAZs; (b) Getis-Ord Gi* of trip numbers.

the variability of the distribution of high-high value clustering. Similarly, the differences in the number of zones in the cold spot are between M1 and M2 (22 zones), M1 and M4 (28 zones), and M1 and M6 (27 zones).

The results of the Wilcoxon signed-rank test between every two apportionment methods are as follows: there are significant differences between M1 and M2, M1 and M4, M1 and M6, M2 and M3, M3 and M4, as well as M3 and M6 at $\alpha = 0.05$ (p-value <0.01). The remaining tests show no significant evidence indicates there are significant differences between methods at $\alpha = 0.05$ (p-value > 0.05).

From mapping the apportionment results between the six methods, the Getis-Ord Gi* hot-spot analysis, and the Wilcoxon signed-rank test, it is obvious that similar differences are observed

Table 2
The number of zones by a range of trip numbers for the six apportionment methods.

Range of trip numbers	Zones for M1	Zones for M2	Zones for M3	Zones for M4	Zones for M5	Zones for M6
0–4,000	166	151	171	149	167	142
4,001–9,000	206	176	192	181	195	185
9,001–14,000	142	154	143	150	142	146
14,001–23,000	69	73	75	72	76	75
23,001–45,000	12	14	14	15	14	15

Table 3
The number of zones composed of hot and cold spots at 95% confidence levels.

Area	M1	M2	M3	M4	M5	M6
Downtown	62	40	59	34	57	35
Inner city	55	31	45	28	43	27

across the six methods. These apportionment differences result from the properties of methods.

M1 apportions data just based on geometry area alone, assuming the trip numbers are distributed evenly with each TAZ boundary. This method can compromise accuracy when this assumption is unavailable. For example, the differences can occur in uninhabitable areas (e.g., water areas) and non-residential areas where no population is specifically not located. Also, this method does not use any information about populations or residential land use strongly associated with trip numbers.

In contrast to M1, M2 makes use of the residential land use data to filter the areas where the population does not reside when apportioning the trip number. In contrast to M1, M2 provides a good indication as to where the residential population resides and where the trips possibly are generated. However, M2 uses the size of residential land use boundaries when apportioning trip numbers data, which does not reflect the actual size of the population at residential locations.

M3 apportions the trip numbers using the percentage of the population at the level of the DAs to account for the population of the source TAZs. This percentage is used as the weight to allocate the trip numbers. This method integrates the population density at the DA level into the TAZs in comparison to M1 and M2. However, the limitation of this method is that it does not explicitly consider where the residential population resides so that the population is allocated at broader-scale DAs rather than at the smaller residential land parcels.

M4 takes into consideration population distribution within each residential land use boundary by integrating population density at the DA level into residential land use boundaries. In contrast to M1, M2, and M3, M4 considers where the residential population is located and allocates the population within the residential land use boundary.

M5 is similar to M3, except that it uses the population data at the DB level. Also, M5 better reflects the population distribution than M3 due to the census data on a small scale comparing to M3. Similarly, M6 is similar to M4, except the census data are from the DB level, which is smaller than M4. Therefore, M6 also reflects better the population distribution than M4.

Based on the above analysis, when the detailed land use and population data are available, M6 is preferred because it integrates the residential population locations and population density within residential land-use boundaries at a finer scale (DB).

The different properties of the apportionment methods lead to the differences in the apportionment results. In turn, different apportionment results can influence travel demand modelling and transportation planning policies. Traditional travel demand modelling is the four-step sequential model comprising four sub-models employed in a sequential process. These sub-models are generally referred to as trip generation, trip distribution, modal split, and trip assignment. Trip generation is concerned with predicting the number of trips produced by and attracted to each zone. Trip distribution is concerned with predicting where the trips go. Thus, the trip distribution model links the origin and destination ends of the trips generated by the trip generation model. Modal split addresses the question of how the various trips are made. That is, the modal split model predicts the proportion of trips by each mode of travel between each origin and destination zone. Trip assignment is concerned with predicting the routes used by the trips from a given origin to a given destination by a particular mode (Hanson and Giuliano 2004). As noted here, trips are the basic unit of analysis for the travel demand model. The validity of trips is the basis for modelling travel demand, and the different apportioned trip numbers can directly impact transportation planning and policymaking.

The impact of the MAUP needs to be further addressed when rezoning TAZs and apportioning trip data. There are two MAUP effects: the scale effect relates to the level of spatial data aggregation, and the zoning effect relates to the definition or partitioning of units for which data are collected (Wong and Amrhein 1996). As demonstrated in previous studies, the delineation of zonal boundaries of TAZs has a direct impact on the results of transportation forecasting models (Viegas et al. 2009; Altan and Ayozen 2018). This study used the boundary of existing TAZs and did not address whether the TAZ boundary delineations are appropriate or optimal. To reduce the MAUP effect, it would be worthwhile to perform a sensitivity analysis on the combined effect of TAZ zoning and different apportioning methods on transportation planning analysis in the future.

It should also be noted that this study did not consider the interconnection and effects of the peripheral area of the City of Toronto. In this study, we did not separate the trips within the City of Toronto and those outside of the city. Also, we did not consider the incoming trips from the outside of the city. However, the focus of this study was to reassign the number of outgoing trips in TAZs, and the boundary effect should not impact much on the apportionment results. When we use these numbers to study the interactions of different zones in the future, this limitation should be addressed.

Conclusion

Trip data appointment is essential when rezoning traffic analysis zones. This study systematically examined six methods in apportioning personal trip numbers, considering six scenarios of data obtainment. The apportionment results between the six methods were compared and analyzed using mapping, Getis-Ord G* hot-spot analysis, and the Wilcoxon signed-rank test. By comparing the results between the six methods, we found that the traditional method and the population counted at the DA level method were significantly different ($p < 0.05$) from the residential land use counted method and the integrated method using both residential land use and the population in DAs and the integrated method using both residential land use and the population in DBs. By analyzing the apportioning differences, the integrated method using residential land use and the population at DBs is proposed as the preferred method because populations are specifically located at a finer scale (DBs) based on the spatial distribution of residential areas.

These results provide a useful reference for choosing a method to apportion trip data that ensures data analysis consistency and comparability across different periods when traffic analysis zones are rezoned. Moreover, appropriately apportioned trip data are a basis for valid transportation demand analysis, planning, and policymaking. Furthermore, the apportionment methods presented in this study can be extended to other trip data. However, the conclusion on the different apportionment methods may not apply to other cities or regions due to the differences in geometry of geographic analysis units, land use, and populations′ spatial distribution. Further examination of the combined effect of TAZ zoning and different apportioning methods on transportation planning analysis is needed in the future.

References

Altan, M. F., and Y. E. Ayozen. 2018. The effect of the size of traffic analysis zones on the quality of transport demand forecasts and travel assignments. *Periodica Polytechnica Civil Engineering* 62(4): 971–979.

Alvanides, S., S. Openshaw, and O. Duke-Williams. 2000. Designing zoning systems for flow data. In *GIS and geocomputation: Innovations in GIS 7*, ed. P. Atkinson. London, UK: Taylor and Francis, 115–134.

Beale, L. 2012. *Apportioning population between areas.* https://www.arcgis.com/home/item.html?id=7f299b57007c40e9a1a3ad3f0c4dfe37

Behan, K. M., H. Maoh, and P. Kanaroglou. 2008. Smart growth strategies, transportation and urban sprawl: Simulated futures for Hamilton, Ontario. *The Canadian Geographer* 52(3): 291–308.

Burt, J. E., and G. M. Barber. 1996. *Elementary statistics for geographers*. New York, NY: The Guilford Press.

Cambridge Systematics, Inc., and AECOM Consult. 2007. A recommended approach to delineating traffic analysis zones in Florida. https://www.camsys.com

Chang, K. T., Z. Khatib, and Y. M. Ou. 2002. Effects of zoning structure and network detail on traffic demand modeling. *Environment and Planning B: Planning and Design* 29(1): 37–52.

CHASS Data Centre. 2001. *2001 census data*. Canadian Census Analyser at CHASS, University of Toronto. http://dc.chass.utoronto.ca/census

City of Toronto. 2017. 2016 census backgrounder reports. https://www.toronto.ca/city-government/data-research-maps/toronto-economy-labour-force-demographics

Ding, C. 1994. Impact analysis of spatial data aggregation on transportation forecasted demand. In *Proceedings of Urban and Regional Information System Association (URISA) Conference*, 362–375.

——. 1998. The GIS-based human-interactive TAZ design algorithm: Examining the impacts of data aggregation on transportation-planning analysis. *Environment and Planning B: Planning and Design* 25(4): 601–616.

DMG (Data Management Group). 2001. *Transportation tomorrow survey boundary files: 2001 TTS traffic zones*. http://dmg.utoronto.ca/survey-boundary-files

——. 2006. *Transportation tomorrow survey boundary files: 2006 TTS traffic zones*. http://dmg.utoronto.ca/survey-boundary-files

——. 2016. *Transportation tomorrow survey: 2016 TTS Data Guide*. http://dmg.utoronto.ca/transportation-tomorrow-survey/tts-reports

DMTI Spatial Inc. 2001a. *CanMap RouteLogistics*. Scholars GeoPortal of the Ontario Council of University Libraries.

——. 2001b. *CanMap RouteLogistics user manual v5.1*. Scholars GeoPortal of the Ontario Council of University Libraries.

Esri. 2020a. *ArcGIS Pro: How hot spot analysis (Getis-Ord Gi*) works*. https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm

——. 2020b. *ArcGIS REST API: Data apportionment*. https://developers.arcgis.com/rest/geoenrichment/api-reference

Getis, A., and J. K. Ord. 2010. The analysis of spatial association by use of distance statistics. In *Perspectives on spatial data analysis*, ed. L. Anselin and S. J. Rey. Berlin: Springer, 127–145.

Gorr, W. L., and K. S. Kurland. 2010. *GIS tutorial 1: Basic workbook*. Redlands, CA: Esri Press.

Hanson, S., and G. Giuliano. 2004. *The geography of urban transportation*. 3rd ed. New York, NY: The Guilford Press.

Harder, C., T. Ormsby, and T. Balstrom. 2013. *Understanding GIS: An ArcGIS project workbook*. 2nd ed. Redlands, CA: Esri Press.

Martinez, L. M., J. M. Viegas, and E. A. Silva. 2009. A traffic analysis zone definition: A new methodology and algorithm. *Transportation* 36(5): 581–599.

Miller, E., and R. Soberman. 2003. *Travel behaviour and urban form*. Issue Paper No. 9. Toronto, ON: Neptis Foundation.

Miller, H. J., and S. L. Shaw. 2001. *Geographic information systems for transportation: Principles and applications*. Toronto, ON: Oxford University Press.

Mustafa, R., and M. Zhong. 2014. Challenges and opportunities in applying high-fidelity travel demand model for improved network-wide traffic estimation: A review and discussion. *The Open Transportation Journal* 8(1): 1–18.

Openshaw, S. and P. Taylor. 1982. The modifiable areal unit problem. In *Quantitative geography: A British view*, ed. N. Wrigley and R. J. Bennett. London, UK: Routledge, 60–69.

Rosu, A., and D. M. Chen. 2016. An improved approach for geocoding Canadian postal code-based data in health-related studies. *The Canadian Geographer* 60(2): 270–281.

Statistics Canada. 2001a. *2001 census boundary files: Blocks*. https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2001-eng.cfm

——. 2001b. *2001 census boundary files: Dissemination areas*. https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2001-eng.cfm

——. 2001c. *Block digital boundary file 2001 census*. https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2001-eng.cfm

——. 2001d. *Cartographic boundary files 2001 census*. https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2001-eng.cfm

——. 2006. *Boundary files 2006*. https://www150.statcan.gc.ca/n1/en/catalogue/92-160-G

Viegas, J., L. M. Martinez, and E. A. Silva. 2009. Effects of the modifiable areal unit problem on the delineation of traffic analysis zones. *Environment and Planning B: Planning and design* 36(4): 625–643.

Wong, D., and C. Amrhein. 1996. Research on the MAUP: Old wine in a new bottle or real breakthrough? *Journal of Geographical Systems: Geographical Information, Analysis, Theory and Decision* 3(2): 73–76.

You, J., Z. Nedović-Budić, and T. J. Kim. 1998. A GIS-based traffic analysis zone design: Technique. *Transportation Planning and Technology* 21(1–2): 45–68.