travel
modelling
group

tmg

TRAFFIC ANALYSIS ZONE DEFINITION: ISSUES & GUIDANCE

Eric J. Miller, Ph.D.

March 2021

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING
Transportation Research Institute

UTTRI

TABLE OF CONTENTS

Page No.

Table of Contents 1
List of Figures 1
Acknowledgements 1

1. INTRODUCTION 2
2. MODELLING CONSIDERATIONS 3
2.1 Introduction 3
2.2 Socio-Economic Aggregation 3
2.3 Spatial Aggregation 5
2.4 Short Trips 8
2.5 Centroids & Centroid Connectors 9
2.6 Zone Population 10
2.7 Study Area Definition & External Zones 11
3. GEOGRAPHICAL CONSIDERATIONS 12
3.1 Introduction 12
3.2 Consistency with Input Data Geographies 12
3.3 Consistency with Political & Planning Spatial Units 14
3.4 Conformity with Major Physical Features & Barriers and Other Considerations 15
3.5 Evolution of Future Year Zone Systems 16
4. RECOMMENDATIONS 17
4.1 Discussion 17
4.2 Recommended TAZ Zone Definition Guidance 18
5. NEXT STEPS 19

REFERENCES 20

LIST OF FIGURES

Page No.

2.1: Example of Potential Socio-Economic Bias in a Travel Demand Model 4
2.2: Spatial Aggregation Bias in Transit Mode Choice Modelling 7
2.3: Centroid Connectors & Adjacent Zone Trips 9
2.4: Distribution of 2006 TAZ Population as Measured in the 2016 TTS 11
3.1: Irregular Zone Shape Example 16

AKNOWLEDGEMENTS

The assistance of Gozde Ozonder, James Vaughan and Yang (Luna) Xi in fact-checking and editing this report is very much appreciated.

1. INTRODUCTION

Traffic analysis zones (TAZs) are universally used in travel demand modelling to represent the spatial distribution of trip[1] origins and destinations, as well as the population, employment and other spatial attributes that generate or otherwise influence travel demand. The urban area is divided into a set of mutually exclusive and collectively exhaustive zones. While travel actually occurs from one point in the urban region to another, all trip origins and destinations in a travel demand model are represented at the spatially aggregate level of the movement from an origin zone to a destination zone. These movements are further aggregated within network assignment models as originating and ending at single points within the origin and destination zones – the *zone centroids*.

The road and transit networks coded into the computer network model reflect this zonal aggregation of space in that they are designed to carry the flows between zones, not within zones. Zone centroids (which are the "sources" and "sinks" of trips) are connected to the road and transit networks via artificial links known as *centroid connectors* which are a highly abstracted representation of the actual fine-grained local street network, thus representing another element in the spatial aggregation associated with the use of traffic zones.[2]

The choice of a TAZ system is thus a first, very critical step in travel demand modelling, since it represents the fundamental level of spatial representation, precision and accuracy in the model. Despite this importance, TAZ definition has received very little attention in either the academic or professional literature. A review of the recent academic literature, frankly, provides little practical guidance for operational traffic zone design; nevertheless, several of the citations found are included in the list of references. Similarly, a scan of the Transportation Research Board (TRB), Travel Model Improvement Program (TMIP) and Travel Forecasting Resource (TFResource) websites turns up little in the way of useful guidance. The one excellent exception found to this statement is Cambridge Systematics (2007), which provides excellent, detailed guidance concerning TAZ design (even if it is US-oriented, as well as sometimes Florida-specific).[3] This report is recommended as "required reading" for any agency considering reassessment of its traffic zone system, and it represents the starting point for the discussion presented below. In particular, this report does not attempt to reproduce/summarize the Cambridge Systematics report in any way, but rather augment it with what is hoped to be key points concerning TAZ design from an explicit travel demand modelling perspective.

---

[1] Although activity-based travel demand models are increasing the norm for best practice (including GTHA models such as GTAModel V4 (http://tmg.utoronto.ca/doc/1.6/gtamodel/index.html), this report consistently deals with the trips that emerge from the participation in out-of-home activities.

[2] A few exceptions to these general statements exist. MATSim (https://matsim.org/), for example, does not require a zone system, but rather can accept trip origins and destinations as points in space. In practical applications, however, these usually still are (perhaps small) zone-based. Some land use models such as UrbanSim (https://urbansim.com/) allocate land uses at the parcel level (an extremely fine-grained level of analysis), but usually aggregate these parcels up to TAZs for travel demand modelling purposes.

[3] The Cambridge Systematics report builds heavily upon an earlier major report on travel demand modelling methods, Barton-Aschman (1998).

TAZ system design is very much an exercise of finding the “best” trade-offs among a wide range of criteria, which are often at least somewhat in conflict with one another. These can be broadly divided into two main (but intertwined) categories: modelling considerations and geographic considerations. Sections 2 and 3 discuss each of these in detail. Section 4 then brings these discussions together into a unified set of recommendations for TA zone design. Section 5 concludes the report with a brief discussion of possible next steps in GTHA TAZ evolution.

2. MODELLING CONSIDERATIONS

2.1 Introduction

A zone-based approach to travel demand modelling has many potential impacts on model design and performance. Issues to be considered in TAZ system design from a modelling perspective include:

- Socio-economic aggregation.
- Spatial aggregation.
- Modelling short trips.
- Designing centroids and centroid connections.
- Study area definition and treatment of external zones.

Note that many of these issues involve network design considerations as well as the design of the zone system per se. These issues are each discussed in turn in the following subsections.

2.2 Socio-Economic Aggregation

A classic concern in the use of zone-based models is the potential for significant *aggregation bias* with respect to the representation of socio-economic factors in travel demand models. Given that socio-economic attributes such as income, auto ownership, age, etc. play a major role in determining the generation, distribution and mode shares of trip-making,[4] if zonally aggregated (average) values are used in these models then they generate significantly biased results. Figure 2.1 illustrates this concern for the case of a logit model choice model that predicts the probability of choosing transit for a given trip, with all explanatory factors held constant except for trip-maker income. Let us assume for the sake of discussion that the logit curve perfectly[5] captures trip-makers’ transit choice behaviour with respect to income. As expected from theory, the probability of a person using transit declines as income increases.

Figure 2.1 depicts a case in which two persons, one with low income and one with high income,[6] living in this zone are making a trip to the same destination for the same purpose. (Y_1, P_1) and (Y_2, P_2) are, the income and transit choice probabilities for Person 1 and Person 2, respectively, and $(\bar{Y}, \bar{P})$ are the average zone income and average zone transit choice probability. The key point to note from Figure 2.1 is that it does not lie on the true demand curve. This is a result of

---

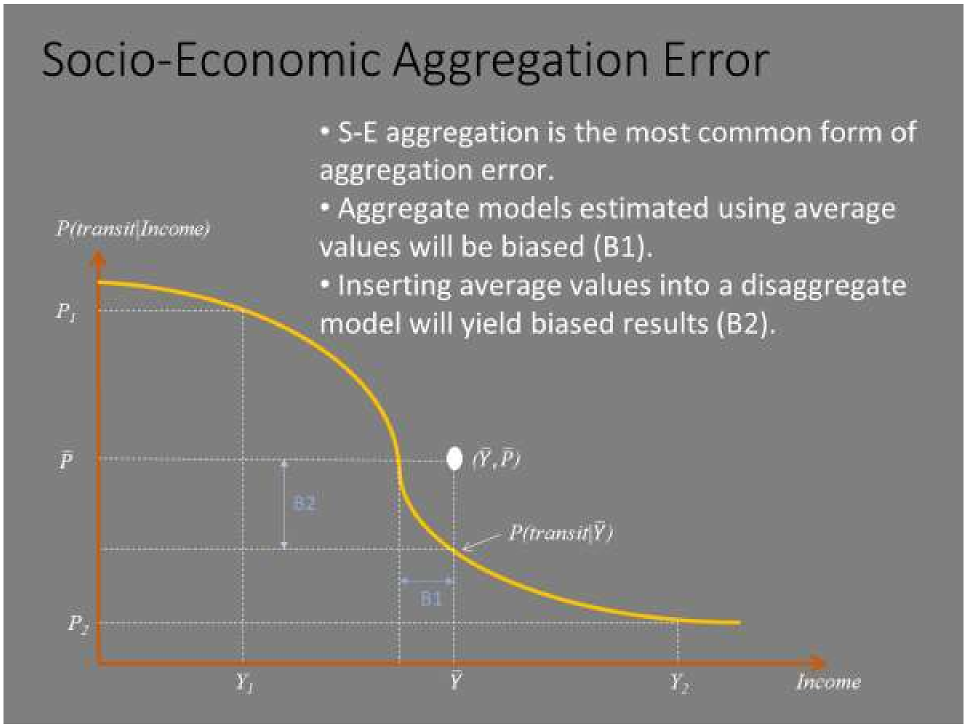
[4] And possibly route choice as well, notably with respect to responses to road pricing and transit fare policies.

[5] I.e., in a probabilistic sense; it is generally impossible to predict a person’s mode choice with absolute certainty.

[6] Or, equivalently, two equally sized groups of persons, one group with the same low income and one group with the same high income.

the demand curve being non-linear in nature.[7] Thus, the use of aggregate (zone average) data in travel demand models is fundamentally problematic[8] in two critical ways.

- Use of aggregate data to calibrate travel demand models introduces an *estimation bias* (B_1 in Figure 2.1), since the demand model is being fit to data points that do not lie on the "true" demand curve. The key point is that the zone average values $(\bar{Y}, \bar{P})$ are completely a function of the *within-zone* heterogeneity of trip-maker attributes (with this information being completely lost through the aggregation process), rather than indicative of actual travel behaviour.
- Use of aggregate data in the application of travel demand models (even if they have been appropriately estimated without bias) introduces a *forecast bias* (B_2 in Figure 2.1).



Socio-Economic Aggregation Error

- S-E aggregation is the most common form of aggregation error.
- Aggregate models estimated using average values will be biased (B1).
- Inserting average values into a disaggregate model will yield biased results (B2).

Figure 2.1: Example of Potential Socio-Economic Bias in a Travel Demand Model

Historically, when constructing aggregate, trip/zone-based travel demand models, one strategy adopted by modellers to reduce socio-economic aggregation bias was to try to define zones that are as socio-economically homogenous as possible. The more homogeneous a zone is, the less the non-linearity in the demand curve matters, and the more representative the zone averages are of the behaviour being modelled. In the limit, if a zone is perfectly homogenous (e.g., in the Figure 2.1 case, everyone in the zone has the same income), then the average values will lie on the demand curve and the data can be used to correctly estimate (and/or apply) the model.

---

[7] Linear regression equations are often used to model trip generation, but these are usually, at best, simple approximations to actual non-linear behaviour. Trip distribution and mode choice models invariably require non-linear specifications (gravity models, logit models, etc.) to capture observed behaviour adequately.

[8] Which is the basis for the argument for the use of disaggregate (microsimulation) models of travel demand (Miller and Salvini, 2002; Miller, 2018).

The problem with this approach is that there is no such thing as a truly homogeneous zone. Even if all households in a zone are perfectly homogeneous with respect to one attribute, say, income (in itself a very unlikely occurrence in practice), they will vary across other attributes, such as household size, age, etc. Further, the distribution of household attributes will inevitably evolve over time and space as households move through lifecycle stages, move about the urban region (as well as in- and out-migration occurring in the region), housing supply changes, etc. Thus, even if a zone system for a base year could be designed that was (relatively) homogeneous within its zones, this homogeneity is unlikely to be preserved as the urban region evolves over time.

Fortunately, this problem is largely eliminated with the adoption of a disaggregate, agent-based microsimulation (ABM) approach to modelling. If each person/household within a zone is being individually modelled, with their individual social-economic attributes, then the "true" demand curve in Figure 2.1 is being used in the model, and unbiased model parameters and predictions can be (in principle at least) obtained. This, however, shifts the problem to ensuring that an "accurate" population synthesis procedure is available to generate the disaggregated population for each TAZ. This issue is discussed further below in Subsection 3.2 with respect aligning TZ boundaries with the zone system(s) that usually provide the data used in the synthesis procedure.

2.3 Spatial Aggregation

A second form of zone-related aggregation of importance in travel demand modelling is the assumption that all trips start and end at zone centroids. This implicitly assumes that:

1. The zone's population, employment and possibly other "trip generators" are uniformly distributed over the zone's area.
2. The zone centroid (however defined in practice), is representative of travel to/from anywhere in the zone.
3. No aggregation bias is introduced into the modelling through the concentration of all trip origins and destinations to the zone centroids.

As with population socio-economics, the distribution of people, jobs, etc. within a zone will rarely, if ever, be perfectly homogeneous. Households and employment, in particular, generally will not be uniformly mixed together. For example, employment will often be concentrated along major streets (which often define zone boundaries and so the employment is located along the edges of the zones), while residential housing may be located more within the zone's interior along local streets.[9] To the extent that relatively homogeneous "employment areas" and "residential areas" exist, it does make sense to try to separate them as much as possible (subject to trade-offs with other zone definition criteria), for at least two reasons:[10]

- The "trip generation" characteristics of these two "and uses" are quite different.
- Reduction in the number of intrazonal trips (discussed further in the next subsection).

---

[9] Exceptions to this generation obviously exist, especially with respect to "mixed use" development, again often along relatively major roads.

[10] The fact that the composition of these zones might change over time (as discussed in the previous section with respect to zonal socio-economic attributes) does not usually invalidate this recommendation. At a minimum, this will introduce useful enhanced homogeneity in the base case for model development and will still result in better zone definitions in future years than might be the case otherwise.

Numerous approaches to defining zone centroid locations exist. The most common are the geographic centroid of the:

- Zone area.
- Zone population or jobs or some combination thereof.
- Geocoded locations of the 24-hour trip origins and destinations (for all purposes) observed in a base year travel survey (e.g., TTS[11]).

The last is the approach adopted in Travel Modelling Group and Data Management Group zone centroid definitions primarily because it represents a robust way of accounting for both population and employment / trip generation locations within the zone. Further points to note concerning centroid definitions are:

- Clearly no definition is "perfect". In particular, as already noted, population, employment and, hence, trip centroids may well shift as the zone evolves over time.
- Of the approaches listed, it is suggested that the simple area centroid is the least preferred (even if it does remain constant over time), since it will invariably be the least representative of the zone's trip generation "centre of gravity".
- A not completely uncommon practice in network coding is to move centroids in some cases so that when zone numbers (or other zone attributes) are plotted on network maps that they "fall nicely" within the zone's boundaries. This practice should never be used, since it clearly invalidates any sensible interpretation of the zone centroid's spatial "meaning: and could well bias modelling results.

Very similar to the socio-economic aggregation problem discussed in the previous section, "pure" spatial aggregation bias can also be introduced into travel demand models through the concentrating of all trip origins and destinations within zone centroids. This is perhaps most critical in the case of transit mode choice modelling, which depends very heavily on the walk access/egress times which potential transit riders face in their decision-making. Figure 2.2 illustrates this issue. In this figure, the logit transit mode choice model from Figure 2.1 is redrawn for the case in which all explanatory factors are held constant except for the time require to walk from the trip-maker's home locations within the zone and the indicated bus stop. As expected from theory, the longer a person has to walk from home to bus stop, the less likely transit is to be chosen for the trip. In all zone-based models (even disaggregate ABMs such as GTAModel), the model assumes that everyone will walk 10 minutes from the zone centroid to the bus stop. But, if trip-makers' homes are actually uniformly distributed over the zone, then some who live closer to the bus stop will have much less than a 10-minute walk (and so be much more likely to take the bus), and some will live much farther away from the bus stop (and so will be much less likely to take the bus). As illustrated in Figure 2.2 (and, again, very similar to the socio-economic aggregation case), the non-linear nature of the travel demand function, means that using the centroid "average" walk time will result in model bias – both in estimation and application.

---

[11] Transportation Tomorrow Survey (http://dmg.utoronto.ca/transportation-tomorrow-survey/tts-introduction).

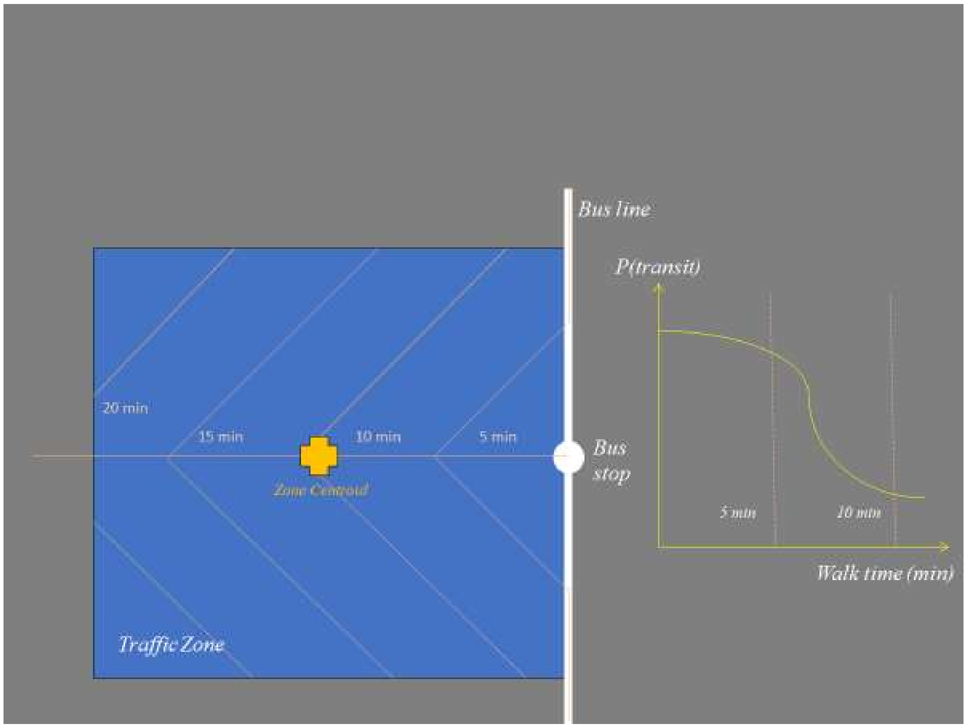


Figure 2.2: Spatial Aggregation Bias in Transit Mode Choice Modelling

As long as a zone/centroid-based network assignment modelling structure (the standard in virtually all operational model systems, including all GTHA model systems), this spatial aggregation problem cannot be completely eliminated, only mitigated to varying degrees. While various two strategies for minimizing this aggregation bias exist, by far the most common strategy, is to simply use smaller zones whenever possible, especially in denser, more transit-oriented locations where walk access effects are most critical. The primary constraint on using small zones is the total number of zones, nodes (centroids and “regular” network nodes) and links (centroid connectors and “regular” network links) required to implement the zone system. Network modelling package licences typically have upper bounds on the number of nodes, links and/or centroid connectors that are permitted, with “larger” licences costing more to purchase and maintain, Model computational and data storage burden using also grows by $O(n^2)$, where n is the number of zones, which can be a practical concern. Smaller and greater number of zones may also add to the population synthesis challenge, depending on the level of detail in the data available for undertaking the synthesis.

Despite these practical constraints, as a general principle, working with smaller zones whenever possible not only alleviates the transit walk access/egress spatial aggregation problem,[12] but virtually all other issues discuss in this report, since it moves the spatial representation close to the “ideal” of modelling point to point travel. With continuously increasing computing power, data storage capacities and supporting data for model building, practical constraints on zone size are reducing.

---

[12] And similar network modelling issues, such as representing walk and bicycle travel times.

2.4 Short Trips

Zone/centroid-based network models introduce two additional modelling challenges in terms of dealing with “short” trips, in particular *intrazonal* trips (in which the origin and the destination are both located within the same zone) and *adjacent zone* trips (in which the destination zone shares a boundary with the origin zone).

Regardless of trip mode (auto, transit, etc.), centroid-based assignment models cannot deal with intrazonal trips since the origin and the destination are located at the same point in space (the centroid of the zone in question). Intrazonal trips are thus stripped out of origin-destination (O-D) matrices before these matrices are fed into the assignment algorithm Thus, no time travel times (or other level of service (LOS) measures such as travel cost) are generated by the assignment model. This creates difficulties for modelling intrazonal trips within both trip distribution and mode choice models, which depend on modal LOS attributes as explanatory variables.

Historically, trip/zone-based models often did not model intrazonal trips at all or treated them in a very simplified way, such as by estimating them at the trip generation stage and then “stripping them out” of further analysis prior to the trip distribution and mode choice model stages. This clearly is not suitable for modelling active transportation (walking and biking) or the impact of mixed-use neighbourhoods (or other spatial densification policies), among other current policy concerns. Instead, intrazonal trips should be maintained throughout the modelling process. This requires auxiliary methods for computing average intrazonal travel times and costs, by mode. These are usually quite simplified calculations based on assumed average trip lengths as a function of the zone’s area.

A second, often-overlooked impact of intrazonal trips not “appearing” within a centroid-based network assignment is that this will underestimate the number of trips using the modelled road network (Manout and Bonnell, 2019; Manout, et al., 2021).

Clearly, the smaller the zones, the less problems that exist with respect to modelling intrazonal trips. First, fewer intrazonal trips will occur as the zone size decreases. Further the intrazonal modal options will generally decrease as size sizes are reduced: transit in particular, but also auto become much less likely modes for the very short distances involved in small-zone intrazonal trips, and bicycle and (particularly) walk modes become predominantly likely.

Similar, but much less discussed issues exist with adjacent-zone trips. As illustrated in Figure 2.3, typical centroid connector configurations result in adjacent zone trips only using centroid connectors and never use “actual” roadway (or transit line) links for the trip, as modelled within the network. For auto-based trips, this may be a somewhat “adequate” representation, since these trips may often be made on local streets that are not explicitly represented in the network model. But there will still certainly be at least some under-representation of volumes on the modelled links, similar to the intrazonal case.

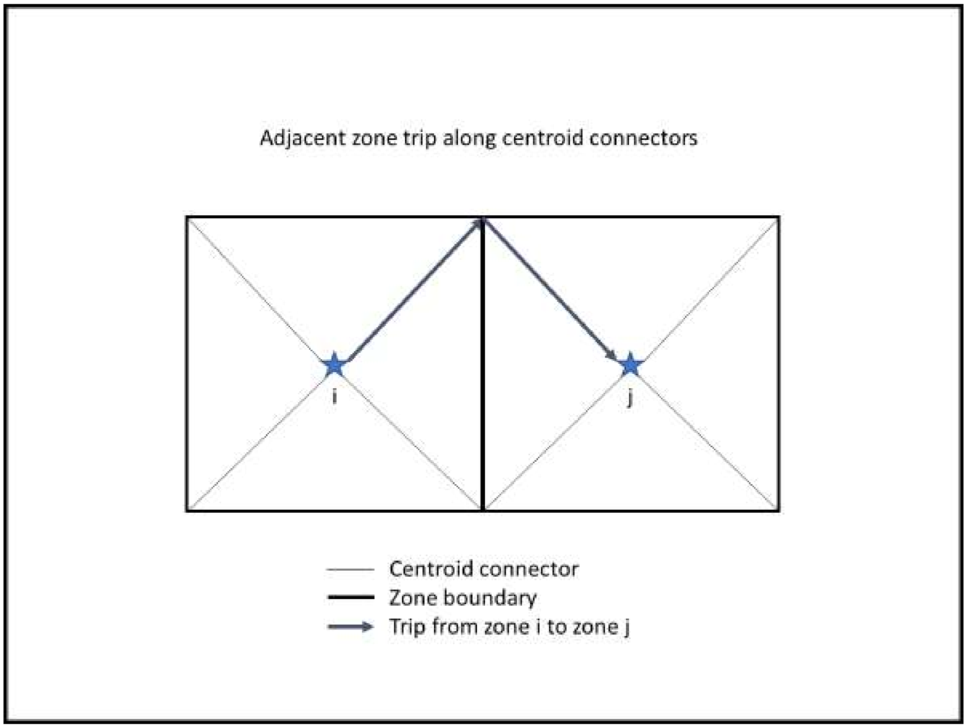


Figure 2.3: Centroid Connectors & Adjacent Zone Trips

The adjacent-zone problem is more critical to transit assignment, since the "all centroid connector" path will yield zero in-vehicle travel and wait times, thereby significantly over-estimating the attractiveness of transit for adjacent-zone trips, when, in fact, these short trips are typically less attractive to be made by transit than by walking, biking and even auto modes. Solutions for mitigating this problem include:

- Not allowing centroid connectors connect with one another on the road network (although this vastly proliferates the number of network nodes required).
- Implementing a constraint within the transit assignment algorithm to enforce a rule that all transit trips must, in fact, board a transit vehicle within the assigned path (Emme can be configured to enforce such a constraint).
- Introduce an "adjacent-zone dummy variable" into the mode choice model to capture the average disutility of making such short trips by transit.
- Introduce a constraint that makes an "all walk on network" path (including use of just centroid connectors) infeasible for transit trips. I.e., all transit trips in the Emme assignment must board at least one transit line. This is the approach currently used in GTAModel V4.

Note that this is one problem that is not reduced through the use of smaller zones, although the problem generally does not "get worse" as zone sizes shrink.

2.5 Centroids & Centroid Connectors

Several other trip assignment issues exist with respect to centroid connectors (and, implicitly, the location of centroids). These are intrinsic to the centroid-based network structure and so can only be mitigated, rather than completely eradicated, by careful centroid connector specifications. These are:

- "Point loading" on the network. Actual traffic from a given zone will load onto an arterial road from the zone's local street system at multiple points. The use of centroid connectors as an abstract replacement for the actual set of local streets results in loading all flow to/from the zone onto the few nodes and links that connect directly with a centroid connector. This artificially increases the level of congestion on these links, possibly distorting overall travel time and roadway congestion calculations. Use more centroid connectors per zone and "mid-block" connections rather than (or in addition to) zone vertex connections helps reduce these impacts.[13]
- The (inevitably) limited number of centroid connectors per zone exacerbates local effects in the network assignment (for both road and transit) networks, in that it will tend to load all flow for a given O-D pair to the centroid (or perhaps two centroids at best) that is on "the minimum path" between the origin and destination centroids.

Finally note that it is very important to allow transit users to be able "walk on the network" beyond their immediate centroid connections to transit line. That is they must be able "walk past" the transit node connected to a centroid connector to be able to access/egress the transit network at a more distant, but more attractive stop/station. An example in which this network coding is essential to adequately capture transit user route choice is a trip from a downtown zone immediately west of Bay Street in downtown Toronto travelling to North York City Centre. Clearly the transit rider will prefer to "walk past" the Bay bus and access the Line 1 subway for a direct trip to his/her destination, rather than be forced to take the Bay Bus north to Bay Station, transfer to Line 2 eastbound for one stop and then transfer again at the Bloor-Yonge interchange station to travel north on Line 1. A second important case is one in which a person lives on one side of a transit fare boundary, such as someone living on the north side of Steeles Avenue in the Region of York who wishes to travel to downtown Toronto. The person will inevitably "walk across the street" to directly access a TTC vehicle, rather than paying a double fare by boarding a southbound bus from York Region to the City of Toronto.

2.6 Zone Population

In addition to geographic size, the zone population[14] is another factor that is of some concern in TAZ design. Very large populations may result in large "point loads" on links connected to centroid connector, as discussed in Section 2.5. In any aggregate model they may dominate parameter estimation. And they will also be more heterogeneous in population composition. On the other hand, zones with an extremely small number of residents will have higher statistical variability with respect to population synthesis, although it is inevitable that many zones ("employment zones", etc.) will have small numbers or (even zero) of population.

Figure 2.4 plots the histogram of 2006 TAZ population as measured in the 2016 TTS. As indicated in the figure, the average zone size (excluding zones with no population) is 3,581

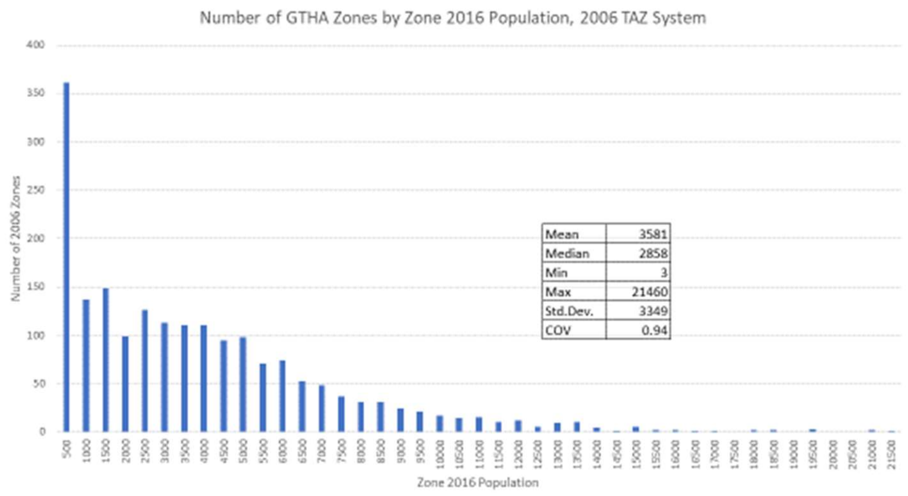
---

[13] This is primarily a road network modelling problem, although it can also be of some importance in transit networks using congested assignment methods.

[14] Similar issues can exist with zonal employment levels. We focus on population in this discussion.

persons (median 2,858), with zone populations ranging from 3 to 21,460 persons. 26% of zones have 1000 persons or less, while 73% have 5000 persons or less.

No absolute guidance exists with respect to recommended zone population size, and is probably unlikely to be possible given the large number of factors that need to be considered in TAZ definition. But, it is suggested that zones with populations of over 5,000 persons should be considered for possible splitting into two or more smaller zones, if all other desirable criteria for zone definition are met.



Mean	3581
Median	2858
Min	3
Max	21460
Std.Dev.	3349
COV	0.94

Each bar represents populations in 500 person increments (0-500, 500-1000, etc.
Zones with zero population in the TTS are excluded from the plot.

Figure 2.4: Distribution of 2006 TAZ Population as Measured in the 2016 TTS

2.7 Study Area Definition & External Zones

Defining the boundary of a travel demand model study area is an essential first step in any modelling exercise. This will depend on the intended purposes of the model and the urban configuration, among other factors. The boundary is often defined on the basis of political jurisdictions, as is the case with both the GTAModel (the 6 six regional municipalities and single-tier cities comprising the GTHA) and GGHMV4 (the municipalities comprising the Greater Golden Horseshoe).

Spatial interactions such as travel, however, do not cease at the study area boundary. Trips to and from adjacent external zones will occur, and can be of importance in terms of generating flows on the study area road and transit networks, especially within the outer regions of the study area. To account for these flows within the model system, a "halo" area surrounding the primary study area, consisting of a set of *external zones* needs to be included in the zone system that can generate trips to and from the primary study area to this external, "halo" region surrounding the study area. Generally, these external zones are defined at a much grosser scale than the study area internal TAZ, often the municipality level (township, county, etc.). The road and transit networks connecting these external zones to the primary study area are similarly defined in a very aggregated, simplified way. And even in advanced ABM, activity-based travel modes, the external-to-internal (E-I) and internal-to-external (I-E) trips are modelled in a simplified, aggregate trip basis.

3. GEOGRAPHICAL CONSIDERATIONS

3.1 Introduction

In addition to modelling considerations, a number of practical issues exist with respect to the geographical definitions of TAZ boundaries. These include:

- Consistency with input data geographies.
- Consistency with required output spatial units, notably political boundaries.
- Conformity with major physical features and barriers.

Each of these sets of issues is discussed in the following subsections. Another practical consideration is the definition of appropriate standard aggregations of TAZs for various purposes, which is discussed in this section's final subsection.

3.2 Consistency with Input Data Geographies[15]

Data concerning population, employment, activity "attractor" (retail floorspace in a zone, etc.) that are used in the construction and application travel demand models are obtained from many sources. Some data may be available in point form (e.g., GIS[16] "Place of Interest" (POI) datasets) or at very fine spatial areas (e.g., land use data at the individual parcel level). But the majority of such data are only available in an aggregated form, including aggregated to some spatial zone system, often due to privacy/confidentiality concerns. Use of such data requires being able to map the data's zone system into the TAZ system. This, obviously, is easiest to do if the data's zone system and the TAZ system are fully nested within one another (e.g., the data zones are wholly contained within a single TAZ or vice versa).

By far the most common source of input information for the GTHA is the Canadian Census, undertaken every five years by Statistics Canada. Census data is a primary source of information concerning population and population attributes (including income, labour force participation, demographics and household structure), employment and employment attributes, place of residence – place of work (PoRPoW) linkages, work trip commuting mode choice and housing stock attributes. This wealth of information is used in a variety of ways in travel demand model building, including providing a primary input to many population synthesis procedures.

Given these considerations, a TAZ system ideally should align "seamlessly" with Census geography, so that either several TAZ zones aggregate to single Census spatial unit or several

---

[15] The other major source of input data for model development is, of course, travel survey data (such as TTS), as well as road and transit count data. These are generally not problematic from a TAZ design perspective. Survey data is essentially always geocoded and so can be aggregated to any required zone system as long as a GIS boundary file is available for this zone system. Count data similarly can essentially always be mapped to network nodes and links within the model system's computerized network model.

[16] Geographic Information System.

Census spatial units aggregate to a single TAZ. Tabulated[17] Census data are available to varying degrees at five levels of geography:

- Dissemination Areas (DAs).
- Census Tracts (CTs).
- Census Sub-Divisions (CSDs).
- Census Metropolitan Areas & Census Agglomerations (CMAs, Cas).
- Custom Tabulations (get counts for the GTHA).

CSDs and CMAs/CAs are generally much too large aggregations for most travel demand modelling applications. Also note that these definitions are based on labour market commuter sheds, rather than regional municipal boundaries. For example, the GTHA corresponds approximately to the combined Toronto, Hamilton and Oshawa CMAs, but the mapping is not perfect, with a few small parts of the CMAs not being included in the GTHA, while the three CMAs similarly do not provide complete coverage of the entire GTHA.

Obtaining custom tabulations coded to a study area's TAZ system is generally feasible to do, but can be very expensive, and is generally limited to a subset of key variables, due both to cost considerations and Statistics Canada privacy/confidentiality constraints. Also, a complete set of all variables of possible interest in travel demand modelling and analysis usually is not feasible to obtain.

Thus, DA- and CT-level tabulations are the primary source of Census data for must travel demand modelling purposes. DAs generally are smaller than TAZs and CTs on average are larger than TAZs. Ideally, a TAZ system should be designed so that each TAZ consists of one or more DAs (with no DA overlapping two or more TAZs) and one or more TAZs map into each CT (with no TAZ overlapping two or more CTs). This tight mapping between TAZs and either DAs or CTs generally has not happened in the GTHA, however, for several reasons, including:

- Census boundary definitions sometimes are not compatible with TAZ boundary requirements with respect to dealing with physical barriers, transportation features, etc. (see Subsection 3.4).
- Historically CT boundaries have changed over time, making establishing a consistent zone system over time challenging. Since 2006, however, CTs have only be subdivided into smaller sub-CTs as needed, without changing the original, larger CT boundaries, so this is now less of an obstacle.
- DA and CT boundaries are not always mutually consistent.
- Census boundaries do not always respect sub-CSD municipal geographies (e.g., ward boundaries; see Subsection 3.3).
- Census boundaries are population based and so can be less conducive to dealing with primarily employment-based zones.

---

[17] Anonymized Census records for containing full information for both individual persons and households are also available for a small sample of households within a CMA, but without any sub-are geography attached. Thus, except for a limited number of special purposes, these data are not directly usable in constructing travel demand models, and so these data are not considered further in this report.

Despite these challenges, improved correspondence between TAZ definitions and DA/CT boundaries is a highly desirable objective, wherever possible. At a minimum, very explicit, standardized GIS-based methods for mapping between Census and TAZ geographies should be established and consistently employed whenever needed.

3.3 Consistency with Political & Planning Spatial Units

Travel demand model outputs can be tabulated, mapped, etc. in a variety of ways, at both the traffic zone level as well as any desired aggregation of TAZs (given that the TAZ may often be too fine a level of spatial detail for many purposes). One important aggregation is to various political and planning zone systems, which are often useful for displaying results to a variety of audiences, including politicians and the general public, as well as planning staff.

Political boundaries of interest generally are wards, area/single-tier municipalities (which generally correspond to Census CSDs) and regional municipalities. At a minimum, TAZs should respect area and regional municipal boundaries. Respecting ward boundaries is also very desirable, but may be more difficult, since these do change from time-to-time. At a minimum, as with mapping to/from Census geographies a standard GIS-based method for mapping TAZs to wards should be developed and used whenever needed.

In addition to political spatial units, planning departments typically maintain various spatial zone systems for a variety of planning purposes. These can include: “neighbourhoods”, “priority development areas”, “transportation investment zones” (TIFs), etc. Definitions of such special-purpose planning zone systems generally will not respect traffic zone boundaries in all cases. They also may change over time, or even go out of usage as planning needs and issues evolve. Hence, these zone definitions usually will not play a significant role in TAZ definitions. One exception to this general statement is if it is clear that certain neighbourhoods/districts will be of considerable planning concern for a significant period of time going forward and if clear criteria for neighbourhood definition exist (e.g., historically low-income and/or otherwise disadvantaged neighbourhoods), then these criteria should be seriously considered when defining/redefining TAZs so that travel demand modelling/analysis can be as helpful as possible in addressing these policy concerns.

By far the most common (and most useful) aggregation of TAZs used by GTHA transportation planning agencies for decades is the 46-zone *Planning District* (PD) system. This consists of the 16 historically defined PDs within the (amalgamated) City of Toronto (formerly the Regional Municipality of Metropolitan Toronto), 6 historically defined PDs within the (amalgamated) City of Hamilton (formerly the Regional Municipality of Hamilton-Wentworth) and 24 area municipalities (CSDs) within the Regional Municipalities of Durham, Halton, Peel and York. The PD system has been very useful in displaying travel survey and modelling results at a spatial scale that is much easier to assimilate than the much finer TAZ level, providing in some cases travel model parameters and/or control inputs into population synthesis procedures for TAZs contained within a given PD, and establishing “K-factors” in trip distribution models, among other applications. Thus, while any custom aggregation of TAZs is always possible for special purposes, the availability of a standard PD system is viewed as an essential adjunct to the TAZ system itself.

The current 46-PD system, however, clearly has its weaknesses and needs to be updated. The City of Mississauga (the second largest City in the GTHA), for example, is represented at the moment by a single PD. The 16 City of Toronto PDs are quite idiosyncratically defined relative to current planning needs, neither corresponding to political wards or other standard geographies. Thus, a need exists to revisit and redefine the GTHA PD system. Criteria for this redefinition should include:

- A relatively limited number of PDs, so that interpretability of results remains strong. 50-60 PDs probably is an appropriate range to consider for the GTHA.
- TAZs must aggregate cleanly to PDs.[18]
- PDs should conform to municipal boundaries.
- While PDs are clearly too large to be homogeneous, they should still be meaningful as "large neighbourhoods" which can act as "control totals" for finer-scale TAZ calculations and which are meaningful to lay people (politicians and public) viewing PD-level results.

3.4 Conformity with Major Physical Features & Barriers and Other Considerations

At least four other important considerations in TAZ design exist. First, it is essential that any traffic zone system conform to natural features and boundaries, such as rivers (and other bodies of water), railway lines, limited access highways, etc. – i.e., any feature which acts a physical barrier travel, either by foot or by vehicle. Zones must not straddle such barriers, since it is impossible to construct meaningful centroid connectors (and, indeed, meaningful centroids) in such cases.

Second, areas containing special, homogeneous land uses should be defined as unique zones. Many examples of such land uses exist, including: cemeteries, parks, undevelopable land (flood plains, etc.), areas designated for heavy industry, purely residential areas, special-purpose areas such as airports, hospital districts, etc. In Section 2.2 it is argued that homogenous zones essentially never exist in terms of their socio-economics, and so this should not be a primary design concern – at least for microsimulation-based demand modelling. But zones that are homogenous with respect to land use (and are likely to remain relatively homogenous over time) are very common, and defining zones to exploit these homogenous areas is extremely useful. In particular, two primary advantages of zones comprising relatively homogeneous land uses are:

- The more that residential and employment/commercial land uses can be separated, the less intrazonal travel there is likely to be, at least with respect to work trip commuting.
- Activity/trip generation can generally be more accurately predicted – to the extent that land use provides a useful explanatory variable for explaining generation rates.

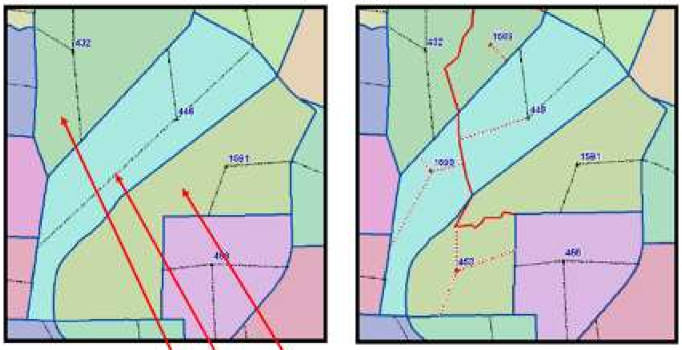
Third, in addition to major transportation infrastructure such as railways and expressways creating physical barriers, in general it is important that the spatial detail of the transportation network be consistent with that of the TAZ system, and vice versa. In particular, road links and transit lines generally should not travel through TAZs, since this creates problems with centroid locations and centroid connector definitions. That is, if a network link passes through a TAZ it effectively subdivides the TAZ into two zones, with two centroids and two sets of associated centroid connectors. In such cases in which this link is deemed essential for properly representing the transportation network, one generally should define two TAZs, one on either

---

[18] Note that this means that PDs will map into Census geography to whatever extent that TAZs do.

side of the network link. Note that this represents yet another repercussion of zone/centroid-based network modelling.

Finally, irregular-shaped zones should be avoided if at all possible. TAZs ideally should be approximately rectangular (and preferably approximately square) in shape. As illustrated in Figure 3.1, highly irregular shapes (even very elongated rectangles) result in awkward centroid connector geometries and will exacerbate spatial aggregation issues in network modelling.



Irregular-shaped zones 432, 446 & 1591

Redraw (red line), creating three new zones: 453, 1569 & 1590

Figure 3.1: Irregular Zone Shape Example

3.5 Evolution of Future Year Zone Systems

Given that the urban area will inevitably grow in population and employment over time, especially in more suburban and rural areas, it is often necessary to subdivide a base-year zone system into finer zones for future year forecasts. The typical case is a large rural zone with minimal population in the base year that is projected to be developed in 20- or 30-years' time. Keeping this need in mind, the following guidelines should be followed whenever possible:

- Sub-divide existing zones; avoid redrawing original zone boundaries whenever possible. This way, the new zone system will aggregate directly to the original base zone system when needed for comparisons to the base year.
- In designing the base-year zone system, try to anticipate where future growth is likely to occur, and, hence, how the base-year zone is likely to be sub-divided to best capture this growth within well-defined sub-zones.
- Remember that the local street network will also have to be augmented to be consistent with the new development and zone system. In particular, the new sub-zones will need well-defined centroid connectors to the road and transit networks.

- The new, sub-divided zones will need to adhere to all the regular guidelines for zone design discussed in this report, as well as in Cambridge Systematic (2007).

Note that an ABM demand model formulation helps guard against overly "locking into" into a specific zone system. That is, it minimizes the amount of socio-economic/spatial aggregation bias that is "baked into" the model parameters. Thus, the model system should be relatively independent of the zone system details and so applicable to the modified future year zone system without serious concern about being "out of calibration" due to the use of the new zone system.

4. RECOMMENDATIONS

4.1 Discussion

Summarizing this report's discussion to this point, two "meta-issues" with respect to TAZ design exist: minimization of spatial aggregation biases due to the inherent structure of zone/centroid-based transportation network modelling (Section 2), and consistency with input data sources and other geographic constraints (Section 3).

Smaller zones consistently reduce spatial aggregation biases in route choice, for both road and transit assignments, but most critically in the case of transit assignment. As a result, they will reduce level-of-service (LOS) related biases in mode choice models as well, which depend so heavily on accurate modal LOS explanatory variables. The ability to accurately predict small area population (and employment) attributes, and hence activity/trip generation, on the other hand, tends to decrease with smaller zones, as stochastic variations in these predictions become more prominent within smaller populations. Note that mode choice models are also dependent on socio-economic (S-E) trip-maker attributes, and hence will also be subject to this increased uncertainty concerning these attributes as zone sizes shrink. The trade-offs between enhanced LOS accuracy and possible increased uncertainty in S-E attributes as zone sizes decrease is not well understood; this represents a potentially fertile area for future investigations.

The relationship between zone size and the accuracy of spatial location/destination choice models (Place of-Residence – Place of Work linkages, non-work/school destination choice, etc. – aka trip distribution) is complex. Clearly, one expects better predictive accuracy at larger spatial scales, such as Planning Districts, with increasing uncertainty in predictions at finer levels of geography. The spatial accuracy of destination choice models is a surprisingly under-researched topic, with operational models often being largely evaluated in terms of comparing predicted and observed trip length frequency distributions and screenline counts, rather in terms of O-D trips at the TAZ or CT level. Limited evidence, however (cf. Smith and Hutchinson, 1979; Wang and Miller, 2014; Elgar, et al., 2015), indicates that predictive accuracy of typical models is not high at the TAZ/CT level. Reducing zone sizes per se will not generally improve this performance. One potentially promising approach to this problem which has not been explored to any significant effect in the literature is to adopt a hierarchical (e.g., nested logit) approach, in which the upper level is the choice of a larger geographic area (perhaps a Planning District or something similar) and the lower level is the choice of a finer geographic area (such as an individual TAZ within the PD). This approach might be able capture the heterogeneity in land use, etc., as well as the precision of TAZ-to-TAZ LOS attributes, within a parsimonious and conceptually attractive structure.[19]

As noted in Section 2, practical computational limits also exist with respect to the number of traffic zones that can be handled within reasonable computing times, notably with respect to both trip assignment and destination choice models. There are 9344 DAs in the GTHA, which is far too many to be practical as the level of spatial detail for general O-D calculations. DAs are also the smallest geographic unit for which Census data are generally available. Thus, it is highly arguable that the DA represents a fairly "hard" lower bound on spatial precision in large-scale regional transportation models. An interesting research question would be to systematically explore how DA's might be best aggregated to develop a computationally tractable traffic zone system that minimizes the various spatial aggregation biases discussed in this report.

4.2 Recommended TAZ Zone Definition Guidance

In addition to the very good guidance provided by Cambridge Systematics (2007), guidance for traffic zone design emerging from this report's discussion includes the following:

1. Zone size is a critical determinant for spatial aggregation issues, especially wrt to modelling transit (and active) trip-making. Keep zone sizes as small as possible, commensurate with computational and other constraints.

2. Consistency with Census small-zone geography (DAs and CTs) is highly desirable and should be maintained wherever possible (recognizing that some deviation from perfect consistency is often inevitable).

3. Consistency with major political boundaries is essential and with key planning boundaries is also highly desirable.

4. Anticipate and plan for sub-division of suburban and rural zones to respond to future-year growth.

5. Avoid use of irregularly-shaped zones.

6. Whenever possible, define zones so that they contain homogenous land uses (especially non-trip generating ones). Separating residential zones and employment zones assists in reducing intrazonal travel (at least for work commuting trips).

7. Consistently locate zone centroids at the "activity centre of gravity" of the zone, as defined by the centroid of trips ends as observed in a base year survey.

8. Zones with very large population or employment should be considered for splitting into two or more zones, provided that all other criteria for zone design are met.

---

[19] It also represents a theoretically attractive approach to dealing with the *modifiable unit area problem* (MAUP) of dealing with zones of varying sizes, as well as spatial autocorrelation issues. Cf., among others, Ferguson and Kanaroglou (2010).

In addition, TAZ-related recommendations concerning network coding and trip assignment modelling to reduce spatial aggregation biases are:

9. Consistently locate zone centroids at the “activity centre of gravity” of the zone, as defined by the centroid of trips ends as observed in a base year survey.

10. Take care in following best practice in terms of centroid connector definitions.

11. Allow transit users to “walk on network” to connect with transit access/egress points other than those that are directly connected to centroid connectors.

12. Explicitly deal with both intrazonal and adjacent zone trips in the network assignment and the overall demand modelling process to minimize travel time calculation inaccuracies.

5. NEXT STEPS

The next step for the Travel Modelling Group (TMG), in collaboration with the Data Management Group and partner agencies, is to review current zone systems with respect to the criteria established in this report. Based on this review changes where deemed necessary will be recommended. The GTAModel external zone system will also be reviewed and updated as required. An updated Planning District (PD) system for the GTHA will also be developed. These activities will result in a new standard TMG TAZ and PD system for travel demand modelling purposes by the end of 2021.

In addition, resources will be sought to undertake a research project investigating the zone size / spatial aggregation / computing burden problems discussed in this report in the hope of further refining generalizable guidance concerning TAZ definition and zone-based network modelling.

REFERENCES

Altan, M. and Y.E. Ayözen (2018) "The Effect of the Size of Traffic Anlysis Zones on the Quality of Transport Demand Forecasts and Travel Assignments", *Periodica Polytechnica Civil Engineering*, 971-979, doi:10.3311/PPci.11885.

Barton-Aschman Associates, Inc. (1998) *NCHRP Report 365 – Travel Estimation Techniques for Urban Planning*, Washington, DC: National Cooperative Highway Research Program.

Cambridge Systematics (2007) *A Recommended Approach to Delineating Traffic Analysis Zones in Florida*, report prepared for the Florida Department of Transportation Systems Planning Office.

Ding, C. (1998) "The GIS-based human-interactive TAZ design algorithm: examining the impacts of data aggregation on transportation-planning analysis', *Environment & Planning B*, 25, 601-616.

Elgar, I., B. Farooq and E.J. Miller (2015) "Simulations of Firm Location Decisions: Replicating Office Location Choices in the Greater Toronto Area", Journal of Choice Modelling, 17: 39-51.

Ferguson, M.R. and P.S. Kanaroglou (2010) "Representing the Shape and Orientation of Destinations in Spatial Choice Models", *Geographical Abakysis*, 30(2), 119-137.

Ghadiri, M., A.A. Rassafi and B. Mirbaha (2019) "The Effects of Traffic Zoning with Regular Geometric Shapes on the Precision of Trip Production Models", *Journal of Transport Goegraphy*, 78, 150-159.

Hadavand, S. F. Tabesh and H. Motevalli (2018) "Presenting traffic Analysis zone method for origin-destination statistics operations in Iran", *International Journal of Current Engineering and Technology*, 8:1, 10-16.

Haider, M. and T. Spurr (2006) "The Design and Development of Large-Scale Traffic Assignment Models Using Geographic Information Systems", presented the 2006 Transportation Research Board Annual Meeting, Washington, D.C., January.

Miller, E.J. (2018) "Agent-Based Activity/Travel Microsimulation: What's Next?", in Briassouli, et al. (eds), *Spatial Analysis: Tools and Land Use, Transport and Environmental Applications*, Springer, 119-150.

Miller, E.J. and P.A. Salvini (2002) "Activity-Based Travel Behavior Modeling in a Microsimulation Framework", invited resource paper, Chapter 26, in H.S. Mahmassani (ed.) *In Perpetual Motion, Travel Behavior Research Opportunities and Application Challenge*s, Amsterdam: Pergamon, 533-558.

Manout, O. and P. Bonnel (2019) "The impact of ignoring intrazonal trips in assignment models: a stochastic approach", *Transportation*, 46, 2397-2417.

https://doi.org/10.1007/s11116-018-9951-y

Manout, O., P. Bonnel and F. Pascull (2021) “Spatial Aggregation Issues in Traffic Assignment Models”, *Netowrks and Spatial Economics*, 21, 1-19. https://doi.org/10.1007/s11067-020-09505-6

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

Smith, D.P. and B.G. Hutchinson (1979) *Goodness of Fit Statistics for Trip Distribution Models*, Waterloo: Department of Civil Engineering, University of Waterloo.

Tischer, V. (2017) “Homogeneous Zones for Urban Mobility Planning: Case Study of Balneário Camboriú, Brazil, *Management Research and Practice*, 9:3, 5-12.

Wang, J.A. and E.J. Miller (2014) “A Prism- and Gap-based Approach to Shopping Destination Choice”, special issue of *Environment & Planning B*, 41, 977-1005.

Wang, S., L. Sun, J. Rong and Z. Yang (2014) “Transit Traffic Analysis Zone Delineating Method Based on Thiessen Polygon”, *Sustainability*, 6, 1821-1832, doi:10.3390/su6041821.

Yang, Y., W. Wang, H. Ding and K. Jin, (2019) “ Quantitative Determination Method for Traffic Analysis Zone Generation”, *Proceedongs of the Sixth International Conference on Transportation Engineering*, Chengsu, China, September 20-22.

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,
DOI: 10.1080/03081069708717601

Zhao, H. and Y. Zhao (2017) “Traffic Analysis Zones – How Do We Move Forward? Presented at the AASHTO, CTPP and TRB Joint Conference, Applying Census Data for Transportation”, Kansas City, Missouri, November.
http://onlinepubs.trb.org/onlinepubs/conferences/2017/censusdata/TAZ_Paper.pdf