"Developing an “intelligent” high-fidelity GIS-based travel demand model framework for improved network-wide traffic estimation"

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Canadian Journal of Civil Engineering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>cjce-2019-0085.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>05-Jul-2019</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Mustafa, Riad; University of New Brunswick Fredericton, Civil Engineering  
Christie, James S.; University of New Brunswick Saint John  
Hildebrand, Eric; University of New Brunswick Fredericton |
| Keyword: | four-step model, traffic analysis zones, high-fidelity travel demand model framework, automatic zone design, remotely-sensed images |
| Is the invited manuscript for consideration in a Special Issue? : | Not applicable (regular submission) |
Developing an “intelligent” high-fidelity GIS-based travel demand model framework for improved network-wide traffic estimation

Riad Mustafa, James Christie, and Eric Hildebrand

Abstract: The four-step travel demand model (FSTDM) is based on coarse rural traffic analysis zones (TAZs) which tends to exaggerate the intrazonal trips resulting in biased and unbalanced trip distribution over the roadway network with high estimation errors. These limitations have necessitated developing a Geographic Information Systems (GIS)-based high-fidelity travel demand model framework (HFTDMF) capable of achieving network-wide traffic volume estimation with improved model accuracy. This requires using an all functional class roadway network and enhancing the census-based coarse TAZ structure with finer-grained spatial resolution TAZs by integrating the travel demand modeling software platform, remotely-sensed images, parcel-based digital property maps, the AZTool aggregation algorithm, and areal interpolation technique. Preliminary results from the Greater Fredericton Area (GFA) showed that increasing the GFA spatial resolution from the coarsest TAZ structure at Census Tract (CT) level (27 CT TAZs) to the finest TAZ structure at 4252 “fine” TAZs resulted in an improvement to modeling accuracy of $R^2$, by 0.4092 (from 0.2490 to 0.6582) and an improvement in traffic assignment coverage by 46 percentage points (from 29% to 75%).

Key words: four-step model, traffic analysis zones, high-fidelity travel demand model framework, travel demand modeling, geographic information systems, areal interpolation technique, remotely-sensed images, automatic zone design.

Introduction

Network-wide traffic estimates with accurate replication of ground truth counts are of great importance to transportation planners, traffic engineers, and policy makers as they base their transportation decisions such as...
planning, design, maintenance, and asset management programs on traffic volume counts from individual highway
links. The traditional traffic monitoring system is diverse, lacks standards, is expensive, and can only cover a limited
number of point locations over a network. Estimating Network-wide traffic volumes with a non-survey approach
using travel demand models is an effective alternative to traditional traffic monitoring approaches for transportation
planners. Traditionally, the four-step travel demand model (FSTDM) is based on traffic analysis zones (TAZs),
which conveniently uses existing census blocks, block groups, or census tracts to take advantage of socioeconomic
and demographic data readily available from Statistics Canada. The traditional aggregate zonal-based model divides
the study area into coarse TAZs, which are assumed to be spatially homogeneous with all attributes uniformly
distributed throughout. Also using Statistics Canada census boundaries can result in some TAZs that are more rural
and tend to be larger with irregularities in shape and relatively low land use density, which is mainly residential in
type and scattered in distribution. This coarse zone structure tends to exaggerate the intrazonal trips resulting in
biased and unbalanced trip distribution over the roadway network and high estimation errors. In the design of the
zone structure, the objective is to maximize the interzonal trips while minimizing the intrazonal ones. Due to the
coarse zone structure used, the FSTDM generally shows inaccuracies in estimating traffic volumes at the roadway
level, especially for lower functional class road groups. Lower functional class roads (LFCRs) are ignored in most
cases in traditional modeling exercises. In fact, according to Statistics Canada 2006 Census Road Network and
Geographic Attribute File about 91% of Canada’s road network is classified as non-highway which refers to urban
lower-class roads and local rural roads (Statistics Canada 2006). In the light of these facts, traditional travel demand
models failed to provide for improved network-wide traffic volume estimation and, therefore, is not an ideal policy
tool for contemporary transportation agencies.

Limitations of the aggregated coarse zonal structure have led to the use of Geographic Information Systems (GIS)
in order to disaggregate the coarse zones into smaller spatial units such as polygons or grid cells, each carrying its
own unique features. GIS could be used to aggregate fine spatial geographic units at parcel level to generate study
area fine grained TAZ structures. This step of improving spatial resolution of coarse-grained TAZs is in essence
transforming the highly aggregated nature of the traditional modeling framework into a high-fidelity one. This type
of modeling framework is a zone design problem and transferring data from one spatial level of geography to
another by following the process of aggregation or disaggregation based on modeler-specified constraints to better
fit model scale and goals. Estimating traffic volumes at a finer spatial resolution requires high-fidelity travel demand
model framework (HFTDMF) input data (socioeconomic and land use) with similar spatial resolution. Spatial resolutions that can be achieved are higher than ever because of the availability of high-resolution remote sensing data and parcel-based tax assessment records. Advanced socio-economic assignment areal interpolation methods can be incorporated based on auxiliary factors such as roadway density (total length or number of roadway segments), and household and employment distribution discovered from high-resolution remote sensing data and parcel-based tax assessment records. Utilizing remote sensing and GIS to disaggregate census data to a fine-grained zone level to reflect more realistic land use distribution is expected to improve the traditional four-step model estimation accuracy.

This research paper briefly reviews the literature on developing network-wide traffic volume estimation using the traditional four-step model, discusses critical issues related to the four-step model such as Traffic Analysis Zone (TAZ) design and selection, and utilizing advanced technology and “intelligent” data sources (e.g., GIS, remote sensing, and parcel-based tax assessment data) for high-fidelity travel demand modeling purposes. This paper then presents the proposed methodology to be applied in developing the HFTDMF. Then, the proposed methodology is applied to a case study of the Greater Fredericton Area (GFA) in the province of New Brunswick. Finally, conclusions and recommendations from this research paper are presented.

**Review of critical issues in modeling improved network-wide traffic estimation**

Examining both the state-of-the-practice and the state-of-the-art of using FSTDM as a tool for estimating network-wide traffic volumes is achieved by considering TAZ structure and roadway network details. Although the interaction between the zoning structure and the detail of the network has been of concern to the transportation community, it appears that it has been largely overlooked in the literature and in the practice. Since this interaction should not be ignored, the general recommendation is to use a more detailed roadway network with a smaller fine-grained TAZ structure and a less detailed one for larger coarse-grained TAZ for network-wide traffic estimation (Khatib et al. 2001). The study investigated the effect of TAZ structure on travel demand forecasts and concluded that the use of fairly large TAZs produced results with fairly high estimation errors reported, mainly related to the unbalanced trip distribution and limited traffic assignment over the highway network on one hand while smaller
TAZs resulted in higher model estimation accuracy on the other hand. Another study by Cambridge Systematic -Inc (2007) mentioned that large TAZs with both production and attractions (dwelling and employment) ignore the assignment of a large number of intrazonal trips to the network. Both studies recommended reducing the number of intrazonal trips by decreasing the size of the TAZs.

This literature review investigated a number of traditional four-step models developed at varying levels of geographies (small size city, county, regional, and Metropolitan Planning Organization-MPO) in the U.S. All developed models failed to achieve a network-wide traffic volume estimate because they ignored local roads or their traffic estimates were fairly inaccurate for lower-level roads (minor arterial and collector) (Horowitz 2006; Ullah et al. 2011; Martin and McGuckin 1998; VHB-Inc 2007). More to the point, TRB Special Report 288 (2007) mentioned that highway networks coded by MPOs usually include all or almost all freeway, expressway, major arterial, minor arterial and feeder/collector routes but not local roads. The document reported that on average, less than 25% of local class road mileage is represented. Local streets are rarely coded into the model network and they have been simulated by the centroid connectors (Ullah et al. 2011). In this regard, the network developed under “the-state-of-the-practice” guidelines recommended coding links to capture or represent at least 85% of the vehicle trips (ODOT 1994). The guideline document claimed that the level of network detail which would be able to capture 90% or more of the vehicle trips should include collector roads. From the point of view of this research, capturing 90% of the traffic volume generated should necessitate the inclusion of local streets, especially in urban zones where the majority of the network consists of local streets and a fairly large amount of traffic is local. It is clear from the literature review that the traditional travel demand model (TDM) was based on a coarse TAZ structure and less detailed road networks, which mostly ignore lower functional class roads and, generally speaking, traditional TDMs failed to achieve improved network-wide traffic estimation.

Traffic demand modeling exercises typically start with defining a study area which is subsequently subdivided into basic geographic units called traffic analysis zones (TAZs) used for data collection, tabulation, and analysis. TAZ selection and configuration plays a major role in reducing the errors in relation to trip generation and traffic assignment. Martínez et al. (2010) stated, “Defining a good set of TAZ is still one of the transportation unsolved problems”. This arises because the transportation community has not reached a consensus about TAZ configuration and design. To date, the TAZ design process is purely a planner and decision-maker exercise related to different criteria being used and knowledge of the study area being modeled. Sometimes, the TAZ selection is as naive as
being based simply on existing coarse census geographies or even mere perception of the situation. On the other hand, scholars have identified census as an “inherently impoverished source of data, which is limited by coarse and arbitrary zonal boundaries” (Schuurman et al. 2006). As the census geographies do not fit modeling purposes in most cases, they should be reconstructed under a set of modeling criteria in order to achieve a set of modeling goals. For this reason, the rest of this section will closely look at the remaining two TAZ design approaches: “rules of thumb” guidelines and rule-based analytical techniques.

Literature on TAZ design for travel demand modeling purpose is scarce and sparse. Mainstream transportation modeling literature is largely silent on TAZ design issues and procedures (Garber and Hoel 1988). TAZ selection and design are addressed in many transportation planning publications and textbooks as a set of criteria or “rules of thumb” without a formal method for guiding the design of TAZs. TAZ design which is based on fulfilling a set of multiple criteria (or optimizing a set of constraints) has proven to be a challenging task due to the complexities involved. Modelers trying to achieve zonal homogeneity will most likely conflict with meeting the zonal contiguity requirement. A TAZ with homogeneous land use activity may conflict with its shape compactness. The whole TAZ design exercise becomes a question of dealing with trade-offs between conflicting constraints to reach a feasible solution to fulfill the modeling needs. Due to the complexity of the problem, several heuristic approaches have been tried to solve the TAZ configuration problems as well as more advanced methods attempting to rationalize the design process in an analytical framework using spatial and statistical techniques. TAZ design analytical techniques have been motivated by advances in GIS, computer speed and storage capabilities, and algorithm software application. A number of clustering algorithms were deployed in the GIS environment using a different set of constraint and objective functions to achieve “optimal” TAZ design. All attempts were faced with problems associated with the “modifiable” or arbitrary nature of the spatial zoning system (O’Neill 1991; Choi and Kim 1995; You et al., 1997; Ding 1998). Hagen-Zanker and Jin, (2013) have introduced a method for modeling spatial interaction. The adaptive zoning establishes a compendium of different zone plans, each of which is applicable to one journey origin or destination only. For every origin zone, destination zones are aggregated together based on their distance from the origin. Hence, a separate map is created for each origin, with nearby destination zones being small and more distant ones larger. They tested the method on a commuting model in England and the findings suggest that adaptive zoning has a significant potential in enhancing the accuracy of mode choice modelling at the city or city region scale, especially where walking and cycling are considered important components of the transport model.
system. Adaptive zoning ignored the well documented criteria required for developing zones for traditional travel demand modeling applications. Molloy & Moeckel, (2017) developed a raster cells-based aggregation algorithm towards zone system design for transportation and spatial modeling applications. The generated zoning system respected the municipal jurisdictional boundaries to facilitate the disaggregation of socioeconomic data. The developed automated zone design algorithm overlooked the treatment of rural zones spatial resolution along with zonal structure “intensity” (i.e., spatial resolution and number count). Also, developed zone systems did not fulfill the zone design criteria required for travel demand modeling purposes such as homogeneity, compactness, sliver zones, spatial contiguity, respect physical barriers and major roads, etc. Finally, the developed zone system is to be tested for improvement in modeling accuracy. To date, the issue of TAZ design for travel demand modeling application is still an open area for further research.

Traditional models failed to provide improved/network-wide traffic estimations. This was mainly due to their dependence on highly aggregated census data reported over coarse spatial geography. Nowadays, the remedy to problems of spatial geography scale and census data aggregation is possible with recent advances in GIS and Remote Sensing (RS), along with tremendous improvements in computing speed, internet accessibility, graphic quality, and data availability. The introduction of Geographic Information Systems (GIS) into the traditional modeling practice has been acknowledged as being the most significant enhancement to advanced practice in the field. Martin and McGuckin (1998) recognized GIS as a significant development in the travel modeling field. The existing traditional travel demand modeling efforts have integrated GIS technology to the point that the use of a GIS has become mandatory (Bacao and Painho 2002) and travel demand model software packages capabilities have been significantly improved with GIS integration. In this regard, one could note the following aspects: (1) input data processing, analysis on the fly, and visualization are performed easily within a GIS environment, (2) output results can be monitored and visualized in a variety of formats, (3) direct interaction between the various model components facilitates monitoring the model at every step, (4) linking between model components by feedback and conducting sensitivity analysis is possible (McCormack and Nyerges 1997).

The traditional four-step travel demand model (FSM) is based on employing census ready sociodemographic and employment data which is problematic due to the data aggregated nature for privacy reasons as well as the data tabulation over fairly coarse census geographies. Original Census data and geographies were prepared, by Statistics Canada, for purposes other than travel demand modeling on one hand and unrealistically assumed to be spatially
homogeneous with all attributes being aggregate averages and uniformly distributed throughout on the other hand. Spatial misrepresentation from demographic and economic census data limited the accuracies of traditional models and their network-wide traffic estimation capabilities. These limitations have led transportation agencies and communities to continuously seek new modeling tools and data sources to improve the static and rigid nature of census data. Remote sensing technology is emerging as an “intelligent” data source, which has the potential to resolve some of the aforementioned issues from census data. The potential contributions from remote-sensing technology to the transportation planning and modeling process could be motivated by the following benefits of using remote-sensing for data collection in transportation planning (Ekern 2001): (1) flexibility in collecting data from places that are costly or impossible to reach, (2) it is a non-intrusive technique, and (3) wide-area coverage at improved spatial and spectral accuracies without intensive labor activities. Ekern (2001) in his efforts to close the knowledge gap between transportation and remote-sensing technology communities suggested that the two communities should become knowledgeable with respect to mutual needs and products. High-resolution images (better than 5 m spatial resolution) contain rich, detailed, accurate spatial information, providing, in a timely and a relatively less-costly manner, land use/land cover (classified land use, impervious surfaces, and vegetation) and urban features (roads and buildings). In this regard, Usher (2000) urged transportation organizations to make changes to their current approaches to transportation planning and operations by employing new tools for storing, retrieving, visually representing, manipulating and analyzing data. Remote sensing is used to improve building and roadway network extraction. On one hand, buildings accommodate human beings’ activities, and they contain vital information such as size, height, geometric shape and orientation. On the other hand, all of them are useful for land use classification and high-fidelity travel demand modeling application. Accurate extraction of roadway network centerlines/boundaries is considered to be an opportunity for validating, updating and conducting various accuracy checks. Remote sensing activities provide useful quantitative information about socioeconomic attributes. They provide travel demand models with real-time or near real-time socioeconomic data as well as population/residential densities in each TAZ, which reduces reliance on census data on one hand and helps in validating census data on the other hand. In this regard, Jensen and Cowen (1999) stated that socioeconomic attributes can be extracted directly from remotely sensed data or indirectly by surrogate information derived from the imagery.
Methodology

The traditional trip-based four-step macroscopic modeling approach ignores lower functional-class roads in most cases and uses coarse rural traffic analysis zone (TAZ) structure coded with highly aggregated sociodemographic and employment (SD & E) data, all of which has failed to provide network-wide traffic estimates with improved accuracy. The main objective of this research was to develop a methodology aiming to enhance the traditional travel demand model in relation to performance which is hindered by the model’s macroscopic planning perspective due to using a coarse rural traffic analysis zonal (TAZ) structure along with highly aggregated sociodemographic and employment (SD & E) data, and ignoring lower functional class roads (locals and collectors) in most cases. The research efforts were directed to full enhancement of the traditional “gross” highly aggregated modeling scope into a high-fidelity “fine” disaggregated modeling framework. This enhancement was achieved by introducing three main measures into the traditional model: 1- Improving the spatial resolution of coarse TAZs into smaller spatial units, each carrying its own unique features; 2- Using high-resolution (disaggregated) modeling input data; 3- Using a detailed all functional class digital roadway network to include local streets and collector roads in order to represent realistic traffic patterns. The promise of the proposed HFTDMF is based on the research hypothesis which stated that improving modeled area TAZ structure spatial resolution (or precision) and modeling an all functional class roadway network will improve the network-wide assignment in terms of both modeling accuracy and coverage.

In light of the above-mentioned modeling issues and facts, this research was initiated and highly motivated by the recent technological advancements and the availability of intelligent data sources, which could be utilized to overcome the aforementioned limitations and shortcomings of the traditional travel demand models by improving spatial resolution of geographic space and allowing for spatial levels of analysis finer than ever before. The effect has been the development of a GIS-based HFTDMF for improved network-wide traffic volume estimation which was based on an integrated methodological framework (Fig. 1). The methodological GIS-based zonal enhancement framework is achieved by integrating the travel demand modeling software platform, remotely-sensed images, parcel-based tax assessment database, automated zone design (AZD) algorithm, and areal interpolation technique. This unique “intelligent” integration of various tools has proven to have tremendous GIS-based functionalities and capabilities in treating spatial resolution of “coarse” rural TAZs as well as building a fully detailed roadway network to include lower functional classification roads (collectors and local streets) which are coded with link and node
attributes as needed for high-fidelity modeling purposes. The promise of this research is to evolve the gross and highly aggregated traditional travel demand model into a GIS-based High-Fidelity Travel Demand Model (HFTDM) by utilizing a zonal enhancement framework to estimate network-wide traffic volumes including lower classification roads with improved estimation accuracy. The proposed zonal enhancement framework is not designed to develop a fully “calibrated” High-Fidelity Travel Demand Model (HFTDM) due to the lack of local travel surveys, such as household survey, cordon counts, roadside interview counts, screenline counts, etc., to adequately calibrate each step of the four-step travel demand model. The zonal enhancement framework is designed to improve the spatial resolution and to reduce aggregation flaws of the traditional four-step travel demand model as it has been applied in both the-state-of-the-art and the real-world modeling practice.

The GIS-based intelligent data framework improves the zonal structure spatial resolution of coarse zones by aggregating fine spatial geographic units at the parcel level. The generation of parcel-based fine spatial resolution TAZ structure is guided by the impervious surface layer (building, roadways, etc.) extracted from high-resolution remotely-sensed images for the study area. This aggregation procedure helps to ensure that fine-grained spatial or attribute data are not lost or averaged through aggregation. The hypothesis is that fine-grained spatial resolution model input data (socioeconomic and land use) allows for development of high-fidelity travel demand models with improved efficiency, resolution, and accuracy for estimating network-wide traffic volume. In order to achieve this goal a GIS-based methodological framework was developed to fulfill the following research tasks:

- Delineating Census-based coarse TAZ structure at the Dissemination Area (DA) level to a Parcel-based fine-grained spatial resolution level using an aggregation (step-up) approach. Spatial improvement is primarily targeting large or coarse-grained spatial resolution rural DA zones which are mostly overlooked in TAZ design practices despite their impact on modeling accuracy.
- Developing an integrated GIS-based aggregation tool for generating a fine TAZ structure. The tool is a methodological framework which incorporates “intelligent” data sources from remote sensing imagery and parcel-based tax assessment data with an Automated Zone Design (AZD) platform which is based on an AZTool aggregation algorithm.
- Creating fine-grained spatial resolution TAZ structures for the modeled area using the integrated GIS-based aggregation tool. Each developed TAZ structure is guided by a specific aggregation rule.
Disaggregating sociodemographic and employment (SD&E) attributes using an Areal Interpolation Method (AIM) for spatial conversion between source zone (large rural DA) and target zones (generated fine zones). The SD&E attributes are disaggregated as a function of impervious area overlap between the target zones (fine TAZs) and the source zone (DA). The “intelligent” areal interpolation method for spatial data conversion is based on an auxiliary variable of impervious area ratio (IAR) which is extracted from remotely sensed images.

- Running the HFTDMF which is based on enhancing zonal spatial resolution and disaggregated sociodemographic and employment attributes at each developed TAZ structure.

- Performing statistical validation and model calibration to establish the research paper intent in relating the effect of improving the modeled area TAZ structure spatial resolution on the network-wide traffic assignment accuracy.

Each research task of the proposed methodology was used to develop a GIS-based HFTDMF using the Greater Fredericton Area (GFA) in the Canadian Province of New Brunswick as the case-study.

**Developing a GIS-based HFTDMF for the Greater Fredericton Area (GFA)**

The study area was chosen from both York and Sunbury Counties in the Canadian Province of New Brunswick which is divided into 15 counties. Each County is basically a Census Division (CD) as described by Statistics Canada (2006) geographic hierarchy. The study area was composed of Fredericton Census Agglomeration (CA 320) and the Town of Oromocto Census Subdivision (CSD 1303012) and was referred to for the purpose of this research as the Greater Fredericton Area (GFA) as shown in Fig. 2. The GFA study area was modeled with a geographic landscape which includes an Urban Core (UC), Rural Fringe (RF) and a nearby Metropolitan Area Influenced Zone (MIZ) to better represent a diverse range of population sociodemographic and employment activities and adequately capturing them in the modeling process (Fig. 3). Fredericton (CA 320) was the main component of the modeled area which contained the City of Fredericton (CYF), the rural Village of New Maryland (NMV), the rural 7 Perishes (P), and the 3 rural First Nations Reserves (FNR). The GFA population count totals 94,090 of which 85,688 lives within Fredericton CA and 8,402 lives in the Town of Oromocto CSD covering an area totaling 4,613.19 km² of which 4,590.82 km² is in the Fredericton CA and 22.37 km² is in the Town of Oromocto CSD.
Modelers should decide on the spatial/geographic level of the study area to be analyzed. The high-fidelity model is based on analyzing the census data at the smallest census units called Dissemination Areas (DAs). Statistics Canada (www.statcan.ca) defines dissemination area (DA) as a small relatively stable geographic unit composed of one or more adjacent dissemination blocks. It is the smallest standard geographic area for which all data are disseminated. In practice, the DA level of Census geography is often used as a building-block for TAZs. The study area was selected so that the 166 DAs covered an entire urban influencing area (UIA), including urban, suburban, and rural area. This ensured all relevant urban activities were considered and this led to a modeling area with three levels of spatial resolution and degree of heterogeneity. Urbanized areas (e.g. CBDs) tend to include zones which are large in number, small in areas, and dense in land use activities. In rural areas the zones tend to be larger in areas and lower in land use density.

After deciding on the study area Census-based TAZ structure at the DA level, the development of the GIS-based HFTDMF for improved network-wide traffic volume estimation was based on the following broad methodological tasks:

1- Collecting and tabulating of all modeling input data from available sources.

2- Attaining GIS-based “Parcel” Digital Property Files for Rural Coarse DA TAZs.

3- Extracting Impervious Surface Area (ISA) from Rural DA TAZs through Object-based Supervised Image Classification.

4- Designing Parcel-based TAZ Structures by Developing an Automated Zone Design (AZD) Methodological Framework.

5- Developing an Areal Interpolation Sociodemographic and Employment (SD&E) Data Disaggregation Framework.

6- Utilizing the “Parcel-based” Fine-Grained TAZ Structure and Disaggregated SD&E Data to Develop and Run HFTDMF.

Census statistics are the primary source for demographic and economic data at zonal level. They are essential for effective and efficient travel demand modeling. Statistics Canada (SCAN) provides aggregated socioeconomic data for each census unit (DA) within the study area. The socioeconomic profile of the population is a key input to the Travel Demand Model (TDM) because it determines the trip-making characteristics of the population. Demographic data includes total population, the number of households, the number of dwelling units, income per
household, and employment by sector (retail, non-retail, etc.). Partial presentation of SD & E data for the base case GFA study area is shown in Table 1. Another vital element in developing the travel demand model is acquiring an accurate up-to-date roadway network (RN) serving the study area. A functionally classified roadway network serves several purposes in developing the HFTDMF because it includes attributes required for each link such as functional classification, area type, link type, link length, link capacity, link speed, lane width, number of lanes, and traffic volume counts. The major role of the roadway network is to estimate link travel impedance between zones in the study area as well as to perform the traffic assignment process in which zonal trips are loaded and distributed onto the network.

The task of developing the AZD framework for the case-study aimed at enhancing the spatial resolution of 52 rural DAs from the 7 Parish CSDs and the Village of New Maryland CSD located in Fredericton (CMA 320) as detailed in Table 2. The rest of the study area which contains 114 DAs in the City of Fredericton, the Town of Oromocto and the 4 First Nations Reserves (FNRs) were left without any further enhancement due to their high spatial resolution with an average area of 1.46 km² per DA.

The AZD methodology started with defining the 2006 base-case Census geography layer for the 52 rural DAs which are highly stable and coded with needed modeling sociodemographic and employment input data. The methodology calls for building a block layer of a basic “fine” spatial units which are spatially related and their boundaries aligned with the Census-based standard zoning system. The most feasible building block layer for aggregation to higher-level output zones is the Parcel-based Digital Property Map (DPM) Database. Because rural DA TAZs are sparse, many parcels are covering open green space and vegetation which are not related to human sociodemographic and employment activities; this is particularly problematic for TAZ design in relation to developing HFTDMF. For the purpose of capturing only the human foot prints within the sparse large rural zones, supervised classification was performed on 54 Softcopy Orthophotomap Data Base (High-Resolution Images) to create an impervious surface area (ISA) vector layer (Shapefile) for the 8 rural CSDs (Fig. 4). Each Orthophoto image tile is identified with a keyword (8-digit unique ID number) at a scale of 1:10 000, which translates to 0.1 degree (7.5 km) in longitude by 0.05 degree (5.5 km) in latitude. The digital color orthophotos of New Brunswick (digital aerial photos) with 1-meter pixel resolution and geo-referencing information with a Datum -NAD83 (CSRS) and a Map Projection - NB Stereographic Double. The ISA layer was overlaid and spatially joined with DPM parcels to guide the classification process of the parcels into: “developed” or “utilized” and “undeveloped” or “un-
utilized”. “Developed” parcels are those with ISA which is a proxy for human activities, and therefore were used in the step-up aggregation process to generate output “fine” TAZs.

Imperviousness can be incorporated as a new auxiliary factor in zone design and data interpolation (disaggregation and/or aggregation) and this requires relating ISA to the digital property map parcels for the purpose of developing a layer which includes only “developed” parcels. The idea is based on transferring the extracted impervious area from the CSD level to the parcels which are contained within the CSD. Then, one can tabulate the amount of imperviousness (area) within each parcel and classify them as “developed” or “undeveloped” based on a well identified statistical threshold value. Areal imperviousness is expressed as Impervious Surface Area Ratio (ISAR) or Index (ISAI) which is equal to ISA/Geographic Area. ISAR was expected to be low across many of the rural parcels because most farmland parcels had only a limited portion developed for residential or livestock purposes and the rest was assigned to general farming purposes; even ISAR could be close to none in many open green fields. Despite the fact that the Imperious Surface Area Ratio (ISAR), which is a proxy for human travel activities is expected to be low in rural DAs with parcels, the ISA (buildings and roads) is still a very useful indicator of human SD&E activity which could translate into demand for travel.

The impervious area threshold was determined through statistical analysis related to ISAR across all parcels in the CSD and served as the threshold to differentiate parcels as “developed” (utilized) or “undeveloped” (unutilized). The ISAR/ISAI threshold analysis included parcels with impervious surfaces related to both building and transportation features and ensures parcel continuity due to the selection of parcels that were classified as “developed” or “utilized”. The ISA threshold verification method was conducted to establish a good percentile rank with ISAR value to "reasonably" classify parcels into either "ISA" or "non-ISA" parcels which ensures continuity of “ISA” parcels and truly matches ground truth levels of development and human activities. Examination of the entire parcels overlaid in the rural CSDs of the study area revealed that ISAR at the 2nd percentile with a value of 0.0048 as shown in Fig. 5 reasonably reflected and replicated the ground truth developed parcels. The selected (developed) parcels at this threshold could be identified as impervious space occupied by sociodemographic and employment activity and seem to be clustered around impervious features with a reasonable continuity and can be sub-zoned into clusters within the DAs to generate “fine” zones by employing an Automated Zone Design (AZD).

The AZD process was applied at the DA level for each one of the 8 rural CSD to generate output zones based on “developed” or “utilized” parcels that fulfil pre-specified attributes including the target threshold variable of ISAR.
(or ISAI) which is defined for each parcel. The AZD framework was based on ArcGIS/AZTool (AZT) software integration for the purpose of automatic generation of geographical areas in an iterative way resulting in 8 GFA TAZ structures for the purpose of running the HFTDMF. The AZTool algorithm, an open source Windows application, was combined with ArcGIS to facilitate the creation of new geographies by applying a number of formatting steps

(Martin and Cockings 2011). The latest version of AZTool is available at: [http://www.geodata.soton.ac.uk/software/AZTool/](http://www.geodata.soton.ac.uk/software/AZTool/).

To run the model at a larger scope, a highly coarse spatial resolution TAZ structure at the Census Tract (CT) level was added, which satisfies a State-wide or Provincial-wide modeling scope. 10 TAZ structures (CT, DA, and 8 from AZT Runs) were developed, which allowed the study area to be modeled at three modeling scopes or perspectives (Table 3):

1- Regional modeling (State-wide or Provincial-wide) using the CT-based TAZ structure at the macroscopic level.

The TAZ structure was developed based on the ready-made Census geographic entities with some GIS processing.

2- Metropolitan Area modeling (MPOs or Greater Metro Area) using the DA-based TAZ structure at the mesoscopic level. The TAZ structure was developed based on the ready-made Census geographic entities with some GIS processing.

3- Urban Area Focused Modeling (small County or City) using DPM Parcel-based TAZ structure at the microscopic level. The TAZ structure was developed based on the introduced AZD methodological framework for the purpose of this research.

An attribute table with features stored at one geographic spatial level were related to another geospatial level by a relational model (for attributes) such as a “one-to-one” or “one-to-many” relationships. Further, the attribute table features were transferred from one geospatial level to varying levels using the same relational models and a common distribution factor. This data spatial analytical operation is called an Areal Interpolation Method (AIM) and is guided by a source zone, target zones and an auxiliary spatial distribution factor (SDF) for splitting attributes from the source zone to the target zones. The challenge of socioeconomic and land use data transfer between varying spatial geographies (source to target) in travel demand model spatial analysis is equally challenging as TAZ design. At this stage of the case-study, after preparing TAZ structures along with ISAR for parcel-based TAZs,
sociodemographic and employment data were transferred from DA and CT Census-based geographies (source zones) to the AZD developed TAZs (target zones) by applying the Areal Interpolation Method (AIM) for Data Transfer. A framework for the Areal Interpolation of Sociodemographic and Employment (SD&E) Data for the GFA case-study is outlined in Table 4.

GFA roadway network (RN) contains 4619 nodes and 5,787 classified routes totaling 2,060.91 km in centerline length. The selected network model for the study area was a detailed one for all functional classes which included freeways and arterials as well as collector and local roads to serve the purpose of developing the high-fidelity travel demand model framework. Lower functional class roads (collectors and locals) comprise a majority of the network, 78% and 71% of the total roadway links and centerline length, respectively. For the purpose of this research, the developed TDM for all 10 generated TAZ structures was based on the traditional 4-step one, which includes trip generation-production/attraction (TG-P/A), trip distribution (TD), and traffic assignment (TA). The modal split (MS) step was omitted because very limited transit service is available in the study area and most of the trips occur by private vehicle. According to Statistics Canada (www.statcan.gc.ca) the proportion of Canadians using their cars to travel to work was 72.3% in year 2006 while those using public transit was only 11.0%. The percentage of private vehicle mode is expected to be much higher than these national averages in areas covered by very limited transit services. The theoretical aspects of the travel demand model sequential three steps were performed using TransCAD for which details can be found in the TransCAD reference manual (TransCAD 4.8 2005).

The first step after TDM development and prior to its implementation is to ensure that the traffic estimates over a roadway network are properly validated. Link-based (network-wide) traffic assignment validation compares predicted and observed volumes via statistical tools and graphical tools. The statistical measure of Percent Assignment Error (PAE) indicates the accuracy of the model in replicating the actual observed traffic counts. PAE is more widely used than absolute numerical difference between estimated and observed values because it better reflects the traffic volume of the roadway and helps to identify any forecasting biases. PAE = (Assigned−Observed) * 100 / Observed. The Coefficient of Determination, R-Squared (R²), as a “goodness of fit”, measures the model’s overall accuracy. “Real” R² is defined as follows:

\[
R^2 = 1 - \left[ \frac{\text{SUM}(Y_o - Y_e)^2}{\text{SUM}(Y_o - Y_m)^2} \right]
\]
Where:

\[ Y_o = \text{Observed AADT.} \]
\[ Y_e = \text{Estimated AADT.} \]
\[ Y_m = \text{Mean Observed AADT.} \]

\( R^2 \) represents and measures the proportion or percentage of the total variation in the observed link traffic counts that is explained by the estimated assigned traffic volume that are generated by the model. In effect, \( R^2 \) measures the strength of the relationship between the assigned volumes and traffic counts for all validated links and shows the proportion of total variation around the mean explained by the linear regression model. The remaining variation is related to random or unexplained error. To establish the degree of assignability or the assignment capability attribute of the assignment model for each of the model runs, Percentage Assignment Coverage (PAC) was applied to measure the model performance at the network-wide level (PAC = No. of Links Assigned Traffic/Total No. of RN Links ×100%).

The most popular graphical validation tool is a scatterplot of observed counts (on Y-axis) for all validated links (links with traffic counts) versus assigned volume (on X-axis) which, when combined with \( R^2 \), provides a quick visual examination of the correlation between data sets including outliers. Identification of outliers is precisely facilitated by including the 45º line from the origin. Hypothetically, any data set (or validated point) laying on the 45-degree line represents a perfect match between assigned volume and observed traffic count. Outliers are links with either high over-assignments or under-assignments compared to their observed counts. For the purpose of this research, statistical validation was conducted after identifying and eliminating suspicious outliers. Because there was not an easy way to simply detect outliers from the statistical data analysis, there was a need to rationalize outlier’s identification and elimination to be applied for each model run at various TAZ structures. The treatment of outliers was based on the following elimination rules:

- **Rule 1:** Eliminate all erroneous validated links with Absolute Percent Assignment Error, APAE > 100. It was found that they were mostly related to lower class functional roads (locals & collectors) and they were over-assigned (assigned volume > observed counts).
- **Rule 2:** Eliminate all extensively under-assigned links (assigned volume/observed counts ≤10%). Basically, eliminating all points on or very close to the Y-axis.
- **Rule 3:** Eliminate all extensively over-assigned links (observed counts/assigned volume ≤10%). Basically,
Statistical validation was based on defining 2 rationales of outliers’ elimination: 1- Rationale 1 which applied elimination rule 1 only. 2- Rationale 2 which applied elimination rules 1, 2, and 3. Both rationale elimination procedures were conducted at the network-wide level for all validated links. Only $R^2$ from rational 2 was documented, because it resulted in the improved $R^2$ value.

Statistical validation of the developed base-year model showed a poor predictive capability which required performing a calibration process. For the purpose of this research and due to the limited resources of the tremendous wealth of data needed to calibrate each step of the 3-step model, calibration focuses on “fine-tuning” the traffic assignment model by modifying BPR (Bureau of Public Roads) link performance function parameters of $\alpha$ and $\beta$ as well as the Relative Gap (RG) value, roadway network links attributes such as link capacity, and relocation of zonal centroids and centroid connectors. Modelers should target a scatterplot of Observed Counts (as Y-variable) versus Assigned Volume (as X-variable) and a linear regression line which results in a theoretical slope very close to 1 (not statistically different from 1) and an intercept very close to 0 (not statistically different from 0). Due to the difficulties in developing a regression model from the calibrated data with a slope and intercept which are not significantly different from 1 and 0, respectively, the recommendation was to use “forced” regression through the origin as the primary measure of goodness of fit and focus on getting a slope not statistically/significantly different from 1. The “forced zero” intercept regression model was developed from the validated links under the 95% and 99% confidence limits. Detailed “Real” $R^2$ values from base-case and calibrated model runs for the 10 TAZ structure levels are included in Table 5. Finally, the calibrated model “Real” $R^2$ % value is plotted against number of TAZs related to each model run as shown in Fig. 6.

The performance of the developed base-case models at 10 TAZ structure levels in relation to their assignment capabilities indicators such as Percentage Assignment Capability (PAC), Centerline Length Assignment Coverage, and Link “Average” Volume Assignment is presented in Table 6.

**Analysis of GFA 10 TAZ structures modeling results**

Results of link-based traffic assignments were presented and analyzed to show the incremental improvement in the model’s network-wide assignment performance with increasing spatial resolution. This comes in recognition of the
benefits that resulted from the proposed integration of technological advancements and intelligent data sources within the GIS environment for the purpose of enhancing the Traditional TDM and developing the HFTDMF. A detailed analysis of the modeling results comes in 2 folds:

1- Network-wide assignment capability (model assignability).
2- Network-wide assignment predictive capability (model accuracy).

**Analysis of the model network-wide assignment capability**

Table 6 indicates that improving the study area spatial resolution from the coarsest level at CT TAZ structure to the finest level at AZT Run1 TAZ structure improved the model assignment coverage by 46 percentage points (from 29% to 75%). Most of the model assignment capability improvements (36%) were gained by improving study area spatial resolution from CT level at 27 “coarse” TAZs to Optimal level at 626 “fine” TAZs. Clearly, the marginal diminishing returns on the spatial resolution improvements has been achieved and the assignment capability gains slowed down as we reached the most disaggregated TAZ structure at AZT Run1 with 4252 “fine” TAZs (from 65% to 75%). Modeling the all functional class roadway network along with fine-grained zone level increased the assignment coverage over the LFCRs (locals and collectors) as well as over HFCRs (Freeways, Highways, Ramps, and Arterials). Most of the assignment coverage gain (51%) was observed on LFCR (from 19% at CT level to 70% at AZT Run1 level). Interestingly, improvement on the lower functional class roads did not result in any loss on the high-class roads improvements which was at a slower rate of 27 percentage points (from 64% at CT level to 91% at AZT Run1 level). Clearly, running the HFTDM with finer spatial resolution TAZs ensured increasing the assignments to lower functional class local roads.

Base-case RN developed for the study area, totaling 2060.69 km, was covered with assignment volume depending on the TAZ structure spatial resolution (Table 6). The network-wide length assignment coverage improved by 45 percentage points: from 26% at CT “coarse” TAZ level to 71% at AZT Run1 “fine” TAZ level). Improving TAZ structure resolution increased the roadway network centerline length covered with assigned traffic volume over both LFCRs and HFCRs. Most of the assignment coverage gain (50 percentage points) was observed on LFCR (from 15% at CT level to 65% at AZT Run1 level). Improvement on the lower functional class roads did not result in any loss on the higher functional class roads improvements which was at a slower rate of 33 percentage points (from 52% at CT level to 85% at AZT Run1 level). Most of the length assignment coverage gain was achieved at optimal...
TAZ level with 626 “fine” TAZs, beyond which the diminishing return from improving spatial resolution was obvious. A constant of about 84% assignment coverage was observed on higher functional class roads from the optimal TAZ level study and stayed constant for any subsequent TAZ structure level scenario. Also, an approximate constant of about 61% assignment coverage was observed on lower functional class roads after the optimal TAZ level and stayed reasonably constant for any subsequent TAZ structure level scenario. The improvement of assignment coverage at both network-wide and functional-class levels due to applying HFTDM at fine-grained spatial resolution TAZ structures resulted in effective and balanced distribution of assigned traffic volume over the roadway network links. This can be observed from the average assigned volume (AAV) per link which has been reduced with improved spatial resolution for all functional class Roadway Network (RN), HFCRs, and LFCRs. Table 6 shows that improving TAZ structure resolution from CT “coarse” TAZ level to AZT Run1 “fine” TAZ level reduced the average assigned traffic volume by 76%, 59%, and 81% for all functional class RN, HFCRs, and LFCRs, respectively. Most of the average assigned traffic volume reduction was observed on LFCR. The maximum reduction gain on average assigned traffic volume was achieved at optimal TAZ level with 626 “fine” TAZs, beyond which the diminishing return from improving spatial resolution was obvious.

Analysis of the model network-wide predictive capability

The primary goal of this research paper was to prove that increasing zonal structure spatial resolution from coarse to fine will improve the model assignment accuracy (or increasing “Real” R² value). For this purpose, the performance of both developed base-case and calibrated models at 10 TAZ structure levels in relation to their assignment predictive accuracies were compared by investigating the network-wide R² value (Table 5). The model developed calibration framework was followed closely to provide a “robust” base for model “real” R² comparison (Fig. 5). The following could be concluded from Table 5 and Fig. 5:

- Improving the model TAZ structure from a highly coarse-grained at CT level to a medium-grained at DA level resulted in a noticeable improvement in “real” R² from 0.2490 to 0.6375.
- Increasing the model TAZ structure spatial resolution from medium-grained DA level to a high-fidelity fine-grained level further improved model “real” R² but at a slower gain rate.
Increasing the model TAZ structure spatial resolution gradation after the DA level generally speaking improved the “real” $R^2$ till Optimal TAZ Run level then at a slower gain rate for the following fine TAZ structure levels. 

Doubling the number of the model “fine” TAZs from Run 2 (2122 TAZs) to Run 1 (4252 TAZs) did not further improve the model “real” $R^2$ and this could be explained by the diminishing marginal returns of the model zonal disaggregation on the “real” $R^2$ gain.

After the Optimal TAZ run the "real" $R^2$ plateaued because the data that were generated from further refinement of the TAZ structure (Runs 4 to 7) did not generate significantly more traffic that could be assigned to the road and street network. The majority of the generated traffic in the study area was captured and assigned to the road and street network at the Optimal TAZ run.

**High-Fidelity travel demand modeling: discussion and recommendations**

**Discussion**

To date, both the transportation community in general and the travel demand forecasting experts in particular have not addressed the fundamental question regarding the required methodological framework to achieve the traditional model enhancement for the purpose of improving the performance. The concern has always been that the highly disaggregated traditional model is likely to shift the traditional modeling framework to a new paradigm which would require changing the basic theoretical foundation of the 4-step model on one hand and to be faced with the challenge of employing the scarce and difficult to attain high-resolution sociodemographic and employment data on the other hand. This research paper addressed the need for improved network-wide traffic volume estimates through full enhancement of the traditional “gross” highly aggregated modeling scope into a high-fidelity “fine” disaggregated modeling framework. This enhancement was characterized by introducing three main measures into the traditional model: 1- Improving the spatial resolution of coarse “rural” TAZs into smaller spatial units, each carrying its own unique features. 2- Using high-resolution (disaggregated) modeling input data. 3- Using a detailed all functional class digital roadway network to include local streets and collector roads in order to represent realistic traffic patterns. The proposed model framework maintains the urbanized parts of the modeled area (e.g. CBD) without enhancement as they tend to include zones which are large in number, small in area, and dense in land use activities.
This research effort built a strong case in favor of transforming the traditional model into a high-fidelity one by having the basic zones model with TAZs at a finer level of geographic spatial resolution and achieving this goal through a full utilization of the available advanced technological means and “intelligent” data sources. This modeling transformation by applying the HFTDMF generated a unique graph (Fig. 5) underlying a solid theoretical concept relating number of TAZs to modeling accuracy (%R²) which clearly showed the behavior of improving modeling spatial resolution to the incremental gains in modeling performance (assignment accuracies and coverage capabilities).

The proposed HFTDMF improved traditional travel demand model estimation accuracy and network-wide assignment coverage by increasing traffic analysis zones (TAZs) “intensity” (i.e., spatial resolution and number count), especially rural coarse TAZs, as well as using all functional roadway classifications to define the roadway network. This “intensity” TAZ structure enhancement procedure (number and size) help to insure that fine-grained spatial or attribute data are not lost or averaged through aggregation. The procedure also captures most of the trips generated at each zone by maximizing the interzonal flow while minimizing the intrazonal flow and more effectively distribute traffic onto local and collector roads. Trips which have been captured by the trip generation step at each fine/uniform TAZ structure are more fully distributed by the trip distribution step and more effectively loaded onto all the functional classification roadway network links in the traffic assignment step. As shown with the GFA case study developed for this research paper, the overall accuracy improvement of enhancing the traditional modeling approach to a high-fidelity one was achieved based on following features: (1) Fine spatial TAZs, (2) Homogeneous TAZs with single or nearly single land-use based, (3) High-resolution sociodemographic and employment attributes, (4) Detailed (all functional) roadway network, (5) Realistic representation of centroid and centroid connectors. The promise of the zonal enhancement procedure was to improve the modeling estimation accuracies of the 3-step traditional model to include trip generation, trip distribution as well as the traffic assignment. This is very important as the state-of-the-art indicated that uncertainty (prediction errors) propagates from one step to another and in the end, it is compounded over the steps of the traditional modeling process. The modeled traffic volume is expected to have a higher uncertainty than the input data. The call for fine homogeneous TAZs resulted in a closer examination of the individual/household as a trip making unit. Consequently, detailed socio-economic household’s attributes and their activity location were identified within each TAZ in a more accurate and realistic way. High-resolution socioeconomic data improved trip generation/production estimates.
Another advantage of using small homogeneous spatial units is that zone centroids have more realistic representation of trip generations and attractions, which is based on population activities rather than on area centers of gravity. This accurate estimation and representation of trip production and attractions at each zone allow for robust trip distribution analysis.

Currently, the trip-based traditional modeling framework is the state-of-the-practice among planning/modeling agencies and switching to advanced models (integrated land-use and transportation models and activity-based models) has been hindered by additional costs encountered, lack of staff expertise, institutional barriers, detailed data requirements, and absence of “better” modeling results (TRB Special Report 288 2007; VHB Inc 2007). The same modeling barriers could be incurred when practitioners (modelers and planners) being offered to adopt the HFTDMF for the purpose of generating improved network-wide traffic volume estimates. The five key barriers and issues that could face the practicing modeling agencies and transportation community when adopting and implementing the HFTDMF are discussed as follows:

1- **Cost**: As requirements and resources for running the HFTDMF does not differ much from those running the traditional FSTDM, the cost of developing each model will be in close proximity. It could be argued that the added cost, in acquiring and processing the remotely sensed imagery, is not significant due to the fact that most jurisdictions are keeping HRSI for various applications and already hiring image processing technical experts.

2- **Staff Modeling Expertise**: Both the FSTDM and the HFTDMF almost have the same modeling paradigm, structure, and specifications which makes the impact of this barrier on adopting the HFTDMF somewhat irrelevant. The few added technical skills required to develop the HFTDMF such as image processing and AZT coding were not sustainability challenging.

3- **Institutional**: Requirements in relation to organizational staffing structure, budgeting and funding for both the FSTDM and the HFTDMF are almost the same due to the similar resources and skills needed to develop both models.

4- **Data**: The added data required for developing the HFTDMF were limited to Digital Property Maps (Parcels) and High-Resolution Satellite Imagery (HRSI). Most jurisdictions provided both GIS-based Parcel Maps and HRSI at a reasonable or no cost.
5- **Modeling Results**: Transportation community shifting from the FSTDM to the HFTDMF would be expected to experience significant modeling accuracy gains resulting from the shift. Improved network-wide traffic estimates were evident in developing and running the HFTDMF.

The above discussion presents the transportation community (agencies and practitioners) with clear evidence that the potential barriers to adopt the HFTDMF are unfounded and the added data as well as processing requirements in developing the HFTDMF from the FSTDM are neither substantial, nor overwhelming.

**Recommendations**

Based on conducting the research components from preparing the theoretical aspects (literature review and methodology) and developing the HFTDMF for GFA case study, the following recommendations are compiled:

- It is recommended that the transportation community (researchers and practitioners) introduce the spatial resolution dimension into the traditional (conventional) travel demand modeling framework for the purpose of improving the model predictive and assignment coverage capabilities to support contemporary decision-makings related to the transportation infrastructure.

- It is recommended that the transportation community give special attention to “rural” zones in the modeled area due to the fact that they highly influence the modeling accuracy. Rural zones contain a very low number of trips (high statistical error), very high intra-zonal trips (missing data error) and very coarse spatial resolution (high geographic scale error). Further research is highly recommended to consider large rural zones in the TAZ design process.

- The transportation community should fully utilize the available technological and data advancements in order to change the scope of the traditional travel demand modeling framework to a higher fidelity (or more disaggregated) one.

- It is recommended that the transportation community incorporate Remote Sensing (RS) and High-Resolution Satellite Imagery (HRSI) which provides a rich source of dynamic, accurate, detailed and fresh spatial data to compensate for the full reliance on the aggregated, rigid, and static census-based data. These high-resolution data...
images (mostly better than 5 m spatial resolution) contain spatial information about land use, land cover (classified land use, impervious surfaces, and vegetation), and urban features (roads and buildings).

- The use of HFTDMF as a resource reliable planning tool is recommended which could provide planners/engineers with improved traffic volume data at the network-wide level including lower functional class roads to support their decisions on various planning, design and operation issues and reduces their dependency on the costly traditional sensor-based traffic monitoring programs.

- It is recommended that improved link-based network-wide assignments generated form the GIS-based HFTDM framework be utilized to back-cast links by monitoring traffic counts at certain testing links and validate accuracies.

- Travel demand modeling software providers should incorporate geospatial functionalities for TAZ design based on aggregation and disaggregation algorithms in their software. In this case, the step of TAZ design would be an endogenous module and part of the 4-step traditional modeling framework.

- It is recommended that the transportation community shift the TAZ design practice, which is mostly governed by experiences and rules of thumb, to a systematic approach by conducting more innovative research in this field.

- High-resolution remotely sensed images should be incorporated to improve the spatial resolution of large rural zones to enhance the modeling accuracy.

- It is recommended to utilize and integrate the GIS-based high-resolution data sources from parcel-based digital property maps (DPM), the tax-assessment data, the land-use maps, and the point-based postal data into their models for purpose of improving modeling estimations.

- The TAZ structure spatial resolution (aggregation or disaggregation) should be limited to the resources available to achieve the model goals. The performance of high-fidelity travel demand models (HFTDMs) are highly dependent on the level of TAZ spatial resolution and limited by the diminishing marginal returns.

In conclusion, this research paper clearly indicated that there are real opportunities in employing advanced technologies, intelligent data sources, aggregation algorithms, and various areal interpolation procedures to enhance the traditional travel model to a HFTDMF one by improving the geographic spatial resolution of the modeled zones and using all functional classes in the roadway network. Fine TAZ structures ensured capturing most trips by
increasing their likelihood of crossing zonal boundaries, which in turn maximizes the interzonal flow, and by the same token, minimizes missing trips due to internalization. The results from the developed HFTDMF for the GFA case study showed that the model is capable of effectively distributing traffic onto all road classes including lower functional class local roads and collectors with improved modeling estimation accuracies.

References


VHB Inc. 2007. Determination of the state of the practice in metropolitan area travel forecasting: findings of the surveys of metropolitan planning organizations. Prepared for Transportation Research Board Committee B0090, Washington, D.C.


Figure Captions

Fig. 1. GIS-based Intelligent Data Framework for Developing HFTDM
Fig. 2. GFA composed of Fredericton (CA 320) and Oromocto (CSD 1303012)
Fig. 3. Modeled area with an urban core, rural fringe and MIZ entities.
Fig. 4. ISA for the rural component of the GFA

Fig. 5. 2nd Percentile threshold parcels (pink) overlaid with original parcels (yellow)

Fig. 6. Relationship between “Real” R² for calibrated models and number of TAZs from GFA 10 TAZ structures
Fig. 1. GIS-based Intelligent Data Framework for Developing HFTDM

Fig. 2. GFA composed of Fredericton (CA 320) and Oromocto (CSD 1303012)
Fig. 3. Modeled area with an urban core, rural fringe and MIZ entities.

Fig. 4. ISA for the rural component of the GFA
Fig. 5. 2nd Percentile threshold parcels (pink) overlaid with original parcels (yellow)

Fig. 6. Relationship between “Real” $R^2$ for calibrated models and number of TAZs from GFA 10 TAZ structures
Table 1. Sociodemographic and employment database for GFA 166 DAs

<table>
<thead>
<tr>
<th>TAZ ID</th>
<th>DA ID</th>
<th>Pop. Count</th>
<th>No. of (HH)</th>
<th>No. of (DU)</th>
<th>Avg. Ann Income/ HH</th>
<th>Non-Retail Emp.</th>
<th>Retail Emp.</th>
<th>CSD Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13100299</td>
<td>595</td>
<td>235</td>
<td>273</td>
<td>43795</td>
<td>225</td>
<td>0</td>
<td>Douglas</td>
</tr>
<tr>
<td>2</td>
<td>13100292</td>
<td>760</td>
<td>300</td>
<td>372</td>
<td>37040</td>
<td>35</td>
<td>10</td>
<td>Bright</td>
</tr>
<tr>
<td>165</td>
<td>13100305</td>
<td>441</td>
<td>190</td>
<td>191</td>
<td>50284</td>
<td>50</td>
<td>10</td>
<td>Fredericton City</td>
</tr>
<tr>
<td>166</td>
<td>13100306</td>
<td>1674</td>
<td>615</td>
<td>618</td>
<td>77977</td>
<td>222</td>
<td>18</td>
<td>Douglas</td>
</tr>
<tr>
<td>Total</td>
<td>166</td>
<td>94,374</td>
<td>38,015</td>
<td>41,186</td>
<td>63,630</td>
<td>41,024</td>
<td>6,405</td>
<td>14 CSDs</td>
</tr>
</tbody>
</table>

Table 2. 52 rural DA TAZs from Fredericton (CMA 320)

<table>
<thead>
<tr>
<th>CSD Name</th>
<th>CSD ID</th>
<th>No. of DAs</th>
<th>CDS Area (km²)</th>
<th>Avg. DA Area(km²/DA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Douglas Parish</td>
<td>1310028</td>
<td>7</td>
<td>1452.72</td>
<td>207.12</td>
</tr>
<tr>
<td>Bright Parish</td>
<td>1310026</td>
<td>5</td>
<td>417.07</td>
<td>83.40</td>
</tr>
<tr>
<td>Kingsclear Parish</td>
<td>1310018</td>
<td>10</td>
<td>320.25</td>
<td>32.03</td>
</tr>
<tr>
<td>New Maryland Parish</td>
<td>1310001</td>
<td>4</td>
<td>387.51</td>
<td>96.88</td>
</tr>
<tr>
<td>Saint Marys Parish</td>
<td>1310031</td>
<td>6</td>
<td>753.94</td>
<td>125.66</td>
</tr>
<tr>
<td>Lincoln Parish</td>
<td>1303008</td>
<td>11</td>
<td>165.76</td>
<td>15.07</td>
</tr>
<tr>
<td>Maugerville Parish</td>
<td>1303016</td>
<td>3</td>
<td>928.20</td>
<td>309.40</td>
</tr>
<tr>
<td>New Maryland Village</td>
<td>1310002</td>
<td>6</td>
<td>21.20</td>
<td>3.53</td>
</tr>
<tr>
<td>Total</td>
<td>8 CSDs</td>
<td>52</td>
<td>4,446.64</td>
<td>85.51</td>
</tr>
</tbody>
</table>

Table 3. Final 10 GFA TAZ structures

<table>
<thead>
<tr>
<th>GFA TAZ Structure</th>
<th>TAZ Type</th>
<th>No. of DAs (City of Fredericton, Town of Oromocto and FNRs)</th>
<th>AZT Run No. of Generated “Fine” TAZs</th>
<th>Total No. of GFA TAZs</th>
<th>Average Area/ TAZ (km²)¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run1</td>
<td>AZD-based</td>
<td>114</td>
<td>4138</td>
<td>4252</td>
<td>1.09</td>
</tr>
<tr>
<td>Run2</td>
<td>AZD-based</td>
<td>114</td>
<td>2008</td>
<td>2122</td>
<td>2.17</td>
</tr>
<tr>
<td>Run3</td>
<td>AZD-based</td>
<td>114</td>
<td>1436</td>
<td>1550</td>
<td>2.98</td>
</tr>
<tr>
<td>Run4</td>
<td>AZD-based</td>
<td>114</td>
<td>1318</td>
<td>1432</td>
<td>3.22</td>
</tr>
<tr>
<td>Run5</td>
<td>AZD-based</td>
<td>114</td>
<td>1286</td>
<td>1400</td>
<td>3.30</td>
</tr>
<tr>
<td>Run6</td>
<td>AZD-based</td>
<td>114</td>
<td>1276</td>
<td>1390</td>
<td>3.32</td>
</tr>
<tr>
<td>Run7</td>
<td>AZD-based</td>
<td>114</td>
<td>1274</td>
<td>1388</td>
<td>3.324</td>
</tr>
<tr>
<td>Optimal Run</td>
<td>AZD-based</td>
<td>114</td>
<td>512</td>
<td>626</td>
<td>7.470</td>
</tr>
<tr>
<td>DA (base-case) Level</td>
<td>Census-based</td>
<td>n/a</td>
<td>114</td>
<td>166</td>
<td>27.790</td>
</tr>
<tr>
<td>CT Level</td>
<td>Census-based</td>
<td>n/a</td>
<td>n/a</td>
<td>27</td>
<td>170.859</td>
</tr>
</tbody>
</table>

Note: ¹ Based on study area totaling 4,613.19 km²

Table 4. Outline of areal interpolation framework for the GFA

<table>
<thead>
<tr>
<th>GFA TAZ Structure</th>
<th>TAZ Type</th>
<th>SD&amp;E Data Generation Method</th>
<th>Source TAZ (No. of TAZs)</th>
<th>Target TAZs (No. of TAZs)</th>
<th>SFD</th>
<th>Relational Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parcels at Run1</td>
<td>AZD-based</td>
<td>AIM</td>
<td>DA (52)</td>
<td>4138</td>
<td>ISAR</td>
<td>“one-to-many”</td>
</tr>
<tr>
<td>Parcels at Run2</td>
<td>AZD-based</td>
<td>AIM</td>
<td>DA (52)</td>
<td>2008</td>
<td>ISAR</td>
<td>“one-to-many”</td>
</tr>
<tr>
<td>Parcels at Run3</td>
<td>AZD-based</td>
<td>AIM</td>
<td>DA (52)</td>
<td>1436</td>
<td>ISAR</td>
<td>“one-to-many”</td>
</tr>
<tr>
<td>Parcels at Run4</td>
<td>AZD-based</td>
<td>AIM</td>
<td>DA (52)</td>
<td>1318</td>
<td>ISAR</td>
<td>“one-to-many”</td>
</tr>
<tr>
<td>Parcels at Run5</td>
<td>AZD-based</td>
<td>AIM</td>
<td>DA (52)</td>
<td>1286</td>
<td>ISAR</td>
<td>“one-to-many”</td>
</tr>
<tr>
<td>Parcels at Run6</td>
<td>AZD-based</td>
<td>AIM</td>
<td>DA (52)</td>
<td>1276</td>
<td>ISAR</td>
<td>“one-to-many”</td>
</tr>
<tr>
<td>Parcels at Run7</td>
<td>AZD-based</td>
<td>AIM</td>
<td>DA (52)</td>
<td>1274</td>
<td>ISAR</td>
<td>“one-to-many”</td>
</tr>
<tr>
<td>Parcels at Optimal Run</td>
<td>AZD-based</td>
<td>AIM</td>
<td>DA (52)</td>
<td>512</td>
<td>ISAR</td>
<td>“one-to-many”</td>
</tr>
<tr>
<td>DA</td>
<td>Census-based</td>
<td>SCAN</td>
<td>DA (166)</td>
<td>DA (166)</td>
<td>n/a</td>
<td>“one-to-one”</td>
</tr>
<tr>
<td>CT</td>
<td>Census-based</td>
<td>SCAN</td>
<td>CT (27)</td>
<td>CT (27)</td>
<td>n/a</td>
<td>“one-to-one”</td>
</tr>
</tbody>
</table>
Table 5. “Real” R² for base-case and calibrated models at 10 TAZ structures

<table>
<thead>
<tr>
<th>TAZ Structure Level</th>
<th>No. of TAZs</th>
<th>3-Step TDM Type</th>
<th>Base-Case Model¹</th>
<th>Final Calibrated Model²</th>
<th>Calibrated at Run #</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT</td>
<td>27</td>
<td>TTDM</td>
<td>NA</td>
<td>0.1544</td>
<td>0.2490 3</td>
</tr>
<tr>
<td>DA</td>
<td>166</td>
<td>TTDM</td>
<td>0.2531</td>
<td>0.4702</td>
<td>0.5070 3</td>
</tr>
<tr>
<td>Optimal</td>
<td>626</td>
<td>HFTDM</td>
<td>0.3413</td>
<td>0.5220</td>
<td>0.6235 4</td>
</tr>
<tr>
<td>AZT RUN7</td>
<td>1388</td>
<td>HFTDM</td>
<td>0.3465</td>
<td>0.5260</td>
<td>0.6085 3</td>
</tr>
<tr>
<td>AZT RUN6</td>
<td>1390</td>
<td>HFTDM</td>
<td>0.3515</td>
<td>0.5257</td>
<td>0.6115 3</td>
</tr>
<tr>
<td>AZT RUN5</td>
<td>1400</td>
<td>HFTDM</td>
<td>0.2527</td>
<td>0.5426</td>
<td>0.5556 3</td>
</tr>
<tr>
<td>AZT RUN4</td>
<td>1432</td>
<td>HFTDM</td>
<td>0.3580</td>
<td>0.5266</td>
<td>0.6089 3</td>
</tr>
<tr>
<td>AZT RUN3</td>
<td>1550</td>
<td>HFTDM</td>
<td>0.3551</td>
<td>0.5049</td>
<td>0.6100 3</td>
</tr>
<tr>
<td>AZT RUN2</td>
<td>2122</td>
<td>HFTDM</td>
<td>0.3694</td>
<td>0.5178</td>
<td>0.6255 4</td>
</tr>
<tr>
<td>AZT RUN1</td>
<td>4252</td>
<td>HFTDM</td>
<td>0.3783</td>
<td>0.5120</td>
<td>0.5992 3</td>
</tr>
</tbody>
</table>

Note: ¹(α = 0.15, β = 4.0, and RG = 0.01); ²(α = 0.84, β = 5.50, and RG = 0.002).

Table 6. Assignment capabilities for base-case models at 10 TAZ structures

<table>
<thead>
<tr>
<th>TAZ Structure Level</th>
<th>No. of Links</th>
<th>PAC_NW²</th>
<th>PAC_LFCRs³</th>
<th>PAC_HFCRs⁴</th>
<th>CLLPAC_NW⁵</th>
<th>CLLPAC_LFCRs⁶</th>
<th>CLLPAC_HFCRs⁷</th>
<th>AAV_NW⁸</th>
<th>AAV_LFCRs⁹</th>
<th>AAV_HFCRs¹º</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZT RUN1</td>
<td>4324</td>
<td>75</td>
<td>70</td>
<td>91</td>
<td>71</td>
<td>65</td>
<td>85</td>
<td>2495</td>
<td>1403</td>
<td>5377</td>
</tr>
<tr>
<td>AZT RUN2</td>
<td>4182</td>
<td>72</td>
<td>67</td>
<td>90</td>
<td>69</td>
<td>63</td>
<td>85</td>
<td>2663</td>
<td>1500</td>
<td>5634</td>
</tr>
<tr>
<td>AZT RUN3</td>
<td>4065</td>
<td>70</td>
<td>65</td>
<td>89</td>
<td>68</td>
<td>61</td>
<td>84</td>
<td>2815</td>
<td>1598</td>
<td>5856</td>
</tr>
<tr>
<td>AZT RUN4</td>
<td>3980</td>
<td>69</td>
<td>63</td>
<td>89</td>
<td>67</td>
<td>60</td>
<td>84</td>
<td>2895</td>
<td>1656</td>
<td>5921</td>
</tr>
<tr>
<td>AZT RUN5</td>
<td>3991</td>
<td>69</td>
<td>63</td>
<td>89</td>
<td>67</td>
<td>60</td>
<td>84</td>
<td>3164</td>
<td>1751</td>
<td>6629</td>
</tr>
<tr>
<td>AZT RUN6</td>
<td>3983</td>
<td>69</td>
<td>63</td>
<td>89</td>
<td>68</td>
<td>61</td>
<td>84</td>
<td>2906</td>
<td>1666</td>
<td>5942</td>
</tr>
<tr>
<td>AZT RUN7</td>
<td>3972</td>
<td>69</td>
<td>62</td>
<td>89</td>
<td>67</td>
<td>60</td>
<td>84</td>
<td>2920</td>
<td>1675</td>
<td>5956</td>
</tr>
<tr>
<td>Optimal</td>
<td>3762</td>
<td>65</td>
<td>59</td>
<td>87</td>
<td>61</td>
<td>53</td>
<td>81</td>
<td>3135</td>
<td>1811</td>
<td>6217</td>
</tr>
<tr>
<td>DA</td>
<td>3014</td>
<td>52</td>
<td>45</td>
<td>78</td>
<td>47</td>
<td>39</td>
<td>65</td>
<td>4144</td>
<td>2659</td>
<td>7095</td>
</tr>
<tr>
<td>CT</td>
<td>1663</td>
<td>29</td>
<td>19</td>
<td>64</td>
<td>26</td>
<td>15</td>
<td>52</td>
<td>10241</td>
<td>7440</td>
<td>13053</td>
</tr>
</tbody>
</table>

Notes:
- Based on Base-Case RN Links of 5787¹
- PAC_NW²: Percentage Assignment Capability at Network-wide Link Level
- PAC_LFCRs³: Percentage Assignment Capability at Lower Functional Class Roads (LFCRs) Level
- PAC_HFCRs⁴: Percentage Assignment Capability at High Functional class Roads (HFCRs) Level
- CLLPAC_NW⁵: Center Line Length %age. Assignment Coverage at Network-wide Link Level
- CLLPAC_LFCRs⁶: Center Line Length %age. Assignment Coverage at LFCRs Level
- CLLPAC_HFCRs⁷: Center Line Length %age. Assignment Coverage at HFCRs Level
- AAV_NW⁸: Avg. Assigned Volume for RN
- AAV_LFCRs⁹: Avg. Assigned Volume for LFCRs
- AAV_HFCRs¹º: Avg. Assigned Volume for HFCRs

https://mc06.manuscriptcentral.com/cjce-pubs