Challenges and Opportunities in Applying High-Fidelity Travel Demand Model for Improved Network-Wide Traffic Estimation: A Review and Discussion

Riad Mustafa[*,1] and Ming Zhong[2]

[1]*Department of Civil Engineering, University of New Brunswick, Fredericton, N.B., E3B 5A3, Canada*

[2]*Intelligent Transport Systems Research Center, Wuhan University of Technology, Wuhan, P.R. China*

Abstract: Estimating traffic volume at a link level is important to transportation planners, traffic engineers, and policy makers. More specifically, this vital parameter has been used in transportation planning, traffic operations, highway geometric design, pavement design, and resource allocation. However, traditional factor approach, regression-based models, and artificial neural network models failed to present network-wide traffic volume estimates because they rely on traffic counts for model development, and they all have inherent weaknesses. A review to previous research work and the state-of-practice clearly indicates that the Traditional Four-step Travel Demand Model (TFTDM) was generally based on large traffic analysis zones (TAZs) and networks consisting of high functional-class roads only. Consequently, this conventional modeling framework yielded a limited number of link traffic assignments with fairly high estimation errors. In the light of these facts and the obvious need of accurate network-wide traffic estimates, this review is conducted. In particular, this paper provides an extensive review of using traditional travel demand models for improved network-wide traffic volume estimation. The paper then focuses on the challenges and opportunities in achieving high-fidelity travel demand model (HFTDM). This review has revealed that, opportunities in relation to both technological advances and intelligent data present a substantial potential in developing the proposed HFTDM for a much more accurate traffic estimation at a network-wide level. Finally, the paper concludes with key findings from the review and provides a few recommendations for future research related to the topic.

Keywords: Four-step travel demand model, high-fidelity travel demand model, traffic analysis zones, network-wide traffic volumes, geographic information systems, remote sensing, intelligent data sources.

INTRODUCTION: BACKGROUND AND PURPOSE

Traffic volume estimates are important to transportation planners, traffic engineers, and policy makers. More precisely, they are used by many Departments of Transportation (DOTs) and highway agencies to help plan, build, and maintain transportation infrastructure at county, provincial, and national levels. Traditionally, traffic volume estimates are most often obtained from count stations (either temporary or permanent) installed on limited locations covering mainly the upper functional class of the roadway network (arterial and up) due to expensive costs of installation and maintenance. Further, attempts to estimate network-wide traffic volume using analytical approaches, such as regression-based models and artificial neural network models (ANN), were criticized for their limited capability of estimating traffic volumes for groups of functionally classified roads rather than individual ones. Furthermore, ANN approach is mainly applied for research purposes and too complicated for practitioners. In addition, all existing traditional methods require traffic counts for model development. However, since traffic counts are not

available for a significant portion of road network, especially for low-class roads, these techniques fall short in providing network-wide traffic estimates.

Traditional Four-step Travel Demand Model (TFTDM), developed in U.S. more than fifty years ago, is still used by various highway and transportation agencies worldwide to predict traffic volumes for specific roadway network links at both local and regional levels despite well documented weaknesses and flaws [1,2]. Traditionally, the four-step travel demand model (FSM) is represented by aggregate zones. It divides a study area into traffic analysis zones (TAZs) sometimes at a very coarse level, which are assumed to be spatially homogeneous and zonal attributes are mostly represented by aggregate averages. Moreover, traditional models use skimmed roadway networks, which mostly ignore low-class roads (collectors and local streets).

Evidently, fairly large TAZs and limited roadway network (arterials and above) used in the traditional travel demand modeling framework caused biased and unbalanced trip distribution and assignment over the roadway network. Large zones tend to yield a higher percentage of intrazonal trips, which are not accounted for because trips generated in a zone were forced to end at the same zone. Therefore, only limited intrazonal traffic was assigned to low-class roads causing underestimated traffic estimates on this class of

*Address correspondence to this author at the Department of Civil Engineering, University of New Brunswick, Fredericton, N.B., E3B 5A3, Canada; Tel: 647-200-4811; Fax: 506-453-3568; E-mail: riad.m@unb.ca

roads and overestimated on higher functional-class of roads resulting in significant errors and limited traffic assignments on low-class road. More precisely, local roads are mostly ignored and presented by "centroid connectors", which are characterized as a simplified representation of the local roadway networks that let individuals access the major roadways. Although local roads account for the majority of any roadway network (about 85% in Canada) both traditional traffic counting schemes and modeling procedures ignored this functional class road, therefore, traffic information of this class is either not available or very inaccurate. The accuracy of link volume estimation is measured by comparing modeled traffic volumes to traffic count data obtained from various traffic count programs on a link-by-link basis through statistical validation parameters. The mostly used link validation parameters are: Percent Assignment Error (PAE), Absolute Percent Assignment Error (APAE), Root Mean-Square Error (RMSE), and Coefficient of Determination (R^2). Then, errors of traffic assignment models are compared with maximum allowed based on documented guidelines or local practices developed by various highway/transportation agencies [3-5].

Generally speaking, traditional travel demand models failed to provide network-wide traffic estimation, which could be validated by investigating the percentage of Zero-Assignment Link **(**ZAL) [6]. The test allows modelers to evaluate the model performance in relation to network-wide traffic assignment capabilities. To this day, the proposed ZAL test lacks documentation because transportation agencies mostly ignored the issue of network-wide volume estimation capabilities of traditional travel demand model as part of their recommended calibration/validation checks. Ultimately, the ZAL test should be an addition to documented "reasonableness checks" included in the guidelines and practice documents published by various transportation/highway agencies dealing with travel demand modeling issues. Limited network assignment capability is a serious problem because of the implications of using link volume estimates in Vehicle Miles of Travel (VMT) calculations for many applications such as emission modeling. Vehicles-Miles of Travel (VMT) is estimated by the ability of the model to replicate observed VMT collected from various counting programs. Modeled VMT is validated using performance measure of percent error conducted at the roadway network level of a study area and at facility type such as local roads level. The former is basically a reasonableness test on the four steps of TFTDM: trip generation, trip distribution, model split, and traffic assignment, while the later is a close check on the assignment algorithm that includes link data summaries of travel time, travel speed and capacity. Various highway/transportation agencies used such guidelines to compare between modeled and measured VMT [3].

In an effort to overcome the aforementioned shortcomings of traditional travel demand models, some recent technological advancements and intelligent data can help disaggregate geographic space allowing for levels of spatial analysis finer than ever before. Today, employing intelligent data-collection technologies and data sources such as remotely sensed images, parcel-based digital property maps (DPM) and tax-assessment data, land-use maps, point-based postal data for improving modeling estimations could be entirely completed within a Geographic Information Systems (GIS) environment. Indeed, GIS hold particular promise to overcome limitations of the aggregated coarse zonal structure. After all, GIS is a predominantly data-processing tool with various functionalities for spatial analysis, including managing, sorting, analyzing, and displaying data, which can greatly facilitate the integration of diverse intelligent data sources as presented in the systematic integration framework shown in Fig. (**1**).

To begin with, the general purpose of this paper is three-fold: (1) to review the current practices of GIS-based traditional travel demand models in relation to network-wide traffic volume estimation; (2) to identify the challenges to achieve network-wide traffic modeling; and (3) to discuss the existing opportunities from the emerging technological and intelligent data perspectives. Accordingly, this paper is organized into the following four sections: in the first section, a critical review of the literature in relation to network-wide traffic volume modeling/estimation and capabilities using the traditional four-step model is conducted; then, we present challenges facing the development of HFTDM together with several opportunities to achieve this goal in second section. In the third section, we introduce the potential contributions from HFTDM. Finally, in the fourth section we discuss a number of conclusions identified from the review and list a set of recommendations for future research in the field.

CRITICAL REVIEW OF LITERATURE

Literature on traditional four-step travel demand models is vast and couldn't be covered in this single review paper. Hence, this literature review is being conducted in the context of exploring the possibility of enhancing traditional travel demand model for network-wide traffic estimates purposes. Thus, the focus of this review is on TFTDM related to the following topics: (1) Modeling practices, (2) Modeling network-wide traffic volumes, (3) Traffic Analysis Zone (TAZ) design and selection, and (4) Advanced technology and "intelligent" data sources for modeling purposes.

Review of Traditional Modeling Practices

Model in transportation planning is a decision-support tool for planning activities and policy scenario analysis. Traditional travel demand models are a set of mathematical/"statistical" relationships constitutes four-step procedure developed as a response to the U.S. major interstate highway construction projects carried out in the 1950s. Therefore, by large, these models were developed with macroscopic perspective to help size capital facilities. After all, it is in this regard, Douglass Lee published in the Journal of American Planners his famous paper "Requiem for large-scale models" [7] where the role of models as a tool for scientific planning was severely criticized. Lee expressed his concern about the planning context resulting from large scale urban models, which weaken their response to well-defined policy issues. Recently, Horowitz voiced the scale concern on his statewide travel forecasting synthesis review document. Although the author acknowledged the many successful applications of statewide models, he mentioned obstacles, which are facing modelers such as the scale of statewide model and the coarse zone system being used [8].
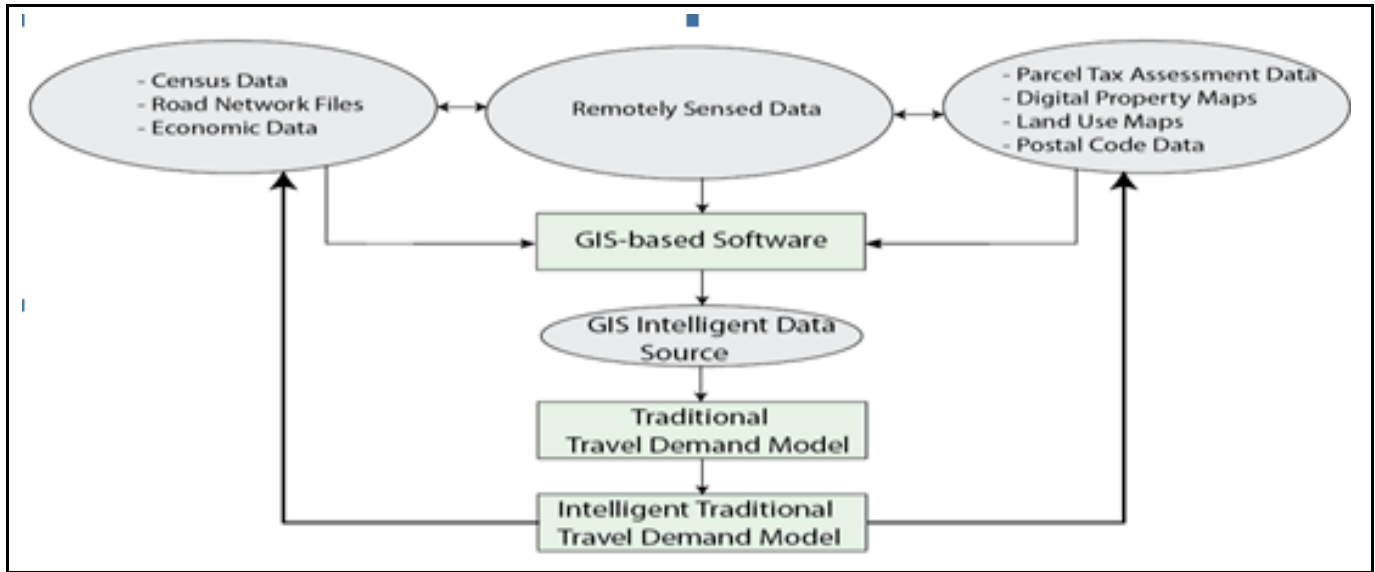
Fig. (1). GIS-based Intelligent Data Framework for Traditional TDMs.

Generally speaking, modeling practices follow three broad scale and organizational hierarchy presented as follows:

- **Macroscopic:** Statewide Models: State Departments of Transportation (DOTs).

- **Mesoscopic:** Regional Models: Metropolitan Planning Organizations (MPOs).

- **Microscopic:** County/Municipal Models: Local planning bodies.

Various jurisdictions around the world and in particular many states in USA developed their models based on the Urban Transportation Planning System (UTPS) process, which is a typical standardized version of the "four-step" models. Hence, this led to deficiencies in statewide travel demand models, which are listed as follows [9]: (1) size of networks may overly burden the computational process, (2) size of some zones led to very coarse traffic assignments; and (3) omission of many roads led to excessive volumes on links near urban areas. Further, Horowitz and Farmer [10] noted that the most prominent problems in implementing statewide models relate to issues of scale. Because of the zonal aggregation in urban areas, there are a large number of intrazonal trips, which are discarded during the assignment step. The authors concluded that, of all statewide travel demand model steps, traffic assignment appears to be the weakest. Additionally, scale issue and coarse zone system were further emphasized by two separate peer [8,11], which also have recognized the maturity of some operational statewide models and further reported their successes.

Metropolitan Planning Organizations (MPOs) in the U.S. were established as part of the Federal-Aid Highway Act of 1962 and charged with regional transportation planning for metropolitan areas with population of fifty thousand or more. Further, MPOs are required to use travel demand forecasting models to estimate future travel demand and analyze the impacts of alternative transportation investment scenarios [12]. Various studies [12-14] noted general satisfaction expressed by MPOs with the four-step model and shifting to advanced modeling procedures was challenged by budget constraints and lack of expertise. Furthermore, MPOs were spectacle that "advanced practices" will yield improved forecasts or will permit the agencies to address questions that now go unanswered. Evidently, the study showed general consistency in the modeling framework used by MPOs which was GIS-based traditional four-step model or three-step by omitting model choice model from the process.

The literature covering modeling practices at the municipal/local level is scarce partly because it is a continuation of small MPO models, which typically are used directly by local municipalities and modified to address their own planning purposes and modeling needs. For example, in Canada, municipalities and urban areas are primarily involved in strategic transportation plans and travel demand forecasting. And yet, most urban areas have a 4-step model and there are no activity-based models in operation within municipalities [13].

Modeling Network-Wide Traffic Volumes with Improved Accuracy

This section of the literature review will examine both the state-of-practice and state-of-art of using TFTDM as a tool for estimating network-wide traffic volumes. This will be achieved by considering TAZ structure and roadway network details. Although the interaction between the zoning structure and the detail of network was of concern to transportation community, it appears that it has been largely overlooked in the literature and in the practice. Whereas, this interaction should not be ignored, the general recommendation is to use a more detailed roadway network with small TAZ structure and a less detailed one for large TAZ for network-wide traffic estimation [15], which sounds like a common sense. As part of their efforts to estimate traffic volumes for low-class roads using traditional travel demand model, Zhong and Hanson [16] developed a GIS-based travel demand model to estimate traffic volume for low-class roads. The model was applied to York County and Beresford Census Consolidated Subdivision (CCS) in the Province of New Brunswick. Their study was in agreement with [15] regarding increased estimation accuracy achieved

by reducing the size of the analysis zone. However, both studies didn't investigate the impact of TAZ disaggregation on the model's capability in terms of network-wide traffic assignment and estimation.

As mentioned previously, Khatib *et al.* [15] studied the effect of TAZ structure on travel demand forecasts. The study concluded that the use of fairly large TAZs produced results with fairly high estimation errors, mainly related to the unbalanced trip distribution and limited traffic assignment over the highway network. Additionally, a GIS-based travel demand model was developed based on a variety of TAZ structures, centroid locations, and road network for Latah County in Idaho, USA. The assignment results were validated based on Idaho Transportation Department (ITD) traffic counts. The study concluded that the smaller the TAZ the higher the model estimation accuracy. Also, the authors noted that TDMs using the detailed road network by including roads from all functional classes achieved better assignment results regardless which level of TAZ being used. The recommendation is consistent with the above "common sense" use a more detailed roadway network with smaller TAZ structure and a less detailed one for larger TAZ for network-wide traffic estimation. It will be of interest for future research to investigate the impact of the combination of 11 TAZ structures with the four levels of centroid locations and detailed road networks (employed by the study) on the model's capability of network-wide assignment and traffic estimation. Finally, Cambridge Systematic -Inc [17] mentioned that large TAZs with both production and attractions (dwelling and employment) will ignore the assignment of large number of intrazonal trips. They recommended reducing the number of intrazonal trips by decreasing the size of the TAZs.

The need to estimate traffic volumes on local roads without excessive cost incurred by traditional data collection schemes and limitations of the aggregated coarse zonal structure inspired Mustafa and Zhong [18] to develop GIS-based improved travel demand model using space disaggregation technique. The grid-based technique disaggregated coarse TAZs into three levels of smaller spatial units, each carrying its own unique features. The use of smaller homogenous zones resulted in effective distribution of traffic onto all road classes including low-class local roads which in turn improved the capability of assignment coverage of TFTDM. Test results from this research resulted in an increase of overall network coverage of 17% with most improvement contributed to low class local roads.

In their efforts to model truck traffic within a municipality on minor arterials, collectors and local roads, Marker and Goulias [19] developed a methodology based the new Quick Response Freight Manual (QRFM) for modeling truck traffic in Centre County, Pennsylvania, using TransCAD (a GIS transportation planning software). A comparison is made in truck traffic estimation between two models when the model resolution is changed. One model used census tracts and the other used census blocks and block groups as TAZs. Both models used the same network, which included all major highways and most local roads in the urbanized region of the county. The results of these two models were compared with the truck counts from Highway Performance Monitoring System (HPMS) databases obtained from the Pennsylvania Department of Transportation (PennDOT). The study reported that the estimated truck traffic link volumes favored the use of the more disaggregate TAZ scheme, which was based on blocks and block groups. The study concluded that the use of disaggregated TAZs improved the overall model accuracy significantly, but with almost all improvements to increased accuracy over low-class local roads.

In this regard, Horowitz [20] attempted improving traffic assignment based on fine spatial resolution and detailed roadway network. First of all, the author identified the following problems in traffic assignment procedures employed in traditional four-step model: (1) large zones with their true trip origins or destinations being considerably distant from their zone's centroid or from the point at which the centroid connectors attach to the arterial network, and (2) traditional network coding methods, using hypothetical centroid connectors and ignoring local streets. For these reasons, the author proposed the following spatial precision in his traffic assignment: sub-zoning around links by creating small TAZs referenced to its corresponding centroid and drawn as polygons on top of the network. Basically, traffic assignment is performed where each TAZ is drawn as a link service area in what the author referred to as "area-spread assignment". The author tested area-spread assignment on a detailed network in Fredericton, New Brunswick that contained all streets in the city and extremely small zones. The study reported that area-spread assignment did a better job of replicating volumes on the detailed network than traditional coarse TAZ assignment, although computation times were much longer. The author pointed out that this difference by itself should be a sufficient incentive to planners to seriously consider spatial issues in network coding and traffic assignment. Finally, the author concluded that a traffic assignment with all local streets should produce better results than an assignment without them. Fortunately, the author's concerns of computational and source data constraints to perform high-resolution spatial analysis is no longer holding with recent technological advances in processing speed and storage along with GIS.

This literature review investigated a number of traditional four-step models developed at varying level of geographies (small city, county, regional, and 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 low-level roads (collector or lower) [8,14,21-23]. More to the point, TRB Special Report 288 [12] mentioned that highway networks coded by MPOs usually include all or almost all freeway, expressways, major arterial, minor arterial mileages and feeder/collector routes but not local roads. The document reported that on average less than 25% of local mileage is represented. As local streets are rarely coded into model network, they have been simulated by the centroid connectors [14]. In this regard, the network developed under "the-state-of-the-practice" guidelines recommended coding links up to collectors to capture at least 85% of the vehicle trips [24]. Moreover, the guideline document claimed that this level of network detail would be able to capture 90% or more of the vehicle trips. From our point of view, 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 fairly large amount of traffic is local. Needless to say, NHCHR 365 [21], a master document of traditional travel demand modeling focusing primarily on the needs of smaller urban areas, considered local roads as "unwanted details", which should be filtered out when building up the model network. The document recommended, an ambiguous rule-of-thumb, to include only local streets that carry a substantial volume of traffic (e.g., those directly connected with arterials and highways). It can be argued that local streets are not functionally designed to carry significant traffic volume at link level, but collectively, they have the potential to carry significant traffic volume because they account for the majority of any roadway network (about 85% or more in Canada). Thus, ignoring this functional class from any roadway network used in modeling analysis will no doubt have serious consequences.

To this end, literature review conducted in this research paper clearly revealed that achieving improved TFTDM is highly dependent on disaggregating the study area with a fine zone structure and transforming related census data into the developed fine zones. To start with, spatial disaggregation is accompanied with the challenge of determining an optimal TAZ structure (number and size), which will result in the best overall model accuracy. Also, it is critical to determine the most suitable statistical measures to be used for model accuracy evaluation. State-of-the-practice revealed the use of R^2 (Coefficient of Determination) or "goodness of fit" as a measure for model accuracy. Optimal R^2 is related to the optimal TAZ structure and can be studied using the concept of diminishing marginal returns. For each iteration run, the model uses a finer TAZ structure, which will improve R^2 value (close to 1.00) towards an optimum value after which the marginal effect of the fine TAZ structure on R^2 diminishes. In our opinion, for the purposes of testing model's sensitivity to the spatial resolution of TAZs, the diminishing marginal effect of fine TAZ structure on the overall model accuracy should be evaluated for each TAZ disaggregation technique proposed in the development of TFTDM. On the other hand, transformation of census data from one level of spatial geography to another (areal interpolation) cannot ignore errors and biases associated with data aggregation, which is related to the way the data being collected and reported before any further spatial disaggregation. No doubt, coarse zones highly influence the modeling accuracy because they may contain 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). 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, which involves spatial disaggregation of TAZs and transformation of socioeconomic data and both are associated with spatial analytical issues, which can substantially affect the modeling results. Miller [25] identified some of the major spatial issue related to travel demand modeling as follows: (1) modifiable area units problem (MAUP) and ecological fallacy (EF), (2) spatial dependence (association) and spatial heterogeneity. The

former is related to scale and zoning effects while the later touches spatial dependence and heterogeneity effects. Both issues can have a strong influence to the results of traffic demand modeling, but they have so far received little recognition from travel demand modeling community.

To conclude, despite the recognition of transportation community for the improved TFTDM, which is featured with higher spatial resolution TAZ structure, disaggregated input data and detailed networks including roadways from all functional classes, no one has addressed the fundamental question regarding the extent to which zonal disaggregation is feasible. In our opinion, the answer to this question is frequently raised during TAZ design process, which is performed outside the traditional modeling framework on one hand and highly dependent on a set of contradicting constraints (irrationalized process) on the other hand. Further, traditional four-step models are lacking a sensitivity analysis mechanism, which can provide a feedback linking the accuracy assessment of the outputs of traffic assignment models with the TAZ design step. Finally, TAZ design process is surrounded by the above spatial analytical issues, which so far received little attention from travel demand modeling community, although they could strongly influence travel demand modeling results. For the previously mentioned reasons, one should note, the relationship between zone spatial resolution and modeling accuracy has been recognized but not fully understood, which makes it an area open for more research and innovations.

Ultimately, the solution to the spatial aggregation issue above (the relationship between zone spatial resolution and modeling accuracy) seems to disaggregate space to the household level and modeling unit to individual person level. However, highly disaggregated TFTDM would likely shift the traditional molding framework to a new paradigm, which has a component of disaggregated urban land use model that interfaces with the traditional four-step model. Clear examples of this hybrid modeling approach are TRANSIMS and UrbanSim models. To start with, four-step model has inherited weaknesses in accurately forecasting traffic estimates under complex urban land-use transportation system. Most fundamentally four-step models are not behavioral in nature, using highly aggregate zones, and its outputs and estimates are not responsive to the complex range of policies from planners and politicians. Increasing environmental concerns to congestion and sprawling land use patterns, combined with four-step model inherited weaknesses, has led to the emergence of activity-based disaggregate microscopic modeling packages such as TRANSIMS.

TRANSIMS (TRansportation ANalysis and SIMulation System) initially developed in 1996 at Los Alamos National Laboratory (LANL) as a part of Travel Model Improvement Program (TMIP). TRANSIMS uses census and survey data to synthesize population (individuals and households) at micro level to simulate (in real time) trips and generate traffic dynamics. TRANSIMS modeling paradigm establishes an integrated, multi-modal, and regional transportation system analysis tool, which is based on a cellular automata microsimulator. Indeed, TRANSIMS is capable of combining demand modeling and traffic operation by modeling discreet individual vehicles or individuals and

platoons of traffic flows. TRANSIMS, an open source travel demand modeling software, which consists of the following five sequentially structured modules: population synthesizer, activity generator, route planner, traffic micro-simulator and emissions estimator.

TRANSIMS could be implemented on modular basis at two distinct tracks of 1 or 2. Track 2 is the typical five modules stating with population synthesizer module, while Track 1 is capable of integrating traditional four-step models of trip generation, trip distribution, and mode choice with the traffic assignment modules (Router and Microsimulator) TRANSIMS. This link between the two models may be able to fulfill the promise of transforming the static assignment from TFSTDMs to dynamic network simulation in TRANSIMS. Further, the integration of the two models will improve the modeling capabilities of TFSTDMs by including better Measures of Effectiveness (MOEs) for traffic operation purposes, and policy analysis at planning level.

The integration of disaggregated models with traditional four-step travel demand model- a hybrid modeling approach- is also evident in UrbanSim model. UrbanSim is a simulation system for supporting planning and analysis of urban land development, which takes into account the interactions between land use, transportation, the economy, and the environment. It is intended for use by various jurisdictions and transportation organizations for the purpose of exploring the impact of infrastructure development and policy choices on various community outcomes such as urban form, motorized and non-motorized accessibility, housing afford-ability, and greenhouse gas emissions (http://www.urbansim.org/Main/WebHome). UrbanSim is open source metro-politan-scale software developed in the late 1990s at the University of Washington. Basically, UrbanSim package integrates eight models to predict the household, employment, and land characteristics for each 150 square meter grid cell covering the modeled area. While its comprehensiveness is attractive to many land use modelers, others are deterred by the numerous data requirements [26].

Travel demand model is exogenous to UrbanSim and the link between the two models is two-way, since that is accessibility from the travel demand model will influence the location decisions of developers, employers, and residents in the long-run largely determines travel demands among different locations. The external travel model provides travel times and utilities to the Accessibility Model. The travel model is typically run only once every five simulated years or when there is a major change to the transportation system, since running it is relatively cumbersome and since its outputs generally change much more slowly than others in the simulation. However, UrbanSim is designed to run annually, updating the accessibility values based on the evolving spatial pattern of activities [27]. Further, UrbanSim can work with various transportation models such as those activity-based travel model with the aggregated zone [28] and general four-step TDM, which allows regional accessibilities to be recalculated at regularly scheduled interval [29].

Developers of the above disaggregated models (TRANSIMS and UrbanSim) are typically confronted by the high-resolution socioeconomic and demographic data required, which is scarce and difficult to attain, for building and calibrating such models. Further, it has been observed that the use of highly disaggregated zone system will result in large highly sparse O-D matrix with too many unpopulated cells. Overall, disaggregated models are recommended for jurisdictions that have sufficient resources and time to overcome data gathering challenges, and that have a need to answer policy questions regarding dynamic behavior, which simply cannot be answered using an aggregate model [30]. It is true that, a closer link between the macro-level traditional travel demand models and micro-level disaggregate (activity-based and microsimulation) models is attractive to the transportation community due to the benefit gains for both modeling paradigms. However, an important aspect of the link between macro-level and micro-level models is the ability of the macro-level model to provide the micro-simulation with right transport demand or accessibly values. This may not be easily achieved as the spatial resolution of the two levels are just too different. Equally important, macro-level modeling process could take advantage of micro simulation results, but it has to reconcile the micro-simulation errors or dynamics from that level. Further, Holyoak and Statzic noted that confidence in the results of "linked" macro and micro models will often depend on strong data-exchange relationships and consistency in the modeling theory and outputs achieved [31].

Traffic Analysis Zone (TAZ) Design and Selection

Traffic demand modeling exercise typically starts with defining a study area. Then, the study area is subsequently subdivided into basic geographic units called traffic analysis zones (TAZs) used for data collection, tabulation, and analysis. And zoning system is a long, established practice for incorporating spatial structure into transportation demand models. Moreover, each TAZ is characterized by socioeconomic and land use variables and therefore, TAZ selection and configuration is perceived to plays a major role in reducing the errors encountered in travel demand modeling.

Martinez and colleagues stated, "Defining a good set of TAZ is still one of the transportation unsolved problems" [32]. This problem stays because transportation community hasn't reached a consensus about TAZ configuration and design. In this regard, Horowitz [20] voiced concerns about choice of spatial precision of travel demand model and related this to the challenge of selecting the right size and placemen of traffic analysis zones (TAZs). Further, the author characterized TAZ selection as a compromise among data preparation costs, computation costs, and model accuracy. Finally, the author observed a recent trend toward greater spatial precision in models due to advances in computer speed and geographic information system (GIS) technology.

To date, TAZ design process is purely a planner and decision maker exercises related to different criteria being used and knowledge of the study area being modeled. Sometimes, the TAZ selection is as naive as being simply adopting exiting census geographies or even based on merely a perception to the study areas. After all, TAZ design and selection is performed outside the four-step models, even

though it is still a vital step in the whole modeling process. Evidently, a well-designed TAZ system will definitely have an impact on modeling results in terms of assignment capabilities and accuracy. Thus, this review will shed the light on TAZ design and selection following three general tactics: (1) employ existing census hierarchical geography (2) employ modeler's judgment guided by a set of "rules of thumb", and (3) employ rule-based analytical techniques (using GIS-based spatial analysis).

Using geographic areas from census at varying spatial scale (Dissemination Block (DB); Dissemination Area (DA); Census Tracts (CT); etc.) provides modelers readily available demographic data and saves efforts, time, and complications that may arise from designing a new TAZ structure. However, concerns have risen from the fact that census geographies are not designed for modeling purposes to start with, and therefore, they are highly aggregated to preserve confidentiality. Moreover, census boundaries are arbitrary in nature and drawn to follow existing roadway network for data collection purposes. Further, their design does not follow any TAZ design criteria and guidelines. Consequently, census boundaries may cross-cut demographic and economic features on the ground without considerations to a wide variation in size, population, and land use density related to various urban and rural census geographic areas. Certainly, this will pose concern to modeling results. Despite those mentioned pitfalls, researchers continue to rely on census because it is the only, affordable, geo-referenced, concise, reliable source of data available to them. On the other hand, scholars have identified census as an "*inherently impoverished source of data, which is limited by coarse and arbitrary zonal boundaries*" [33]. Ultimately, 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" and rule-based analytical techniques. To start with, Cambridge Systematics [17] prepared detailed guidelines to the best practices in TAZ disaggregation. The document included a set of specific guiding "rules of thumb", which are basically a reasonability checklist recommended for TAZ delineation. Further, the document provided illustrative procedures for splitting TAZs and distributing socioeconomic data.

Literature on TAZ design for travel demand modeling purpose is scarce and sparse. And mainstream transportation modeling literature is largely silent on TAZ design issues and procedures [33]. At best, TAZ selection and design are addressed in many transportation planning publications and textbooks as a set of criteria or "rules of thumb." However, to date, there is no formal method for guiding the design of TAZs. Table **1** on the next page summarizes these design criteria or "rules of thumb" and corresponding studies.

Review from Table **1** (on the next page) reveals the following findings:

- GIS software packages (e.g., TransCAD or EMME-3) used to develop their models does not include any TAZ design module and TAZ file is added as exogenous input data.

- Literature review reveals that TAZ design issues have been overlooked by most transportation community and further it is challenged by lacking a well-defined concept.

- Each study included a defined list of criteria for TAZ structure design, which are set by researchers and practitioners to fit their modeling needs, but not all of these criteria are technically justified, and some of them are included merely to facilitate data processing.

- To-date, both the-state-of-the-art and the-state-of-the-practice haven't provide a solid methodology to rationalize TAZ design process in an analytical framework, which can satisfy various modeling needs.

- It is almost impossible for a single TAZ design to fulfill all criteria due to the complexities resulting from the reality that some of these constraints conflict with each other.

- The impact of each TAZ design criterion onto modeling accuracy has not fully understood or has been largely ignored in most studies.

- Large rural zones are mostly overlooked in TAZ design presented in the above studies. Special attention to rural zones is motivated by the fact that they highly influence the modeling accuracy. Indeed, rural zones tend to contain 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). Thus, further research is highly recommended to consider large rural zones in TAZ design process.

- The complexity and ambiguity surrounding TAZ definition and selection justify the call for further research in the field.

In the real world modeling process, transportation planners conducting various travel demand modeling activities require disaggregating/aggregating socioeconomic data from the census available for varying geographies such as dissemination blocks (DBs), dissemination areas (DAs) and census tracts (CTs) to customized geographic areas-TAZs to best fit their models. This is another influential aspect in the spatial analysis related to travel demand models and is equally challenging as TAZ design. Further, a significant aspect in data transformation between two geographical levels at varying scale (for example, from a census geographical unit- census tracts to TAZ) is to ensure that the two zonal systems boundaries are compatible or congruent. In this case, the data transfer is utterly an aggregation procedure as discussed before. Unfortunately, in most cases, data transfer is between incompatible (incongruent) zonal systems and under these circumstances, data transfer requires more sophisticated techniques than just aggregating and adding fully contained TAZs. Thus, modelers need to build techniques to assign data between these varying spatial units or to transfer data to newly generated TAZs. In general, areal interpolation techniques have been used for data processing between two or more incompatible TAZ systems.

Table 1.   TAZ Selection Criterion and their Impacts on 4-Step Model

TAZ Design Criteria	Criteria Description	Criteria Impact on 4-Step Model	References
Homogeneity	Similar land-use and socioeconomic characteristics within TAZs.	Minimize intrazonal trips. Improve trip generation estimation.	[17,20,32,34-41]
Contiguity	Spatially connected TAZs.	Well defined TAZ for Spatial dependence and facilitating data interpolation. Improve O-D estimates.	[35-39,42-44,46]
Convexity and compactness	Compact and well defined regular TAZ shape.	Improve O-D estimate. Irregular TAZ shape might result in unrealistic "lumping" assignment of trips to road network due to long centroid connectors.	[35-37,41,43-45]
Exclusiveness	Avoid zones (islands or doughnuts) within TAZ.	Avoid calculation problems in 4-steps travel demand models.	[34,35,37,42-44,46]
Boundary compatibility with census and statistical boundaries	Nesting census geography within TAZ boundaries.	Rich source for readily available census data and facilitate data transfer. Improve generation /production, O/D estimation and overall modeling results	[32,34-37,40-46]
Boundary compatibility with physical features	Respect of physical geographic separators placed on territory such as railways, rivers etc.	Impose travel barriers and thus impact trip distribution and traffic assignments.	[34-36,40,42-44,46]
Boundary Compatibility with network geography	Respect roadway network details (number of links/TAZ and functional class composition.	Ensure balanced assignment of traffic and improve traffic assignment capability and accuracy.	[17,40,47]
Equal trip generations per TAZ and setting up maximum threshold	Equal number of trip generation (productions & attractions) per TAZ.	Prevent overloading road links in TAZ with high trip generations and improve traffic assignment.	[34-37,42-44,47]
TAZ numbers	Modeler's choice. Selected based on modeling need to result geographically balanced data. Consider maximum threshold.	Improve trip generation and distribution.	[32,35,36,40-42,48]

As revealed from the literature review, traditional travel demand modeling over the years has been focusing more on the process itself while paying little attention or even ignoring the aspects of TAZ definition, configuration and design. Therefore, modelers cannot find formal methods for guiding the design of the zone system and characteristics of those zones must be decided on an empirical basis in every specific modeling situation [49]. In general, TAZ structure characterized by the number and size of zones inside the study area is closely connected to the analysis level that modelers try to achieve. Further, it is important to understand that the quality of a zoning solution is determined by the model objectives. Clearly, TAZ design based on fulfilling a set of multiple criteria (or optimizing a set of constraints) has been proven to be a challenging task due to the complexities involved. For example, trying to achieve zonal homogeneity will most likely conflict with meeting zonal contiguity requirement. Furthermore, TAZ with homogeneous land use activity may conflict with its shape compactness. Ultimately, the whole TAZ design exercise becomes a question of dealing with trade-offs between conflicting constraints to reach a reasonable and feasible solution to fulfill the modeling needs.

Formulating a multi-objective aggregation or disaggregation criteria is difficult to achieve in practice, because the number of possible TAZ structure configuration is infinite. On the other hand, "rule-of-thumb"-based TAZ selection process was criticized for loosely reflecting fundamental spatial analytic principles and not suggesting an objective, scientific procedure [50]. Thus, due to the complexity of the problem, several heuristic approaches have been tried to provide solution to the TAZ configuration problems. Today, more advanced methods attempting to rationalize the design process in an analytical framework are being developed using spatial and statistical techniques. In particular, TAZ design analytical techniques have been motivated by advances in GIS, computer speed and storage capabilities, and computing algorithms. Consequently, a number of clustering algorithms have been deployed in GIS environment using a different set of constraint and objective functions to achieve an "optimal" TAZ design. In particular, O'Neill [36], You *et al*. [39], Ding [43], and Choi and Kim [38] attempted to solve the problems associated with the "modifiable" or arbitrary nature of the spatial zoning system. Unfortunately, the spatial analytic issues have been selectively recognized by the transportation researchers, who have attempted the GIS-based analytic solution for optimal

TAZs design. Specifically, the modifiable areal unit problem (MAUP) have not received much attention among both transportation researchers and practitioners. Evidently, MAUP arises when the spatial zoning system used to collect geographic data is arbitrary and consequently, the results based on that system can be arbitrary. To this end, the pervious discussion related to TAZ design reveals that the issue has been overlooked by transportation community and further it is challenged by lacking a well-defined concept, a solid methodology and a sound spatial analytical process. In addition, ambiguity surrounding TAZ design is related to defining a same TAZ structure for a variety of modeling goals, because quite often zones designed for one purpose may not be suitable for another [51].

To date, the issue of TAZ design for travel demand modeling applications is still an open area for further research. What definitely needed is the introduction of new methodological approach for zoning scheme, which is based on an algorithm-based delineation process defined to fulfill a set of constraints (objective function) and to achieve multi modeling goals. Equally important, TAZ design step should be dealt with as an integral part of the traditional modeling framework rather than an exogenous data input module.

HFTDM with Advanced Technology and "Intelligent" Data Sources

As previously mentioned, 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 such a problem 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. This section will briefly discuss how GIS as a data-processing tool and remotely sensed images as data-collection technology can be used to in develop high-fidelity travel demand models (HFTDM). Detailed discussion of both GIS and RS along with other intelligent data sources such as parcel-based tax-assessment data, and point-based postal data in developing HFTDM for improved modeling accuracy will be included in the coming section "Opportunities and Challenges".

GIS in Relation to HFTDM

The introduction of Geographic Information System (GIS) into the traditional travel demand modeling has been acknowledged as being the most significant enhancement to advance the practices in this field. Generally speaking, the basic practice of travel demand forecasting has changed little since the days of Urban Transportation Planning System (UTPS) (early 70's). The most significant advances have been in computer technology, software enhancements such as improved graphical displays and geographic information system (GIS) [12]. Further, Sutton [52] attributes the adaptation of GIS in transportation (GIS-T) to the requirements of the Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA), the Clean Air Act Amendments of 1990, and legislation by states that mandate the development of transportation models/programs to reduce traffic impacts.

Further, NCHRP 365 [21] recognizes GIS as a significant development of travel modeling field. The manual sums up the role of GIS in the modeling process as follows: (1) prepare, map and batch roadway network at detail needed for the model, (2) map demographic data at a census block/tract level, (3) convert census blocks/tracts to traffic analysis zones (TAZs), and (4) export TAZ structure to form a demarcation file for use in the model. It is worth mentioning at this point that, in travel demand model, spatial data are identified, stored, processed and presented in GIS environment as points, lines, and polygons. A list of traditional modeling activities facilitated by GIS in the reviewed literature is included in Table **2**. Table **2** lists a variety of GIS spatial analysis tools, which can facilitate data preparation, building TAZ and roadway network elements. Traditionally, TDM starts with data perpetration and GIS is used to store, retrieve, analyze, modify, import, export and display data required or produced by travel demand models. GIS provides powerful, graphical, data base query capabilities that are useful in testing various attributes and their impact on demand modeling activities. Further, GIS is proven to be an efficient tool in defining and designing TAZs by applying various spatial functions such as clip, merge, selection, overlay, split, districting, aggregate etc. For example, GIS can be used to achieve a multiple map layer overlay, which helps in detecting and identifying TAZ and census boundaries. Also, map editing has been used widely to redefine TAZ boundaries. Finally, GIS provides a systematic and consistent way to build, code, update, and check connectivity of roadway network.

Remote Sensing (RS) Technologies in Relation to HFTDM

Spatial misrepresentation of demographic and economic data from census limits the accuracy of traditional travel demand models and their network-wide traffic estimation capabilities. Thus, these limitations have led transportation agencies and communities to continuously seek for new modeling tools and data sources to improve the static and "confidential" nature of census data. Remote sensing technology has been emerged as a powerful "intelligent" data source, which has the potential to resolve some of the aforementioned issues from census data. Certainly, this has been motivated by the following benefits of using remote-sensing for data collection in transportation planning [54]: (1) flexibility in collecting data from places that are costly or impossible to reach, (2) it is non-intrusive technique, and (3) wide-area coverage at improved spatial and spectral accuracies without intensive labor activities. Moreover, Ekern [54] in his efforts to close the knowledge gap between transportation and remote-sensing communities suggests that the two communities should become knowledgeable to the needs and products of each other. Indeed, closing the gap is needed nowadays more than ever with the recent advent of high-Resolution Satellite Imagery (HRSI) such as IKONOS satellite from Space Imaging and QuickBird satellite from Digital Global. Truly, high-resolution images (mostly better than 1 m spatial resolution) contain rich, detailed, accurate spatial information, in a timely and a relatively less-costly manner about land use/land cover (classified land use, impervious surfaces, and vegetation) and urban features (roads and buildings). In this regard, Usher [55] urges

Table 2.   Literature with Modeling Activates Facilitated by GIS Functionalities

Modeling Activity	Possible GIS Functionalities Employed	Reviewed Literature
Designing TAZ structure	Clip, Merge, Selection, Overlay, Import, Export, Split, Districting, Aggregate, Calculate Area, and Measure Distance.	[32,36,37,39,43]
Building Model Network	Ckeck Line Layer Connectivity, Set One-way Segment, Clip, Distance Matrix, Selection, Overly, Buffer, Create Route, Dissolve, Dataview, and Modify Table.	[20,45,53]
Census Socioeconomic Land-use Data Preparation	Dataview, Join Dataview, Modify Table, Selection, Overlay, Spatial Join, Import, and Export.	[20,42,43,45]

transportation organizations to make changes to their current planning and operation approaches by employing new tools for storing, retrieving, visually representing, manipulating and analyzing data.

In short, this section presents remote-sensing as an enabling tool and data collection technology which have a significant potential to empower and improve traditional travel demand modeling. Table **3** summarizes literature using remote-sensing for enhancing various traditional modeling activities. Clearly, Table **3** shows remote sensing as an opportunity to reduce efforts in field data collection and surveys activities, which are laborious, costly and time consuming and to provide a rich source of data for various travel demand modeling activities. First of all, it is essential to precisely specify data items that are to be collected from remotely sensed images. For Land-Use/Land-Cover (LULC) classification, the focus is on LULC classification categories and functional classification. The ultimate challenge to using remote sensing technology in transportation planning field is related to the lack of standardization in LULC classification and the presence of a range of ways in which standard image classification techniques are being developed/used in urban transportation applications. Further, the quality of classifications may become highly dependent on effective integration with exogenous ancillary information such as land use or digital property maps. No doubt, improving LULC classification will provide an additional powerful data source for land- use-based TAZ design.

For building and extracting roadway network, Table **3** reveals numerous efforts by remote sensing communities. Due to the fact that on one hand, buildings accommodate human beings activities, and the information of buildings, such as size, height, geometric shape and orientation, can be used to roughly estimate the location, scale and even purpose of various human activities, on the other hand, all of them are useful for land use classification and high-fidelity travel demand modeling application. Further, accurate extraction of roadway network centerlines/boundaries is considered to be an opportunity for validating, updating and conducting various accuracy checks. It should be emphasized that one possible reason of traffic estimation error is partly attributed to the differences between the assumed/modeled roadway network and the actual one in-place. Finally, feature extraction is challenged by misclassification due to feature complexities and heterogeneity of urban surfaces, for example, roof tops *vs* street pavement, shades from trees, noise from cars etc.

In our opinion, literature review reveals that research efforts to improve four-step model in general has not been

focused on the "high-fidelity" (including both fine zones and detailed roadway network) concept and fall short in the following aspects: (1) creating fine zones by disaggregating census data from geographically defined census area, e.g., dissemination area (DA) or census tract (CT) level to even higher spatial resolution (e.g., blocks or even parcels) through integrating GIS/RS/tax assessment data, (2) utilizing data source with fine spatial unit (e.g., dissemination blocks, tax assessment digital property maps, and postal code point data) which has the potential of forming a new geographic building block for improved TAZ structure, (3) emphasizing the need for fine-grained, detailed, and up-to-date socioeconomic information that go beyond census by employing a variety of innovative "intelligent" tools and data sources, (4) utilizing remote sensing technology for transportation molding applications based on potential benefits from remotely sensed data, (5) dealing with large rural zones to ensure an effective distribution of the assigned traffic over the entire roadway network including low-class roads, and (6) defining a set of guidelines for the optimal TAZ structure/ network details to meet a specific modeling goal.

As part of a remedy effort for the previously mentioned limitations of traditional travel demand models, the following sections of this research paper will highlight the potential challenges and opportunities in relation to developing HFTDM based on what have been revealed from the above extensive literature review and related conclusions.

CHALLENGES/OPPORTUNITIES FOR IMPROVING TFTDM ESTIMATION

Literature review from this paper reveals the challenges facing researchers and practitioners in developing the proposed high-fidelity travel demand model for a variety of improved decision-makings in transportation planning field. In the following paragraphs, we will discuss challenges identified in the review and they are as follows:

Challenge 1: Census data issues -aggregation, confidentiality, and assumptions. Census data are heavily employed in developing traditional travel demand models. Unfortunately, census agencies in various countries collect data at individual and household levels subsequently aggregate them at a specific geographic level for reasons of confidentiality (non-disclosure). Openshaw *et al.* [76] identified the following three implications from the aggregation process of census data due to the privacy issue: (1) the loss of precision of dissimilar personal records for the same geography, (2) the loss of considerable uniqueness and details, and (3) major reduction in multivariate

Table 3.   Summary of Remote Sensing Research Activities in Relation to Transportation Planning

Research Focus	References	Modeling Application	Anticipated Problems
Road-network Extraction	[56-60]	Update existing road networks, check road networks connectivity, check TAZ boundaries and centroid connectors.	Local-street misclassification due to shadows using HRI and asphalt paved roofs. Rural roads weak boundaries.
Building Extraction	[61-65]	Validate and improve census data, extract building footprints for reducing rural zones size, improve land use classification for urban areas for TAZ design and selection.	Shades by tall buildings or trees. Different materials for same feature result in classification errors.
Land-Use/Land-Cover Classification (LULC)	[66-69]	Functional classification used in TAZ design, place of work employment data, trip generation, and trip distribution validation, and centroid location.	Relationship between LU and LC is not direct and sometimes is ambiguous. Confusion and non uniformity in classification standards. Complexity and diversity of urban structures and materials.
Socio-economic Data Population Density Residential Density	[70-73,75]	Census data validation and trip generation estimation.	Accuracy depends on classification results and study area scale. Difficulties in determining buildings functional use and problems in covering large area.

dimensionality. No doubt, personal data collected then aggregated are changed dramatically as individual characteristics recorded with immense details are aggregated to become frequency counts with the loss of linkage between different variables. For these reasons, Statistics Canada (SCAN) asked users of census data to be aware of data and area suppression issues, so that they can assess the usefulness of census data for their purposes and the risks involved in basing conclusions or decisions on these data.

Unfortunately, this aggregation is accompanied with troubles for travel demand modeling analysis. Specifically, as a result of this forced aggregation, two kinds of biases arise in the modeling analysis: (1) from using areally grouped data for model variables (income, age, auto ownership, etc....), and (2) from using virtual geographic centroid to represent the location of human activities (dwelling units and employment centers) rather than using the real spatial allocation of individual human activity in each areal unit [77]. More to the point, working with aggregated objects at coarse zones is basically assuming all objects are treated similarly, and all intrazonal processes are ignored since they do not appear on any transportation network and can be considered as a problem of missing data [78]. Thus, ignoring intrazonal trips will have a major impact on modeling accuracy and network-wide traffic assignability. Equally important, census data are prepared for purposes other than transportation planning and therefore, most of the time modelers and transportation planners have to manipulate them to fit their modeling needs. Generally speaking, with an aim of the development of HFTDMs, manipulating census data is carried out by transforming them from a state of aggregation over coarse geography to a state of disaggregation over fine geography. The main challenge (among others) in the whole transformation process is defining the required level of spatial disaggregation to achieve the required number, shape, and size of zones to fulfill the model goals. Ultimately, the final choice is a compromise between available resources and modeling needs. Further, census data suffer from the assumption that socioeconomic attributes are uniformly distributed across areal polygons and trips are generated from a hypothetical or virtual center called "area centroid". Thus, the intensity and

distribution of population and their socioeconomic activities over space within a polygon-TAZ are missing from census data. Certainly, the aforementioned census data issues will impose a challenge when they are used at the level of geography required in order to meet the needs of travel demand modeling.

Challenge 2: Lack of longitudinal and timely data. Accurate travel demand forecasting requires data that are representative of the current conditions, which should be continuously updated as well. Unfortunately, transportation community continues relying on the census partly because they often lack the resources needed to continuously update socio-economic and demographic input data for their travel demand models. To date, census data are locked not only in spatial details provided to the public, but also in the temporal depth. Data extracted from the census is like a snap shot taken every ten years (US and UK), or every five years (Canada and Australia). Ford and Fricker [79] recognized the need for socioeconomic and demographic data that are complete, current, representative and appropriate for the applications of travel demand models. Notably, full reliance on census data would leave us to a "no-data" situation for those periods between the two censuses, which can only be manipulated or estimated by planners using growth rate and trend analysis. To address this issue the authors called for continuous "real-time" data collection efforts. Further, maintaining an up-to-date inventory for the periods between censuses is one of the thorniest problems in data collection efforts [80]. Thus, Goulias [80] called for year-to-year update to attain what he called "fresh" data in order to move data inventory to an accurate and more complete picture. Certainly, using out-dated census data in the development of travel demand model involves a much higher uncertainty, and therefore, the challenge is to find an affordable way to update data and supplement the census.

Challenge 3: Heterogeneous land-use patterns. Heterogeneity is basically reflected by the degree to which residential units and employment centers (housing/job mix) are located in the same areal unit or the same TAZ. Although land use heterogeneity is promoted as one of the smart growth principles, it is discouraged in the spatial analysis for

travel demand modeling applications. And this is due to the fact that heterogeneous zonal fabric characterized by mixed land uses will result in missing intrazonal "short" trips while capturing only interzonal trips and consequently, reducing estimation accuracy of travel demand models. Therefore, improving traditional TDM implies creating homogeneous TAZ structure in terms of size, shape and land use purpose of the created TAZs. However, defining homogeneity based on a set of constraints and variables such as, population, employment, socio-economic factors and land-use characteristics, is a challenging task. Further, modelers are challenged with a dynamic heterogeneous system within each spatial unit irrespective of its level of geographic hierarchy, which also changes substantially according to the temporal scale used. Finally, developing a homogeneous TAZ structure using the finest levels of spatial geography (i.e., parcel and block levels) with heterogeneous land uses requires establishing a complicated optimization question and corresponding solution.

Challenge 4: Lack of data integration and standardization. Despite the widespread use of census socioeconomic data in travel demand modeling, no consensus exits as to which variables should be used or at which level of spatial geography. Data collection efforts come from a variety of organizations such as governmental (federal, provincial, and local), commercial, and non-profit/educational institutions. Further, several data types exist (census and non-census), each with different variables, geographic levels, and updating frequency. These data have to be integrated before being used in modeling analysis. Retrieving and integrating geo-spatial information from multi-database systems and from various sources is a critical, non-trivial issue. For example, when integrating GIS-based roadway network from vector layer and another one extracted from satellite imagery, we need to ensure that roads from the two layers line up and have the same geo-referencing details.

Challenge 5: Place of work (POW) data. Information on the origins and destinations of trips is crucial to travel demand modeling. So that, poor location data may introduce misleading estimates. Unfortunately, Place-of-Work (POW) data provided by census agencies such as Statistics Canada (SCAN) is problematic for two reasons: First, POW data at DA level is population-based and does not reflect the real number of employment by industry at the specified geographic area they represent. Therefore, there is uncertainty surrounding employment data, which will result in misleading model estimation. Today, researchers were required to purchase special tabulations on employment-based POW data at DA level. Second, POW data usually are not provided at finer levels of spatial geography than DA due to confidentiality issue. Accordingly, transforming POW data between various levels of spatial geography may introduce errors due to geo-coding inconsistencies and scale problem. In this regard, it should be noted that Statistics Canada (SCAN) recently released POW employment data at the census tract (CT) level (bigger than DA), for all tracted Census Metropolitan Areas (CMAs) and Census Agglomerations (CAs) in Canada, for the years 2001 and 2006. In 2006 Census, SCAN tabulated POW employment data at CT Level for 48 urban areas (33 CMAs and 15 CAs) and 5,076 CTs. CT level of spatial geography may not be a

good TAZ structure for urban and local modeling purposes, especially in the context of developing HFTDMs. Finally, census agencies, e.g., SCAN suppresses data for POW population due to confidentiality issues. The impact of data suppression on modeling accuracy needs to be further investigated.

Challenge 6: Spatial resolution of rural area. Developing an improved TDM require dealing with rural zones. In the context of North America they are large, irregular in shape with relatively low land use density, which is mainly residential in type and scattered in distribution. Further, rural zones usually have much sparse roadway network (or even non-existent) when compared to their urban counterparts. This characterization of rural large zones tends to generate high portion of intrazonal trips, which are mostly unaccounted for in the traffic assignment process. However, the issue of rural zones has been long challenging to transportation planners and it appears such a problem is caused by the fact that they are designing TAZs irrespective of the level of spatial geography they use. Indeed, large rural TAZs are in existence at all levels of census geography, including DA and DB levels. Generally speaking, rural DAs tend to be more homogenous in terms of the population they contain, but not their size, leading to less accurate modeling results. Also, as mentioned before, postal codes suffer from significant variation in geographical extent, which could range from a single building in a CBD area to hundreds of square miles in rural areas (e.g., average size of a ZIP code area in the State of Wyoming is 1,430 km²). Clearly, this tends to make their use for fine spatial analysis misleading and this is generally true for rural areas and even for some urban areas. Additionally, rural land-use categories, which tend to be more homogenous, but occupied only at scattered areas, may be lost because their scale of resolution is so low that their presence is underestimated (e.g. census undercounting problem in rural areas). On the other hand, land-use types occupying much less area and tend to be concentrated and contiguous their presence is overestimated because their scale of resolution is so high. With all that being said, there is a need to disaggregate rural zones by reducing their spatial resolution for creating a homogeneous, compact zonal structure in order to improve modeling accuracy. Despite the growing recognition of the problem, no methodology has been developed as of yet to deal with this issue. Further research integrating GIS and remote-sensing technologies is recommended to disaggregate large rural zones for improved modeling accuracy.

Challenge 7: Roadway network (RN) representation. Clearly, socioeconomic data and roadway network play a significant role in modeling accuracies. Three issues in relation to roadway network are of concern to traditional modeling activities:

a)   RN should be up-to-date, with high connectivity and detailed attributes. Any model validation and sensitivity test must begin with a good network design and coding procedure in place. Without this solid basis for modeling, the best validation procedures and efforts are exercised in futility.

b)   RN should be compiled from various sources for accuracy.

c)   Deciding on the detail level of roadway network needed to match the developed TAZ structure is another challenging question need to be answered in developing HFTDM. The interaction between the zoning structure/spatial resolution and network details has been mostly overlooked in both the states of art and practice. And the recommendation always has been to use a more detailed roadway network with fine TAZ structure and a less detailed one for coarse TAZ structure, but not quantitative guidelines are available.

Basically, detailed, well-connected network containing streets of diverse sizes and functions will facilitate dispersing traffic by offering differing routes to a variety of destinations. Consequently, this will improve network-wide traffic assignment, including low-class local roads. Finally, A Peer Exchange on Travel Model Validation Practices conducted by Travel Model Improvement Program (TMIP) and sponsored by the Federal Highway Administration (FHWA) recommends testing network topology, including balancing between roadway network details and zone details among a number of primary and secondary model validation tests [81]. However, the document does not provide any details regarding the specific test procedures or minimum standards for the tests. Thus, further research is required to quantify the match between the two for various modeling scopes to achieve improved network-wide modeling estimations.

Challenge 8: Lack of incorporating spatial analytical issues. Transportation community always considered travel as a derived demand for personal activities that are distributed in space and time [41]. Therefore, space is an integral part of transportation analysis, which cannot be overlooked. The essence of traditional travel demand modeling appears to be analyzing spatially aggregated census-based data defined by coarse and arbitrary zonal boundaries. As mentioned before, census geographies do not fit modeling purposes in most cases they have to be reconstructed under a set of modeling criteria in order to achieve a set of modeling goals. Although, the focus of designing "optimal" TAZ structures based on the census geographies have been enhanced by advancements in GIS and spatial analysis techniques, unfortunately, this opportunity is accompanied by added troubles resulted from overlooking the sensitivities of disaggregating aggregated census data to the ideal areal units or final TAZs used. The truth of the matter, travel demand modeling involves analyzing aggregated spatial data and modeling spatial geographies and both are associated with spatial analytical issues, which can substantially affect modeling results.

As discussed earlier in this review, modifiable area units problem (MAUP) arise when changing study area boundaries to produce a variety of spatial units, which will definitely have a significant influence on the modeling results. Yet, ecological fallacy (EF) occurs when individual characteristics recorded in immense details are aggregated to become frequency counts with the loss of linkage between different variables [76]. EF is related to reporting census data with the assumption of homogeneity and uniform distribution over geographical units. In travel demand modeling context, it is generally believed that MAUP and EF

effects could be mitigated by disaggregating data over high spatial resolution geography [50]. Unfortunately, to date, there is no predefined method to avoid the spatial analytical pitfalls related to MAUP/EF problems and the uncertainties induced by spatial interpolation. As long as traditional travel demand models are limited by coarse and arbitrary zonal boundaries and continue to rely on aggregated census data, those spatial/statistical issues should be recognized and accounted for as part of our efforts to the model improvement scheme.

The aforementioned impediments of traditional modeling framework should not be considered fatal flaws but challenges to overcome. This section aims to identify a series of opportunities that, if implemented, would enhance traditional travel demand models and help in development of HFTDM. They are as follows:

Opportunity 1: GIS Software. Remarkably, we have entered an era in which it is possible to take advantage of the explosion of digital geographical information, which could only be managed through GIS technology. Evidently, GIS has led the way to the growth of techniques for improving various commercial transportation travel demand forecasting software. For example, Caliper's TansCAD software package is a fully integrated GIS and transportation analysis tool. This software combines GIS and transportation modeling capabilities in a single integrated platform. EMME-2, Cube Voyager and VISUM are other similar software packages. Further, in recent years, there has been an enormous increase in the availability of open-source software, which is being widely adapted by research communities. Today, the world of GIS goes from analyzing static geographic and census data to conducting various spatial analyses over dynamic data in real-time, which in turn provide vital information to planners and decision makers.

Opportunity 2: Remote-Sensing (RS). The widely used four-step travel demand model was mainly criticized as being a static one ignoring land-use/transportation interaction. Capturing this real-world interaction in travel demand analysis is regarded as an opportunity to improve the precision and estimation accuracy of traditional travel demand modeling. Today, various kinds of remote sensing (RS) systems and sensors are utilized for capturing changes in land use and transportation system. High-resolution images from multiple sensors (such as Optical, LiDAR and Radar) contain rich source of spatial data, which can support various traditional travel demand forecasting and modeling activities, such as the design of TAZs and disaggregating census data. Accordingly, information extracted from these images provides the hints about where people live and work within the biophysical context. Furthermore, it provides a measure associated with human activities such as impervious surfaces, building footprints, and land use/land cover classification. In addition, nowadays, remotely sensed data are covering nearly every corner of the earth with increasing spatial and temporal resolution, which has a huge potential for providing traditional travel demand modeling with fine-grained high-resolution socioeconomic data for enhanced modeling experiences and accuracy.

Intuitively, detailed classification of the urban landscape is required for developing HFTDMs, which has been more

and more accessible due to the increasing availability of high-to-very-high-resolution images, advancements in rule-based/object-based image classification algorithms, and availability of fine-resolution census geography (dissemination block level). Consequently, detailed land-use classification, roadway network and socioeconomic attributes can be visually identified and finally extracted by computers at fine levels of geography. This opportunity is extremely vital for improving traditional TDM in the following areas: (1) achieving fine TAZ structure with homogenous land-use, and (2) disaggregating socio-economic data (population, households, and employment) from census based on classification and quantification of building footprints. In addition, keeping roadway network updated in a detailed and timely manner is crucial for developing HFTDM in that (1) it will ensure high-resolution traffic assignment can be carried out over a network including roads from- all functional classes, and at the same time, (2) it will eliminate our reliance on costly, time-consuming and labor intensive ground surveying (traditional or automatic) techniques for updating roadway network.

Opportunity 3: Computer advancements. Recent advancements in computer hardware and software technologies make it possible to perform modeling tasks that have been treated as impediments in the past decade or so. Faust *et al.* [82] recognize the lack in computer capabilities as the reason for preventing the integration of GIS and RS. It appears that improvement efforts of traditional travel demand models focus on disaggregating space and integrating "fine" data into the modeling process, and this, in turn, requires performing analytical tasks, which are becoming very processor and memory hungry. Fortunately, because of the exponential increase in computing speed and memory space, the computational time required related to extensive detailed roadway network or to disaggregate space to the "roof top" is no longer a limitation. With these tremendous computing capabilities, "data hungry" urban models characterized by Lee back in 1973 [7] are no more of a concern. Furthermore, the accelerated increase in the size of computer memory and processor speed will facilitate innovative system designs that manage both optimize speed and accuracy in handling large amounts of geo-referenced transportation data and sizable roadway networks [83].

Opportunity 4: Remote-Sensing/GIS Integration. Geospatial modeling and data handling functionalities from GIS and cost-effective data collection and timely object detection/extraction capabilities from RS call for efforts from both GIS and RS communities to integrate both technologies for various applications, including those in transportation area. Two data issues ultimately have warranted and led to the integration of both GIS and RS: (1) the need for combining data from the two sources (GIS static and RS dynamic data), and (2) the need to utilize unique functionalities of both technologies: GIS -managing, processing and visualizing of data and RS -object-detection and timely collection of data at a large scale. Therefore, the integration is based on a mutual interest of both communities for data enhancements to support policy and decision making efforts. Wilkinson [84] conceptualized the integration between remote sensing and GIS technologies as follows: First, remote sensing can be used as a tool for gathering data to be analyzed in GIS. Second, GIS data can be used as

auxiliary information to improve image processing. Third, remote sensing and GIS can be used together for modeling and analysis. Afterwards, Weng [85] elaborated on Wilkinson's ways of integration by relating each one of them into a specific modeling application. First, remote-sensing images have been used in extracting and updating roadway networks, providing land-use/cover data and attributes of urban features and detecting urban expansion. Second, census data have been used to improve image classification in urban areas. Third, the integration of remote sensing and census data has been applied to estimate population and residential density, to assess socioeconomic conditions and to evaluate the quality of life.

Opportunity 5: Internet-based mapping services. Advancements in the internet and satellite communication have paved the way for various internet mapping programs, which offers users an interactive mapping of the globe covered with high-resolution imagery, streets, and points of interest. They start changing the way we see and model the world by providing users with freely available, easy to use, rich and dynamic geographic data. Evidently, a variety of internet mapping services could potentially provide useful data for developing HFTDM. For example, Google Earth (GE) offers free world-level coverage of high-resolution images with capabilities of downloading them using Google Satellite Downloader 6.7. Then, downloaded GE images can be exported to GIS-based software such as ArcGIS or TansCAD for further analysis. Moreover, users can add geospatial data to Google Earth KML (Keyhole Markup Language) files by converting Shapefiles to KLM files. More importantly, these image can be used to extract rich land cover/land use information and travel demand models can overlay such information with differing GIS vector layers such as study area at both standard census geography and disaggregated levels, road network, centroids and centroid connectors, digital property maps (DPM), land-use maps, and classified images, for various modeling tasks. This overlay of diverse vector layers will enable modelers to perform analysis with a clear picture of the whole study area and details of related modeling components.

Opportunity 6: Parcel-based tax assessor data. Human activities are directly related to land-use signature. For example, residential and employment land-uses define where people live, and work respectively. The rest of land-use classes define all other activities such as recreational, commercial, etc. Tax assessment parcel data and tax maps accurately reflect size, shape, and location, of each parcel of land in a municipality. Moreover, parcel-based tax assessor database uniquely identifies each parcel by parcel identifier number (PIN), property account number (PAN), property description (residential, commercial, vacant, etc.), property type code, and property assessment value. Further, parcels are graphically displayed in digital property maps (DPMs), which defines the property lines for each parcel. Schuurman *et al.* [33] recognize that property (cadastral) data include a seamless spatial fabric, which makes it possible to assign socioeconomic attributes at the rooftop (household) level. Eventually, using cadastral data as a framework for socioeconomic analysis is fulfilling the call for the need for a data-rich GIS-based model building basis. Thus, data at the cadastral level creates the possibility of conducting socioeconomic analysis at the household level. Yet, many of

the variables used for socioeconomic analysis, such as income, are not associated with the cadastre. However, information regarding dwelling values and tenure type is usually assigned to each legal parcel. Further, Cambridge Systematics [17] proposed to use the parcel tax information in a digital format to delineate TAZs based on clusters of similar land uses. It appears that this is definitely a new opportunity for improving traditional four-step TDM by employing property attributes at parcel level in the modeling process.

Opportunity 7: Point-based postal code data. Various postal agencies keep postal or zip code for delivery purposes. This intelligent data source is also regarded as an opportunity, which has not been utilized for improving traditional travel demand model. To begin with, this opportunity was initially captured by researchers in the health community. They considered ZIP code data an interesting and appealing alternative to commonly used area or zonal-based counterpart [86]. This has been suggested for many reasons, including that ZIP code data is relatively easy to collect at various locations, point of sales; and ZIP code areas are often spatially finer than census geographic areas offering a unique opportunity for analysis at a finer level of spatial geographies. Finally, ZIP codes are roadway network-based geographies, which strongly correspond to the spatial location of individual, families, and business and definitely provide insight to the location of human activities. In addition, with postal code data, one can map urban activities as individual points along roadways with a high precision, which is a more realistic way than the traditional areal or zonal-based representation. Further, these days, increased computing power, advancements in matching algorithm and powerful geo-coding tools in various commercial GIS software have led to the development of "crosswalk", "relationship" or "conversion" files that associate Postal Service Codes (ZIP or Postal) to any legal and statistical census geography in countries such as US, UK and Canada. For example, SCAN provides Postal Code Conversion File (PCCF), which is a digital file linking the Canada Post Corporation (CPC) six-character postal code to SCAN's standard geography at all levels, including block-face, dissemination block (DB) or dissemination area (DA). Indeed, this linkage can greatly facilitate data integration from varying sources at diverse levels of geography. This will, in turn, much benefit the development of the proposed HFTDM.

POTENTIAL CONTRIBUTIONS FROM HFTDM

The overall accuracy improvement of traditional modeling approach can be coined by the efforts of developing high-fidelity travel demand model (HFTDM), which shows the following features: (1) Fine spatial TAZs, (2) Homogeneous TAZs with single or nearly single land-use based, (3) High-resolution socioeconomic attributes, (4) Detailed (all functional) roadway network, and (5) Realistic representation of centroid and centroid connectors. More to the point, the proposed HFTDM is expected to enhance each of four steps of the traditional travel demand: Trip Generation, Trip Distribution, Modal Split, and Network Assignment. This is very important as the literature review indicates that the four-step travel demand models are still the mainstream planning tool in most developed countries. On the other hand, the state-of-the-art indicated that uncertainty (prediction errors) propagates from one step to another and it is compounded

over the four steps of the traditional modeling process. Further, the modeled traffic volume is expected to have a higher uncertainty than the input data. In practice, the call for fine homogeneous TAZs will result in a closer examination of individual/household as a trip making unit. Consequently, detailed socio-economic household's attributes and their activity location can be identified within each TAZ in a more accurate and realistic way. Equally important, high-resolution socioeconomic data will improve 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 more accurate representation and estimation of trip production and attractions at each zone will allow for a more robust trip distribution analysis.

No doubt, fine TAZ structure will ensure capturing most trips by increasing their likelihood of crossing zonal boundaries, which in turn maximizes the interzonal flow, but by the same token, minimizes the missing trips due to internalization. Further, the use of fine zones will improve the location of trips origins and destinations and shorten the length of centroid connectors. In particular, this will allow for more accurate estimates of travel impedance between zones, which enhances O/D matrices between zones and simply improve trip distribution results from the gravity models. Not to mention, using fine TAZs could provide travel impedances related to short trips produced and attracted within a short distance along local streets. If modelers have information about impedances (distances, travel times, travel speed, etc...) along the local streets, then it is possible to include them in each step of HFTDM. As far as modal split step is concerned, it can benefit from using high-fidelity modeling framework. More precisely, fine/homogeneous spatial zones will provide an opportunity to define and construct the accessibility factor of different modes at very fine spatial scale, land uses, as well as travel impedances by all modes when traveling along detailed roadway network consisting roads from all functional classes.

As previously mentioned, traditional traffic assignments, which are based on coarse zones, virtual centroids and hypothetical centroid connectors, tend to produce very large estimation errors of traffic flows. Obviously, using fine TAZs will result in shorter and less imaginary centroid connectors, which will reduce lumpiness and loadings of trips to centroid connectors. This is because all trips are originated from centroid in one zone and destined at another zones *via* more realistic centroid connectors. Realistic representation of centroid connectors by including local roads and driveways in the network and dividing the coarse zones into fine zones can be easily realized in the proposed HFTDM. After all, centroid connectors are used to truly represent local streets and collectors that connect true trip ends (origins/destinations) to the nearest arterial street system. In fact, increasing levels of zoning homogeneity, spatial resolution and roadway network details will increase the spatial precision in traffic assignment and result in: (1) Capturing higher proportions of interzonal trips which will help lower Percentage Root Mean Square Errors (PRMSE). (2) Lower portion of intrazonal trips, which usually do not cross TAZs in traditional TDM, are mostly unaccounted for and therefore result in higher PRMSE. (3) Effective and balanced distribution of traffic assigned to both

high-class and low-class roads. The proposed high-fidelity modeling framework will perform traffic assignment with detailed roadway network consisting of roads from all functional classes, which is compatible with the fine-zone TAZ structure utilized, with centroids representing trip ends in a realistic manner and with a much better treatment to all four steps as mentioned above.

CONCLUSIONS AND RECOMMENDATIONS

Literature review conducted in this research paper indicated that the traditional, sequential "four-step" modeling process is still holding firm as it is being used by the majority of transportation communities since it has been introduced in the 1950's to evaluate capital-intensive transportation infrastructure projects. In fact, the shortcomings of conventional model originated from such macroscopic scope, which has been recognized as insufficient to consider policy alternatives and incapable of providing accurate network-wide traffic volume estimates. Although the general and structural framework of the model stay the same, it is not adapting and improving with new emerging technologies and intelligent data sources. Today, much of the promise of available technological and data advancements comes from changing the scope of the traditional travel demand modeling framework to higher fidelity (or more disaggregated) one. Practically speaking, this could be achieved by handling the two critical issues that have been hindering the development of high-fidelity travel demand model (HFDTM): (1) relying on census-based socio-economic/land-use data and (2) designing "optimized" TAZ structure/roadway network details to best fit the modeling goals. In general, TAZ design practice among transportation modelers is mostly governed by experiences rather than a systematic approach, which makes it an area open for more research and innovations.

Needless to say, the performance of a HFTDM is highly dependent on the level of TAZ disaggregation. However, the level of disaggregation should be limited to resources available to achieve the model goals. Therefore, modelers should be clear about this point through some sensitivity analysis. With this, testing the relationship between levels of TAZ disaggregation and the overall model accuracy should be explored in the future research. Furthermore, sensitivity analysis relating to TAZ disaggregation with the estimation accuracy of traffic flow on each functional- class of roads and different disaggregation technique should be investigated.

Based on this review and after more than 50 years of traditional modeling practices, it is strongly felt that we shall ponder the following research topics that have arisen based on the development and advancements from new technologies and from the need for developing a more robust, high-fidelity modeling framework to achieve improved network-wide traffic estimates. In this regard, we propose that our transportation community further envisage and investigate the following topics:

➢ Developing GIS-based integrated spatial data infrastructure for travel demand modeling applications.

➢ Introducing advanced systematic analytical solutions to TAZ structure design, based on the advancements in algorithm and computer's capabilities.

➢ Enhancing various areal interpolation techniques for enhancing travel demand modeling accuracy.

➢ Incorporating high-resolution imagery for disaggregating census data for modeling applications, with a special emphasis on large rural zones problem.

➢ Integrating postal code data into TAZ design process.

➢ Linking traditional TDMs to seamless GIS mapping tools and geospatial database from diverse datasets (Google, WikiMapia, MapQuest, digital property maps, parcel-based municipal/county tax data and tax maps).

➢ Understanding the relationship between TAZ structure and roadway network details.

No doubt, the potential of integrating external data knowledge and tools into the traditional modeling procedures to achieve a well-crafted "intelligent" HFTDM is tremendous. This research paper shows that transportation community has recognized but did not fully utilize the above mentioned "intelligent" sources in developing HFTDM. Today, employing intelligent data-collection technologies and data sources such as remotely sensed images, parcel-based digital property maps (DPM) and tax-assessment data, land-use maps, and point-based postal data for improved modeling accuracy can be accomplished completely within a GIS environment. GIS integration with all above data sources and technological tools will adequately capture geographical space and human activities at a much finer resolution, which will lead to a much better representation of all modeling components and result in a much more accurate traffic estimates. Evidently, successful implementation of the high-fidelity modeling framework will provide transportation community with a much better and affordable resource planning tool.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   M.G. McNally and W.W. Recker, "On the Formation of Household Travel/Activity Patterns", Final Report prepared for USDOT, Institute of Transportation Studies, University of California, 1986.

[2]   J. Bates, "History of travel demand modeling" In: Handbook of Transportation Modeling, D. Hensher, and K. Button, Eds. Elsevier Science, Oxford, 2000.

[3]   Cambridge Systematics Inc. Travel Model Validation and Reasonableness Checking Manual. Cambridge Systematics Inc, second edition, 2010.

[4]   FHWA. Calibration and Adjustment of System Planning Models. Federal Highway Administration, U.S. Department of Transportation, Washington, D.C., 1990.

[5]   NCHRP. Travel Model Validation Practices Peer Exchange White Paper Report No. 255. National Cooperative Highway Research Program Research Program (NCHRP), Transportation Research Board, Washington, D.C., 2008.

[6]   R. Mustafa and M. Zhong, "Testing Traffic Assignment Capability of Improved Traditional Travel Demand Model: New Brunswick Case Study", Proceedings of Canadian Society Civil Engineers (CSCE) of the 9th International Transportation Specialty Conference, Edmonton, Alberta, June 69, 2012.

[7]   D. Lee, "Requiem for large-scale models", *J. Am. Plan. Assoc.,* vol. 39 no. 3, pp. 163-178, 1973.

[8]   NCHRP. Synthesis 358, Statewide Travel Forecasting Models. National Cooperative Highway Research Program Research Program (NCHRP), Transportation Research Board, Washington, D.C., 2006.

[9]   A.J. Horowitz, "Guidebook on Statewide Travel Forecasting", Transportation Research Circular (TRC E-C011). Federal Highway Administration, Wahington, D.C., 1999.

[10]   A.J. Horowitz, and D.D. Farmer, "A Critical Review of Statewide Travel Forecasting Practice", Transportation Research Record 1685, Washington D.C, 1999.

[11]   G.T. Giaimo and R. Schiffer, "Statewide Travel Demand Modeling: A Peer Exchange", Transportation Research Circular, No. E-C075. Transportation Research Board of the National Academics, Wahington, D.C., August 2005.

[12]   TRB. Special Report 288, Metropolitan Travel Forecasting: Current Practice and Future Direction. Transportation Research Board (TRB), National Research Council, Washington, D.C, 2007.

[13]   E.J. Miller and M. Hatzopoulou, "Organizational Structures for Regional Travel Demand Modeling" Prepared for: Metrolinks by University of Toronto, 2008.

[14]   M. Ullah, U. Molakatalla, R. Morocoima-Black and A. Mohideen, "Travel Demand Modeling for the Small and Medium Sized MPOs in Illinois" Illinois Center for Transportation, Research Report ICT-11-091, Illinois DOT, 2011.

[15]   Z. Khatib, K. Chang, and Y. Ou, "Impacts of analysis zone structures on modeled statewide traffic" *J. Transport. Eng.,* vol.127, no.1, 2001.

[16]   M. Zhong, and B.L. Hanson, "GIS-based travel demand modeling for estimating traffic on low-class roads", *J. Transport. Plan. Technol.,* vol. 32, no. 5, 2009.

[17]   Cambridge Systematics. A Recommended Approach to Delineating Traffic Analysis Zones in Florida. Prepared for: Federal Highway Administration Report, Washington, D.C, 2007.

[18]   R. Mustafa and M. Zhong, Testing Traffic Assignment Capability of Improved Traditional Travel Demand Model: New Brunswick Case Study, Proceedings of Canadian Society Civil Engineers (CSCE) of the 9th International Transportation Specialty Conference, Edmonton, Alberta, June 6-9, 2012.

[19]   J. Marker and K. Goulias, "Truck Traffic Prediction Using Quick Response Freight Model Under Different Degrees of Geographic Resolution Geographic Information System Application in Pennsylvania", Transportation Research Record 1625. Paper no. 0361-1981, TRB, National Research Council, Washington D.C, 1998, pp. 118-123.

[20]   A.J. Horowitz, "Computational Issues in Increasing Spatial Precision of Traffic Assignment", Transportation Research Record: *J. Transport. Res. Board.,* no 1777, TRB National Research Council, Washington, D.C., 2001, pp. 68-74.

[21]   NCHRP. Travel Estimation Techniques for Urban Planning. National Cooperative Highway Research Program Research Program (NCHRP) Report No. 365. Transportation Research Board, Washington, D.C., 1998.

[22]   D.B. Zimmerman, "Madison County Travel Demand Model. Challenges and Innovations". [Online] Available: http://www. planning.kytc.ky.gov/traffic/traffic_forecasting.asp

[23]   VHB Inc. Determination of the State of the Practice in Metropolitan Area Travel Forecasting: Findings of the Surveys of Metropolitan Planning Organizations. Report to the Transportation Research Board Committee B0090, Washington DC, 2007.

[24]   ODOT. Travel Demand Model Development and Application Guidelines. Oregon Department of Transportation (ODOT). Prepared by Parsons Brinckerhoff Quade & Douglas, Inc, 1994.

[25]   H .Miller, "Potential contributions of spatial analysis to geographic information systems for transportation (GIS-T)", *Geogr. Anal.,* vol. 31, pp. 373-399, 1999.

[26]   J. Duthie, K. Kockelman, V. Valsaraj and B. Zhou, "Applications of integrated models of land use and transport: a comparison of ITLUP and UrabnSim land use models" Presented at the 54 Annual North American Meetings of the Regional Science Association International, November 2007, Georgia.

[27]   P. Waddell, "UrbanSim: modeling urban development for land use, transportation and environmental planning", *J. Am. Plan. Assoc.,* vol. 68, no. 3, pp.297- 314, 2002.

[28]   P. Waddell, L. Wang, B. Charlton and A. Olsen, "Microsimulating parcel-based land use and activity based travel: development of a prototype application in San Francisco", *J.Transp. Land Use.,* vol. 3, no. 2, pp. 65-84, 2010.

[29]   A. Troy, and B. Voigt, "Dynamic transportation and land use modeling" rubenstein school of environment and natural resources. final project report to USDOT federal highway administration (USDOT FHWA Project: DTFH61-06-H-00022), 2009.

[30]   S. Kakaraparthi and K. Kockelman, "An application of urbanism to the Austin, Texas region: integrated- model forecasts for the year 2030", *J. Urban Plan. Develop.,* vol. 137, no. 3, pp. 238-247, 2011.

[31]   N. Holyoak and B. Statzic, "Benefits of Linking Macro-Demand Forecasting Models and Microsimulation Models" *ITE J.,* October, 2009.

[32]   L. Martínez, D.K. Ariane, and J. Viegas, "An integrated application of Zoning for Mobility Analysis and Planning: the case of Paris Region", In the 12th World Conference on Transport Research, Lisbonne: Portugal, 2010.

[33]   N. Schuurman, A. Leszczynski, R. Fiedler, D Grund, and N. Bell, "Building an Integrated Cadastral Fabric for Higher Resolution Socioeconomic Spatial Data Analysis" Progress in Spatial Data Handling (2006): pp. 897-920, 2006.

[34]   N.J. Garber and L.A. Hoel, *Traffic and Highway Engineering*. New York: West Publishing Co., p. 383, 1988.

[35]   K.G. Baass, "Design of zonal systems for aggregate transportation planning models". Transportation Research Record 807 -Travel Demand Forecasting and Data Considerations, vol. 807, pp. 1-6, 1981.

[36]   W.A. O'Neill, "Developing optimal traffic analysis zones using GIS" *ITE J.,* no. 61, pp. 33-36, 1991.

[37]   W.M. Bennion, and W.A O'Neill, "Building transportation analysis zones using geographic information system", *Transport. Res. Rec.,* 1429, pp. 49-56, 1994.

[38]   K. Choi and T.J Kim, "GIS-based traffic analysis zone design", Paper presented at the 7th Wo Conference in Transportation Research, Sydney, Australia, July, 1995.

[39]   J. You, Z. Nedovic-Budic, and T.J Kim, "A GIS-based traffic analysis zone design: technique", *Transport. Plan. Technol.,* vol. 21, pp. 45-68, 1997.

[40]   TMIP Technical Synthesis. Defining Traffic Analysis Zones, 2007. [Online] Available: https://www.tmiponline.org/Clearinghouse/Items/Technical_Synthesis__Defining_ Traffic_Analysis_Zones.aspx

[41]   J. de D. Ortuzar and L.G. Willumsen, Modelling Transport 2nd ed. John Wiley and Sons, 2011.

[42]   C. Ding, "Impact Analysis of Spatial Data Aggregation on Transportation Forecasted Demand: a GIS Approach", In: URISA Proceedings URISA, 1460. Renaissance Drive, Park Ridge, IL 60068, pp. 362-375, 1994.

[43]   C. Ding, "The GIS-based Human-Interactive TAZ Design Algorithm: Examining the Impacts of Data Aggregation on Transportation-Planning Analysis", *Environ. Plan. B: Plan. Design.,* vol. 25, pp. 601-616, 1998.

[44]   C. Ding, K. Choi and T.J. Kim, "GIS-Based Traffic Analysis Zone Design". In the 3rd International Conference on Computers in Urban Planning and Urban Management, Atlanta, 1993.

[45]   K.T.Z. Chang, Khatib and Y. M. Ou, "Effects of Zoning Structure and Network Detail on Traffic Demand Modeling", *Environ. Plan. B: Plan. Des.,* vol. 29, no.1, pp. 37-52, 2002.

[46]   L. Martínez, J. Viegas, and E.A. Silva, "Zoning Decisions in Transport Planning and Their Impact on the Precision of Results" Transportation Research Record: *J. Transport. Res. Board.,* No. 1994, Transportation Research Board of the National Academies, Washington, D.C., pp. 58-65, 2007.

[47]   W. Barrett, Building a Regional Transportation Model with GIS Software, Proceedings for the 1991 GIS for Transportation Symposium, Ed. D. D. Moyer and B.L. Larson. AASHTO. Washington, D,C, pp.117-125, 1991.

[48]   C.C. Crevo, "Impacts of zonal reconfigurations on travel demand forecasts", *Transport Res. Rec.,* 1477, pp. 72-80, 1991.

[49]   N. Oppenheim, "*Urban travel demand modeling: from individual choices to general equilibrium*", John Wiley & Sons, Inc, New York, NY, 1995.

[50]   H .Miller, "Potential contributions of spatial analysis to geographic information systems for transportation (GIS-T)", *Geogr. Anal.,* vol. 31, pp. 373-399, 1999.

[51] A.J. Horowitz, "Computational Issues in Increasing Spatial Precision of Traffic Assignment", Transportation Research Record: *J. Transport. Res. Board.,* No. 1777, pp. 68-74, 2001.

[52] J. Sutton, "Role of geographic information systems in regional transportation planning", *Transport. Res. Rec.,* 1518, 1996.

[53] M. Hider and T. Spurr, "The Design and Development of Large-Scale Assignment Models Using Geographic Information Systems", *TRB.,* National Research Council, Washington, D.C., 2006.

[54] D. Ekern, "Remote Sensing Technology for Transportation", Presentation at the GIS-T Conference, April 8-12, 2001, Washington, DC. [Online] Available: http://www.bts.gov/gis/reference/gist/2001/23/231_remote_sensing/tsld001.html

[55] J.M. Usher, "Remote Sensing Applications in Transportation Modeling", In: Remote SensingTechnologies Center Final Report, 2000. [Online] Available: http://www.rstc.msstate.edu/publications/proposal1999-2001.html.

[56] X. Hu, and V. Tao, "Automatic extraction of main road centerlines from high resolution satellite imagery using hierarchical grouping", *Photogramm. Eng. Remote. Sens.,* vol. 73, no. 9, pp. 1049-1056, 2007.

[57] P. Doucette, P. Agoouris and A. Stefanidis, "Automated road extraction from high resolution multispectral imagery", *Photogramm. Eng. Remote. Sens.,* vol. 70, pp. 1405- 1416, 2004.

[58] T. Kim, S. Park, and M. Kim, "Tracking road centerlines from high resolution remote sensing images by least squares correlation matching", *Photogramm. Eng. Remote. Sens.,* vol. 70, pp.1417-1422, 2004.

[59] Y. Wang, X. Li, L. Zhang and W. Zhang, "Automatic Road Extraction of Urban Area from High Spatial Resolution Remotely Sensed Imagery", *Int. Arch. Photogrammetry, Remote Sen. Spatial Inform. Sci.*, vol. XXXVIL. Part B6b, Beijing, 2008.

[60] Grote and Heipke, "Road extraction for the update of road databases in suburban areas" The International Archive of the Photogrammerty Remote Sensing and Spatial Information Sciences, vol. XXXVII, part B3b, Beijing, 2008.

[61] H. Mayer, "Automatic object extraction from aerial imagery: a survey focusing on building", Computer Vision and Image Understanding, vol. 74, pp.138-139, 1999.

[62] D.S. Lee, J. Shan, and J.S. Bethel, "Class-guided building extraction from IKONOS imagery", *Photogramm. Eng. Remote. Sens.*, vol. 69, pp. 143-150, 2003.

[63] D.H. Lee, K.M. Lee and S.U. Lee, "Fusion of lidar and imagery for reliable building extraction", *Photogramm. Eng. Remote. Sens.*, vol. 74, pp. 215-225, 2008.

[64] X. Jin and C.H. Davis, "Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information", *EURASIP J. Appl.Signal. Process.,* vol. 14, pp. 2196-2206, 2005.

[65] G. Miliaresis, and N. Kokkas, "Segmentation and object-based classification for the extraction of the building class from LiDAR DEMs", *Comput. Geosci.,* vol. 33, pp. 1076-1087, 2007.

[66] Y. Zhang and B. Guindon, "Using satellite remote sensing to survey transport-related urban sustainability: methodologies for indicator quantification" Part 1: *Int. J. Appl. Earth Obs. Geoinfo.,* no. 8, pp. 149-164, 2006.

[67] M. Ehlers, M.A. Jadkowski, R.R. Howard and D.E. Brostuen, "Application of SPOT data for regional growth analysis and local planning", *Photogramm. Eng. Remote. Sens.*, vol. 56, pp.175-180, 1990.

[68] P.M. Treitz, P.J. Howard, and P. Gong, "Application of satellite and GIS technologies for land- cover and land-use mapping at the rural-urban fringe: a case study", *Photogramm. Eng. Remote. Sens.*, vol. 58, pp. 439-448, 1992.

[69] D. Lu, and Q. Weng, "Spectral mixture analysis of the urban landscapes in Indianapolis with landsat ETM+ imagery", *Photogramm. Eng. Remote. Sens.*, vol. 70, pp.1053- 1062, 2004.

[70] C.P. Lo, "Automated population and dwelling unit estimation from high resolution satellite images: a GIS approach", *Int. J. Remote. Sens.,* vol. 16, pp.17-34, 1995.

[71] P. Sutton, "Modeling population density with nighttime satellite imagery an GIS", *Comput. Environ. Urban. Sys.,* vol. 21, pp. 227-244, 1997.

[72] M. Langford, D.J. Maguire, and D.J. Unwin, "The areal interpolation problem: estimating population using remote sensing in a GIS framework. In: Geographical Information: Methodology and Potential Applications, I. Masser and M. Blakemore (eds), Longman, New York, pp. 55, 1991.

[73] J.T. Harvey, "Estimation census district population from satellite imagery: some approaches and limitations", *Int. J. Remote. Sens.,* vol. 23, pp. 2071-2095, 2002a.

[74] J.T. Harvey, "Population estimation models based on individual TM pixels", *Photogramm. Eng. Remote. Sens.*, vol. 68, pp. 1181-1192, 2002b.

[75] Y. Yuan, R.M. Smith, and W. F. Limp, "Remodeling census population with spatial information from LandSat TM imagery," Comput. *Environ. Urban. Syts.*, vol. 21, no. 3-4, pp. 245-258, 1997.

[76] S. Openshaw, O. Duke-Williams, and P. Rees, "Measuring confidentiality risks in census data". Working Paper, School of Geography, University of Leeds, 1997.

[77] A. Okabe, and N. Tagashira, "Spatial aggregation bias in a regression model containing a distance variable", *Geograph. Syst.,* vol. 3, pp.77- 100, 1996.

[78] B. Bhatta and O. Larsen, "Are intrazonal trips ignorable?" *Transport Policy.*, vol. 18, no. 11, pp. 13-22, Elsevier Ltd, 2010.

[79] K.M. Ford, and J. D. Fricker, "Real-time socio-economic data for travel demand modeling and project evaluation", Publication FHWA/IN/JTRP-2008/22. Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana, 2009.

[80] K.G. Goulias, "Transportation planning and modeling", In: Handbook of Transportation Engineering (Editor M. Kutz), Chapter 29. McGraw-Hill, New York, NY, pp. 3-25, 2004.

[81] Cambridge Systematics. Travel Model Validation Practices Peer Exchange White Paper. Peer Exchange Report. Prepared for Federal Highway Administration (FHWA), 2008.

[82] N.L. Faust, W.H. Anderson, and J.L. Star, "Geographic information systems and remote sensing future computing environment", *Photogramm. Eng. Remote. Sens.*, vol. 57, pp.655-668, 1991.

[83] Jean-Claude, Thill, "Geographic information systems for transportation in perspective", *Transport. Res.,* vol. 8, Elsevier Science Ltd., UK, 2000.

[84] G.G. Wilkinson, "A review of current issues in the integration of GIS and remote sensing data", In: *Int. J. GIS.,* vol.10, pp. 85-101, 1996.

[85] Q. Weng, "*Remote sensing and GIS integration: theories, methods, and applications*", New York: McGraw-Hill Companies, Inc. ISBN: 978-0-07- 160654-7, 2010.

[86] T.C. Matisziw, T.H. Grubesic, and H. Wei, "Downscaling spatial structure for the analysis of epidemiological data", *Comput. Environ. Urban. Syst.,* vol. 32, pp. 81-93, 2008.