# VALIDATION AND EVALUATION OF INTEGRATED URBAN MODELS

# VALIDATION AND EVALUATION OF INTEGRATED URBAN MODELS

By

Genhan Chen, M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Arts

McMaster University

MASTER OF ARTS (2009) McMaster University

(School of Geography and Earth Sciences) Hamilton, Ontario

TITLE: Validation and Evaluation of Integrated Urban Models

AUTHOR: Genhan Chen, M.Sc. (Wuhan University)
B.Sc. (Wuhan University)

SUPERVISOR: Dr. Pavlos S. Kanaroglou

NUMBER OF PAGES: viii, 84

# Abstract

The objective of this thesis is to validate the results of the Integrated Model of Urban LAnd use and Transportation for Environmental analysis (IMULATE), as well as the Integrated Model for Population Aging Consequences on Transportation (IMPACT) in the Census Metropolitan Area (CMA) of Hamilton, Ontario. The land use/demographic modules of these two models are validated using observed data from Statistics Canada, while data from the Transportation Tomorrow Survey (TTS) and the City of Hamilton are used to validate the transportation modules. Statistical, graphical and GIS visualization techniques are incorporated into this validation.

This thesis illustrates some sub-modules in IMULATE and IMPACT can work very well, while the predictive ability of others is not as good. IMULATE considers more factors to simulate land use development. It generates accurate simulations of household dynamic using observed data as exogenous input. We used the "final demands" for economic sectors as the exogenous input to estimate the employment's distribution. After recalculating the "final demands" in its employment location model, the generated employment is also found to be close to the observed value. However, we found that its transportation module was not able to produce accurate predictions of inter-zonal trips and traffic flows over the Hamilton's road network.

IMPACT can predict the growth of population by gender and age with good accuracy. The simulation results for males are better than females. The inter-zonal trips generated by IMPACT are found to be much closer to the observed value than the inter-zonal trips generated by IMULATE. However, we found that the simulated trips have lower dispersion across the city than normally observed.

# Acknowledgements

Firstly, I would like to express my gratitude to my supervisor, Dr. Pavlos Kanaroglou, for his support and encouragement in my study. With his enthusiasm, his inspiration and his insightful guidance, I have learned how to think carefully as a researcher. He provided sound advice and good teaching not only for my research, but also for the living skills in Canada.

I would also like to thank Dr. Hanna Maoh. With his exceptional clarity of explanation, I could resolve each problem successfully. He also brought plenty of good ideas into my research. I would like to thank Dr. Juliet Wallace. She provided a lot of useful information to my work, and helped me revise some chapters of my thesis.

Then I would like to thank Dr. Darren Scott, Dr. Bruce Newbold and Dr. Michael Mercier. Thanks for their involvements in my defense. They provided a lot of good advices revising my thesis. I really studied much from them.

I am also grateful for the students, staff and associated faculty members at the Centre of Spatial Analysis (CSpA). They made my two-year experience here a pleasure.

Lastly, I wish to thank my family and friends. They supported me and loved me. To them I dedicate this thesis.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1. Introduction

Over the past several decades, urban systems have become increasingly complex as an integrated entity of land use, transportation and environment. Integrated Urban Models (IUMs) are developed to reduce complexity and provide a concise understanding of some aspects of urban dynamics. Information obtained from the results of those models supports the users on their decision-making process. Many researchers use IUMs to help them simulate urban changes and draw up a plan for developing the city. For example, MEPLAN model has been applied to the Sacramento, California region by Abraham et al. (1999a, 1999b), CUF model has been used to evaluate the policies for San Francisco Bay and Sacramento areas (Landis, 1995), and Kang et al. (2009) applied IMULATE to analyze the transportation issues in Hamilton area. The individuals affected by those decisions are rightly concerned with whether those models can predict the future accurately. This concern is addressed through model validation work.

We present a diagram that relates validation to developing simulation models in Figure 1.1. This diagram shows the relationship between a real world and a simulation world. We understand the urban system as a complex entity in the real world. We construct system theories by abstracting its characteristics and by experimenting with the real data observed from the system. Based on system theories, we develop the IUMs on a computer system, implementing the simulation process. The simulation results are the data obtained from scenarios conducted on the models. Model validation is conducted to determine whether the simulation results are consistent with the system theories and the urban behaviors in the real world. We often validate the models using the data observed from the urban system to serve as a comparison with model predictions.

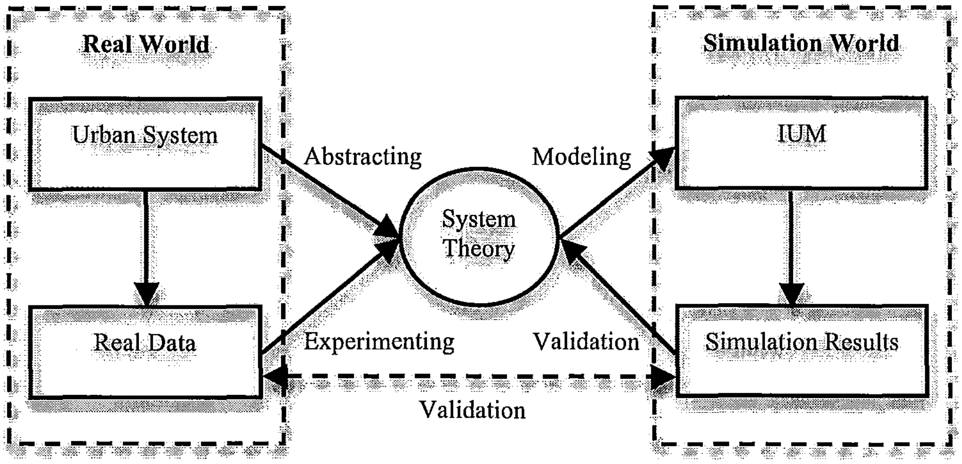



**Figure 1.1 The relationship between real world and simulation world**

Thus, model validation is a crucial process to test the relationship between the IUMs and the urban system. Franklin et al. (2002) reported that, during the development process, the model validation was the last stage before applications. In practice, some validation work should also be conducted after applications since collecting real-world data can be a time-consuming process. In some cases, observed data may become available many years after the system developed.

Center of Spatial Analysis at McMaster University has developed two IUMs, the Integrated Model of Urban LAnd Use and Transportation for Environmental analysis (IMULATE) and the Integrated Model for Population Ageing Consequences on Transportation (IMPACT), to simulate the interactions between land use, transportation and environment in the Census Metropolitan Area (CMA) of Hamilton (Anderson et al., 1996; Maoh, et al., 2008). Although some researchers have used these two models analyzing some social problems, until now, these two models have not been validated systematically. This research aims to complete this validation work.

Validation is a complex process. We will solve three problems in this research. First is the structure analysis of IMULATE and IMPACT. These two models are composed of many sub-models, each of which will generate the corresponding simulated results. The final misestimates could be caused by any of those sub-models. Thus, the key to validating these two models is to find out the outputs of each sub-model, and test each sub-model one by one. This will help us to locate the sources of potential problems.

The second problem to be solved is the collection and integration of observed data from different sources. Statistics Canada, Transportation Tomorrow Survey (TTS) and the City of Hamilton are all potential sources for observed data. However, those data are always stored with different formats. Prior to using those data in validation work, we should clean and transfer the observed data into the formats utilized by IMULATE and IMPACT.

The last problem is how to choose suitable methodologies to validate these models. There are a variety of methods that could be used in this work. Each method has its own benefits and drawbacks. We will use several methods to validate the same model, and provide our interpretation of results in different aspects.

In the following chapters we will focus on these three problems. In Chapter 2, we introduce the development of IUMs and state the related research. We also list the empirical validation work for some main IUMs. After that, we analyze the structures of IMULATE and IMPACT, and indicate the main function of each sub-model.

In Chapter 3, we describe the overall data cleaning process, including the obtaining of simulation results from each sub-model and the transformation of the data observed from the real world. Then we introduce the main methods employed to validate the models in this research.

Chapter 4 is the core of this thesis. After we use the methods mentioned in Chapter 3 to validate the models, we present tables and figures of the validation results. We analyze these results and discuss the relationship between the errors and their related sub-models. Then we report the potential causes of the misestimates.

In the final chapter, we discuss the main achievements of this thesis and provide some potential avenues for future work.

# Chapter 2. Background

## 2.1 Literature review

### 2.1.1 Integrated urban models

Integrated urban models (IUMs) are usually made up of several simulation sub-models used to predict the changes over time in the spatial distribution of land use, and the corresponding travel behavior in an urban area. Typically, these sub-models are interlinked, and any changes in the input of the models will produce varying simulated results.

IUMs are widely applied to predict the outcomes of alternative policies in many metropolitan areas (Kanaroglou et al., 2002). Transportation planning in cities has been dominated for a while by the use of the four-stage Urban Transportation Modelling System (UTMS). The UTMS consists of four interlinked stages: trip generation, trip distribution, modal split, and traffic assignment (Meyer et al., 1984; Kanaroglou et al., 2002).

Trip generation is the prediction of the total traffic flows generated into and out of each traffic analysis zone (TAZ) in the study area. Regression and category analysis are two methods that are usually employed to estimate the generated trips in each TAZ (Federal Highway Administrator, 1975). Paez et al. (2006) applied an ordered probit model to estimate trip generation in Hamilton, Ontario. This kind of trip generation model could simulate generated trips at the household level, as opposed to the zonal level, which is the case with regression and category analysis.

Trip distribution is the prediction of the trips between any origin zone and any destination zone. The outputs of trip distribution models are origin-destination (O-D) matrices. The cells of the O-D matrix consist of an origin-destination pair in which the rows represent the origin locations and the columns represent the destination locations. Growth factor techniques (Fratar et al., 1954; Brokke, 1958; Hutchinson, 1974) and gravity models (Meyer et al., 1984) are usually used to formulate trip distribution models.

The objective of the third stage of a UTMS is to differentiate motorized from non-motorized trips and to estimate an O-D matrix of motorized trips. To this end, modal split models are used to predict the proportion of trips using each transportation mode. The multinomial logit model (MNL) is a typical method used to estimate the people's choice of traffic mode (Ben-Akiva, 1974).

The last stage in UTMS is to distribute the motorized O-D trips into actual routes in the road network. One possible methodology to estimate the traffic assignment is the "all or nothing" network assignment, which fails to take account of the peak-hour congestion and other flow characteristics. A more popular approach is the stochastic user equilibrium assignment, which overcomes many of the difficulties of the "all or nothing" approach (Nguyen, 1974; LeBlanc et al., 1976; Daganzo et al., 1977).

The earliest work to model the interactions between the distribution of residents and places of work was by Lowry (1964). Since then, there have been several improvements to Lowry's model, especially when it is combined with UTMS. Notable among them is the Integrated Transportation and Land Use Package (ITLUP) initiated by Putman (1983). This model consists of two interlinked sub-modules to simulate land use changes. They are the Employment Allocation Model (EMPAL) and the Disaggregate Residential Allocation Model (DRAM). EMPAL predicts the location of employment using exogenous economic forecasts, while DRAM simulates the distribution of households using a model based on a distance decay function. This model uses standard four-stage UTMS to simulate the transportation system. The outputs from EMPAL and DRAM determine the first two stages of UTMS, and the four-stage process provides feedback to those two sub-modules.

Wilson (1970) employed the table of journey to work trips as an example, to explain the relationship between entropy and the probability that a specific trip distribution occurs. He pointed out that the entropy was a useful measure to indicate the state of a system. He reported that the entropy was the logarithm of the probability that a distribution of the system would happen. To consider various constraints in the urban system, he embedded the entropy maximization model into the spatial interaction model. He used this theory to simulate the distribution of various activities over zones and the interactions between zones at the same time (Wilson et al., 1981).

After Alonso (1964) proposed urban economic theory in his work, the utility function has often been used to estimate people's location choice. The utility function measures the relative satisfaction for people making a decision. In theory, no one will move elsewhere when his utility, or the satisfaction from its residence given income constraints, is as close to the maximum as it can be. The market clearing mechanism is proposed based on this idea. This mechanism indicates a balance between supply and demand in a market. To reflect changes in the housing market, housing prices can be estimated with the market clearing mechanism. Anas (1986) utilized a combination of utility maximization and entropy maximization to estimate housing rental in his research. He predicted the rents, keeping the balance between the house owners and housing consumers. Furthermore, he used more complicated probabilistic choice functions to simulate the distribution of households and employment, instead of the earlier gravity models. MUSSA and UrbanSim are notable applications that derive from the combination of entropy and utility maximization (Martinez, 1996; Waddell, 1998, 2002).

Leontief (1941) initiated an application to predict economic development using input-output (I-O) models. An I-O model simulates the transactions between different economic sectors in the form of a matrix. In an I-O matrix, each column represents the monetary inputs from an economic sector, while each row represents the monetary outputs from an economic sector. An array of column vectors, on the right of the matrix, recording the total monetary value for every economic sector, is called "final demand". Macgill (1977) conducted some experiments to adapt the I-O model into the Lowry model, and estimated its final demand with entropy maximization. In his later research, he applied I-O models to discussing the transactions between economic sectors across different regions (Macgill et al., 1979). The notable applications of the I-O model in the context of IUMs are TRANUS and MEPLAN. They both estimate employment in an economic sector using the related final demand and the technical coefficients for that region (la de Barra et al., 1984; la de Barra, 1989; Hunt et al., 1993; Echenique et al., 1994; Abraham et al., 1999).

All of the aforementioned IUMs' analysis is simulated based on zone units. In the future, urban modeling will be implemented from aggregate level to disaggregate level. Earlier researchers have used activity-based strategies to analyze people's travel behaviors (Kitamura, 1988; Jones et al., 1990; Pas, 1990). Their research focused on analyzing a sequence of activities in a specific period, in order to understand complex travel behaviors. Until now, a lot of applications have been developed based on activity-based approaches. Important work done on this topic includes: Kawakami et al. (1990), Golledge et al. (1994), Ettema et al. (1996), Kitamura et al. (1996) and Miller et al., (2003).

With the development of the computing and information sciences, object-oriented theory has been applied to urban forecasting systems. Based on this theory, the microsimulation modeling technique is generated, and it operates at the disaggregate level of individual units (Miller, 1996). The Microsimulation-based system is dynamic and complex in behavior. It represents the complex interactions between different types of individual actors. This kind of system provides a very simple way to computer programming. The notable efforts of microsimulation applications are IRPUD (Wegener, 1998), UrbanSim (Waddell, 2000), ILUTE (Miller et al., 1987; Salvini et al., 2005), ILUMASS (Moeckel et al., 2002; Wagner et al., 2007) and PUMA (Ettema et al, 2006).

### 2.1.2 Empirical model validation

Rodier (2005) defined the model validation as the demonstration of "how well a model predicts actual observed behavior". It is a crucial section during the model development process. It provides users enough confidence in the system (Waddell et al., 2004). Franklin et al. (2002) reported two general approaches to assess the model's validity: historical validation and sensitivity analysis.

Historical validation is a commonly used methodology to evaluate the accuracy of the model. It assesses the quality of the model by examining whether the simulated values conform to reality. To use this methodology, two groups of values are often needed. The model generates one group, while the other is collected from the real world. In the following sections, the first group of data is called simulated data, and the latter is referred to as observed data. To analyze the difference between those two groups of data, several goodness-of-fit statistical methods are widely used by geographers. The choice of a suitable statistic depends on the user's subjectivity and the data's complexity. Physical geographers often used some simple ways to test the goodness-of-fit, such as simply calculating the sum of squared errors (Tait et al., 1998; Paktunc, 1998; Lee et al., 1998; Ferguson, 1999). Around the same time, Legates et al. (1999) conducted goodness-of-fit tests of hydrologic and hydro-climatic models using some correlation-based measures, such as the coefficient of determination. In human geography, Fotheringham et al. (1987) proposed different techniques to test the goodness-of-fit. He reported that r-squared ($R^2$) test and the Standardized Root Mean Square Error (SRMSE) were two important measures to evaluate spatial interaction models (Fotheringham et al., 1989). Before that, he used $R^2$ to determine the differences in some linear interaction models (Fotheringham, 1983). Lewis (1975) also applied the $R^2$ test. He compared the gravity model with the Heckscher-Ohlin model using the $R^2$ statistic.

In addition, the chi-squared statistics has been used in model validations (Hathaway, 1975; Baxer et al., 1981; Knudsen et al., 1986). Voas et al. (2001) reviewed some commonly used chi-squared statistics and discussed their weaknesses and strengths. They reported that a significant problem with the classic Pearson chi-squared test was that it could not handle the presence of zeros in the simulated values. The best ways to resolve this problem were the Neyman-Pearson hybrid ($X_H^2$) statistic and the Freeman-Tukey ($FT^2$) statistic. They indicated those three methodologies were all based on chi-squared distribution and their tested results were quite similar. Furthermore, they also introduced some measures of fit based on information theory. In the early years, Shannon (1956) reported that the mathematics was the core of information theory. Those ideas included in the information theory have produced some goodness-of-fit test methodologies. Many researchers occasionally used these information-theoretic methodologies in their research. (Smith et al., 1981; Knudsen et al., 1986; Fotheringham et al., 1987).

Geographers have applied various methodologies to validating different IUMs. Waddell (2002) conducted validation of the UrbanSim model for Eugene-Springfield, Oregon. He used $R^2$ statistics to compare the data simulated from 1980 to 1994 with the 1994 observed data. He also validated another application of UrbanSim model in combination with the Wasatch Front Regional Council (WFRC) travel model system (Waddell et al., 2003). In this research, they used two methods of visualization to test the simulated results. One way was to use the line chart to present the simulated and historical development over time, and then compare those two lines. The other way was to display the differences between the simulated and historical data on the map, to check the spatial distribution of the errors.

Wegener (1982) conducted some goodness-of-fit tests to assess the credibility of a multilevel dynamic simulation model for economic and demographic development of the Dortmund, Germany urban region. These tests demonstrate how well the model reproduces the general spatial development in the Dortmund region, with the comparisons between observed and predicted population data of the 30 zones. $R^2$ statistic and mean absolute percentage error (MAPE) statistic were the two main measures used to test the validity of this model for its population and migration predictions.

Martinez (1996) assessed the goodness-of-fit of the MUSSA model using $R^2$ statistic. A higher $R^2$ indicated the more successful simulation. Using this method, he found the simulation of high-income residents' location choice was worse than the simulations of the other income groups since its related $R^2$ was very low. The allocation of industry and education were less successful than of retail and service since the related $R^2$ were relative low.

Abraham (2000) conducted goodness-of-fit tests for the Sacramento MEPLAN model. He compared simulated and observed values using the proportion of trips in each O-D pair. The difference between simulated and observed proportion represented the simulated error. He assessed the simulated accuracy with the sum of the square errors. He also depicted the comparisons of auto trips by trip type and travel time using histogram charts. Rodier (2005) validated the 2000 Sacramento MEPLAN model using various methods. He compared the simulated 2000 data with the observed data that were not used to calibrate the model. The difference between them was indicated with the algebraic error (ALE), the mean algebraic error (MALE), the algebraic percent error (ALPE), the absolute percent error (APE), the mean algebraic percent error (MALPE), the model algebraic change (MALC), the model algebraic percent change (MALPC), the model absolute percent change (MAPC), the estimated algebraic change (EALC) and the estimated algebraic percentage change (EALPC).

Mohammadian et al. (2003) validated the ILUTE microsimulation model using $R^2$ statistic and likelihood ratio ($\rho^2$) test. Mohammed et al. (2008) adopted some hypothesis tests to assess the predictive power of the ILUTE model. The main goodness-of-fit methods utilized in their research were the $R^2$ statistic and F-statistic. The differences between the simulated results and the observed values were depicted with histogram charts.

Roorda et al. (2008) conducted research to validate the Travel Activity Scheduler for Household Agents (TASHA) model, in Greater Toronto Area (GTA), Canada. This validation aimed to verify whether TASHA could simulate the activity generation, activity location choice and activity scheduling correctly. The observed data from Transportation Tomorrow Survey (TTS) database were utilized to compare with the simulated data from TASHA using Kolmogorov-Smirnov statistics. They represented the comparisons with histogram charts and line charts.

Landis et al. (1998) tested the second generation of California Urban Futures Model (CUFM) according to whether the simulated results were classified correctly by category when compared with the observed category. The main methods utilized in his research were maximum-probability method and the case-constrained method.

The sensitivity analysis of a model examines whether the model responds suitably to related policy changes (Borning et al., 2006). A shortcoming of the historical validation is that the needed historical data are always not always easy to obtain. Thus, after a model has been developed, the developers often do sensitivity analysis to test the system, in order to demonstrate that it could simulate the urban dynamics according the theoretical expectations. The objective of sensitivity analysis is to check how outputs change when different modifications are applied to the initial database, and suggests specific potential sources of model errors (Rodier, 2005). An IUM usually consists of several modules and associated sub-models. Each of these sub-models does a specific job in the overall simulation. Any changes in those models will make the final simulated results different. If the final results are not as accurate as expected, we can find out which models generate the errors with changing the inputs of each model. During sensitivity analysis of the IUM system, a variety of different land use policies and transportation scenarios are used to test the integrated model. These scenarios are intended to evaluate the validity of model itself (Franklin et al., 2002).

Waddell et al. (2003) conducted a project to evaluate the application of the UrbanSim, integrated with WFRC travel model system. He tested the system with scenarios based on two changes for the starting database: 1) changes in the spatial distribution of population and employment, such as the limit of urban expansion; 2) significant changes to transportation supply, such as the new investments on highway and transit. They addressed the analysis results from four aspects, including the incentive of this scenario, the theory expectation of this scenario, the responses of the model system to this scenario, and some indicators to measure these responses. Paradhan et al. (2002) also conducted some sensitivity analysis to assess the uncertainty propagation in UrbanSim.

Abraham et al. (1995, 1999) conducted sensitivity analysis to demonstrate the richness of the Sacramento MEPLAN model. His research explained how land use activities and transportation behaviors responded to each other by the comparisons between five development scenarios and a trend scenario. Rodier (2004) employed several scenarios to test the integration of the Sacramento MEPLAN model and the UPLAN travel model. The tested results addressed the strengths and weaknesses of the model, the factors appreciably reducing congestions and emissions, and the significance of the results of the alternative scenarios simulated.

Various teams of researchers conducted several applications for different land use and transport interaction models in Sacramento region. Hunt et al. (2001) conducted some comparisons among the different models and compared their base case simulated results to the observed values. Those comparison results showed some limitations of aggregate models and provided important insight into how models should be calibrated and how their results should be used.

Landis (1994, 1995) conducted sensitivity analysis during the CUFM development process, such as business-as-usual scenario, maximum environmental protection scenario and compact city scenario. He tested how the model simulates these policy changes upon the location, amount and intensity of urban growth.

Multivariate sensitivity analysis, examining the effects of input uncertainty on output, helps identify the degree to which model outputs are affected by changes in inputs, while controlling for variations in other inputs (Krishnamurthy et al., 2003). Aramu et al. (2006) conducted some policy tests to show the response of the system to some specific inputs for a new implementation of DELTA land-use modeling package.

## 2.2 Introduction to IMULATE and IMPACT

### 2.2.1 Overview of IMULATE

The Integrated Model of Urban LAnd use and Transportation for Environmental analysis (IMULATE) is an integrated urban model developed to study urban issues in the Hamilton CMA (Anderson et al, 1996). It is used to assess land use change, the generated travel behaviors and their corresponding environmental impacts. IMULATE is composed of three interlinked modules: land use module, transportation module and environment module. Figure 2.1 graphically depicts the general structure of the entire model.

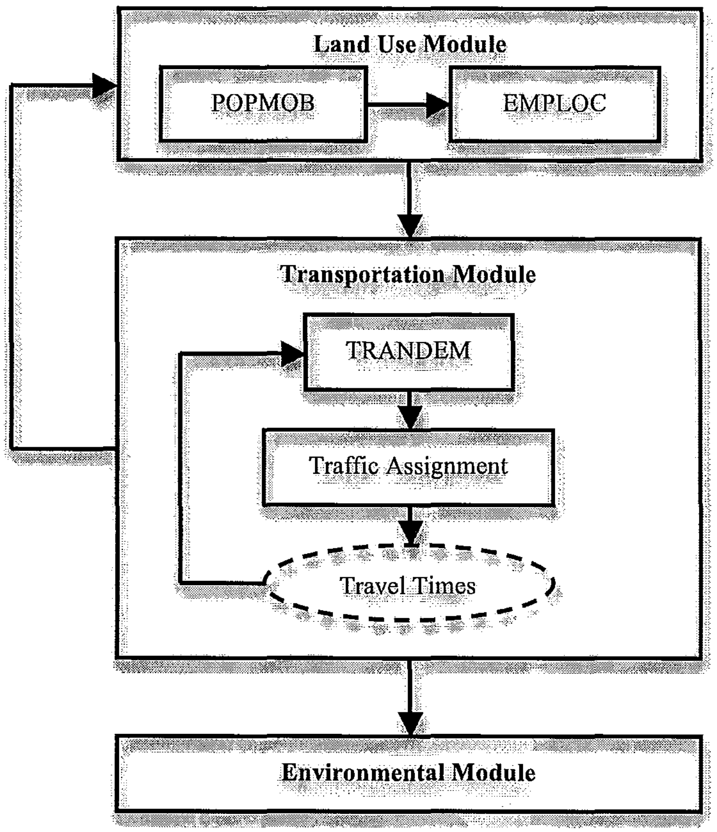



**Figure 2.1 General structure of IMULATE**

The land use module is made up of a population mobility sub-module (POPMOB) and an employment location sub-module (EMPLOC). POPMOB predicts the spatial distribution of urban households across the Hamilton CMA. The key element in this simulation is the multinomial logit (MNL) function, which estimates the destination choice probabilities of households. The utility of those probabilities is affected by the zonal characteristics such as moving distance, the average yearly income, their housing expenditure and the level of accessibility to the destination. An important zonal characteristic affecting people's mobility is the proximity to hazardous industrial sites in Hamilton. In order to project the housing expenditure, market clearing equilibrium theory is applied to estimate the housing prices and rents. New vacancies generated by newly constructed dwellings play an important role in this equilibrium function.

EMPLOC determines the location of firms in the Hamilton CMA as the places of employment. This sub-module predicts the distribution of firms at the city level and census tract level by three main processes. At the city level, the total number of firms is estimated by the input-output function using the exogenous input of "final demands". The "final demands" are used as the exogenous input to simulate the development of employment. For each census tract, the number of lost firms in each census tract, which is determined using proximity to downtown, proximity to regional malls and the location of census tracts in the city, is aggregated into the city level. The total number of gained firms is calculated with the total number of firms and lost firms in the city. The number of gained firms in each census tract is estimated by MNL. The parameters related to firms gained are the proximity to Central Business District (CBD), highways and regional malls, are affected by the measures of agglomeration economies, and also relates to the population size and household density. Thus, the number of firms in each census tract can be estimated by the corresponding number of lost firms and gained firms. The distribution of employment is simulated with the distribution of firms and the average size of each firm.

The transportation module consists of the travel demand sub-module (TRANDEM) and the traffic assignment sub-module. Based on the matrix that links place of residence to place of work generated from the land use module, TRANDEM simulates the distribution of trips in the morning peak time between each pair of traffic analysis zones (TAZs) for all the workers in the city. The distribution of work trips, school trips and discretionary trips is projected over all TAZs using the Fratar algorithm, using the outputs of the land use module. Those forecast trips are broken down using the MNL function into four modes, including auto-driver, auto-passenger, transit and other modes.

The traffic assignment sub-module converts inter-zonal trips into traffic flows on a link-by-link basis using a stochastic user equilibrium (SUE) assignment model. The equilibrium of the transportation module is reached when the travel time along each link generated by the SUE assignment model does not change significantly from one iteration to the next. The effect of congestion on travel speed estimated by this sub-module affects the travel demand analysis. The transportation simulation will not stop until the equilibrium is reached.

Different from classic UTMS, the transportation model used in IMULATE has a relationship with its land use module. The equilibrium of the land use module and transportation module will be established using the "inclusive value" which is a measure of household's mobility. The "inclusive value" is generated by the transportation module and represents the interactions between the land use module and transportation module. Since the outputs from transportation module change the mobility of households and employments, the overall simulation will be completed after the system achieves equilibrium.

The environment module predicts emissions for each road link, using the congested travel times and average congested speeds simulated from transportation module. The emissions of carbon monoxide (CO), hydrocarbons (HC), nitrogen oxides (NOx) and particulate matter (PM2.5 and PM10) for each link is estimated by an automobile emission model, MOBILE 6.2C, which is developed by the US Environmental Protection Agency and customized for the Canadian fleet of vehicles. This module simulates the emission factors of pollutants for the four seasons of the year and estimates the energy consumption for each road link

### 2.2.2 Overview of IMPACT

Integrated Model for Population Ageing Consequences on Transportation (IMPACT) is developed to simulate the interactions between the demographic changes and the performance of transportation system in Hamilton CMA (Maoh et al., 2008). This model is used to analyze urban dynamic related issues.

IMPACT consists of three main modules: demographic module, transportation module and environmental module, as shown in Figure 2.2.

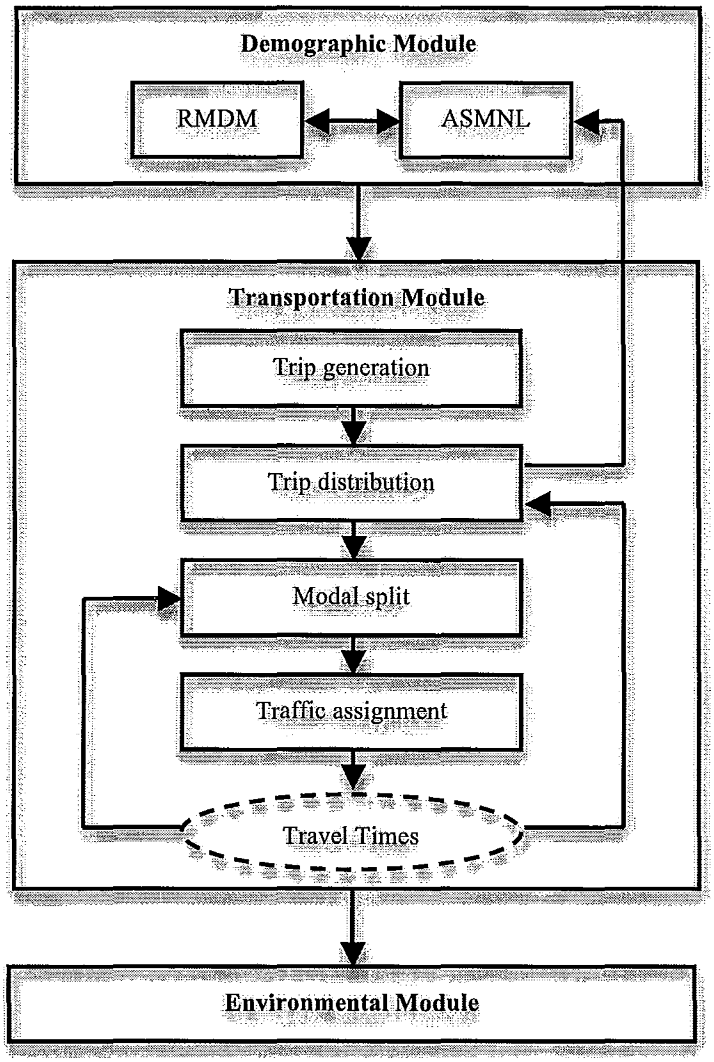



**Figure 2.2 General structure of IMPACT**

The demographic module simulates the population development by age and gender at two levels: municipal level and census tract level. This module is made up of municipal population projection using Rogers' multiregional demographic model (RMDM) (Rogers, 1995) and the distribution forecast of population over census tracts using an aggregated spatial multinomial logit (ASMNL) model. RMDM predicts the progression based on the age and gender related data, including birth rate, death rate and migration between different municipalities. ASMNL is a destination choice model that estimates the probability using zonal characteristics, including new constructed dwellings, housing prices, number of schools and recreational land uses.

The four-stage urban transportation modeling system (UTMS) is adopted in the transportation module (Maoh et al., 2008). Work and non-work trips are generated by an ordered probit-based trip generation model (Paez et al., 2006), which is used to estimate the probability of generating 0, 1, 2 and 3-or-more trips from an age cohort with the information on personal attributes, household attributes and zonal attributes. The distribution of generated trips is estimated using a gravity-based trip distribution model with the information on travel times between different zones in a congested situation and the attractiveness of the travel destination. The results of trip distribution model affect the residential mobility value in ASMNL.

The same as IMULATE, the calculated trips for each origin-destination pair generated by trip distribution model is split by three modes: auto-driver, auto-passenger and other modes using MNL. Personal attributes, household attributes and zonal attributes are employed into the formulation of MNL mode choice models. The results are distributed into each link of road network in Hamilton CMA using a SUE traffic assignment model, which is used in IMULATE. The generated congested effect and travel speed are used to recalculate trip distribution and modal split models, and the iterations will stop after the transportation module achieves equilibrium.

The link flows and average link speeds simulated from transportation module are used to estimate emissions for each road link by the environmental module. Different from IMULATE, MOBILE 6C emission model is employed in IMPACT to estimate the emissions of carbon monoxide (CO), hydrocarbons (HC), nitrogen oxides (NOx) and particulate matter (PM2.5 and PM10) for each link.

# Chapter 3. Data Preparation and Methodologies

## 3.1 Data preparation

### 3.1.1 Overview

IMULATE and IMPACT are computer simulation models that aim to predict land use changes and related traffic volumes in the transportation network of an urban area. Both models are designed to simulate the urban development in five-year time steps. The objective of this thesis is to evaluate the performance of the two models by comparing the simulated results against land use and transportation data collected for the purposes of this evaluation. In this chapter we describe the collected data and the methods that were used to facilitate this comparison. Both models are tested for their implementation in the Census Metropolitan Area (CMA) of Hamilton.

From 1986 to 2006, Statistics Canada, Transportation Tomorrow Survey (TTS) and the City of Hamilton have collected relevant data for the Hamilton Census Metropolitan Area (CMA). Statistics Canada provides the information related to population, employment and housing, while the transportation related data are provided by TTS and the City of Hamilton. We utilize these kinds of data as the observed data in this research. The availability of the observed data provides an opportunity to determine whether the model can accurately predict reality.

In both IMULATE and IMPACT, two GIS layers are used to depict the Hamilton CMA, including a traffic analysis zones (TAZs) layer based on census tract divisions and a road network layer representing all the major roads in the CMA. The TAZs and road links in IMULATE are identified with 1986 IDs, while in IMPACT they are identified with 1996 IDs. Since the data used in this research are collected from several data sources, the collected data are always shown using different formats. Thus, prior to the use of the real data for model validation, all the data should be cleaned and transformed.

### 3.1.2 Data for IMULATE

In IMULATE, TAZs correspond to the 1986 census tract divisions, which include 151 zones. The road network is represented in a layer that includes 1,540 traffic links and 1,100 nodes. There are three main modules in IMULATE: the land use module, the transportation module and the environmental module. Since the environmental module only has a one-way relationship with the transportation module and actual traffic emission data are not available, we did not complete its validation work in this research. To the other two modules, the simulated results of households, employments, inter-zonal trips and traffic flows are validated against the observed data provided by Statistics Canada, TTS and City of Hamilton.

**Land use module**

The land use module is composed of the POPMOB sub-module and the EMPLOC sub-module. The validation of POPMOB involves the comparison of model predicted households at the TAZs level to those observed in data provided by Statistics Canada. The model was calibrated for the year 1986. Table 3.1 indicates that the number of census tracts in the Hamilton CMA has grown from 151 for 1986 to 177 in 2006, mainly by sub-dividing larger areas in peripheral tracts, as their population has grown because of suburbanization.

**Table 3.1 The number of census tracts for different simulated years**

| Year | 1986 | 1991 | 1996 | 2001 | 2006 |
|---|---|---|---|---|---|
| # Census Tracts | 151 | 163 | 163 | 173 | 177 |

The IDs of those new zones have been redefined. In order to match the TAZs in IMULATE, the observed data should be transformed into the census tract divisions of 1986, and the corresponding households' information will be reprocessed with the change of tracts.

This sub-module simulates the distribution of households for 151 TAZs in the city. All the generated data are stored according to the order of census tract IDs that are defined by Statistics Canada. However, during the simulation process, the system only simulates the data in 148 TAZs with missing information for the zones with census tract ID 1479, 1481 and 3542. The first two zones are located in the downtown area of Hamilton, while the last zone is near the Queen Elizabeth Way (QEW). The final simulated data used to validate the POPMOB sub-module are for 145 TAZs without zones with census tract ID 4124, 4125 and 4126. Those three lost zones are located in the downtown of Burlington.

During the simulation process, IMULATE uses some exogenous data to estimate the distribution of households. Exogenous data are normally used in the design of scenarios that are to be investigated with the system. In this case, in order to test the performance of the module we made use of the observed data on new dwellings from Statistics Canada as the exogenous input. Under these circumstances, if the model performs in a satisfactory way, it should be able to replicate well the household distribution at the end of each time period.

The EMPLOC sub-module simulates the distribution of employment across 151 TAZs in the city. Statistics Canada provides the observed data of employment to test this sub-module. The observed data in 1996 and 2001 are calculated using journey-to-work matrices, and the 2006 data are calculated with the number of jobs for TAZs in 2006. The journey-to-work matrices used in this validation work identify the places of all non-institutional residents older than 15 years of age and their work locations. Figure 3.1 describes the process to calculate the observed employments using the journey-to-work matrix.

Place of Work

$$
\begin{array}{r|cccccc|}
\cline{2-7}
 & T_{01} & T_{02} & & {}^{*}T_{0j} & & T_{0(151)} \\
 & T_{11} & T_{12} & \cdots & T_{1j} & \cdots & T_{1(151)} \\
 & T_{21} & T_{22} & \cdots & T_{2j} & \cdots & T_{2(151)} \\
\text{Place of Residence} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & T_{i1} & T_{i2} & \cdots & T_{ij} & \cdots & T_{i(151)} \\
 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & T_{(151)1} & T_{(151)2} & \cdots & T_{(151)j} & \cdots & T_{(151)(151)} \\
\cline{2-7}
\text{Employment} & \sum_{i=0}^{151} T_{i1} & \sum_{i=0}^{151} T_{i2} & \cdots & \sum_{i=0}^{151} T_{ij} & \cdots & \sum_{i=0}^{151} T_{i151}
\end{array}
$$

* If i = 0, $T_{ij}$ is the number of workers who live outside of Hamilton CMA but work in zone j of Hamilton CMA. Otherwise, $T_{ij}$ is the number of workers who live in zone i and work in zone j.

**Figure 3.1 Observed employments in each TAZ**

Thus, employment in census tract i is estimated by the residents who work there living across all TAZs in the CMA of Hamilton or outside of it. Since the journey-to-work data in 2006 are not available, the 2006 occupied jobs data provided by Statistics Canada are used to compare with the simulated results. The outputs of EMPLOC are also generated according to the 1986 census tract IDs. Thus, the simulated results in 1996, 2001 and 2006 should be reprocessed into 151 TAZs using similar methods as in POPMOB.

EMPLOC simulates the distribution of employment using the exogenous final demands. We have used the 1986 final demands to project the final demands for the other simulated periods. The projected final demands make the simulations different from the actual developments. In this research, we recalculate a new group of final demands to make the simulated results closer to the real values.

**Transportation module**

The transportation module is composed of two sub-modules: the TRANDEM sub-module, which simulates modal split and the distribution of motorized trips between different zones, and the traffic assignment sub-module, which simulates the traffic flows for each link in Hamilton road network. The inter-zonal trips are simulated based on 151 TAZs in the city, while the simulated traffic flows are distributed into 1,540 links of the road network.

The TRANDEM sub-module predicts the inter-zonal trips across TAZs for the morning peak period of a typical weekday. The main outputs of this sub-module are the origin-destination (O-D) matrices by transportation modes and travel purposes. In this research, the observed O-D matrices for auto-driver mode provided by TTS for 1991, 1996, 2001 and 2006 are used to validate this sub-module. The TTS is a large-scale cross-sectional trip survey of a random sample of 5% of the households in the Greater

Toronto Area (GTA) and surrounding area of Central Ontario every five years since 1986. TTS only provides the information for 116 census tracts in Hamilton CMA, excluding Burlington area and Grimsby area. Thus, we only validate the simulations of inter-zonal trips for 116 census tracts in this research. In IMULATE, we projected the inter-zonal trips in the morning peak hour (7:00 am – 8:00 am) using the trips generated in the morning period (6:00 am – 9:00am) by the corresponding coefficient. We estimated this coefficient using historical data and its value was 0.422. The observed data collected from TTS is from 6:00 am to 9:00 am for the morning time. Since the simulated inter-zonal trips are for the morning peak time (7:00 am – 8:00 am), the observed data should be multiplied by the coefficient 0.422 to estimate the data for the peak hour in the morning. The same value for this coefficient is also defined in IMULATE.

Traffic assignment sub-module generates the amount of automobile flows for each link in the road network of the Hamilton CMA. The City of Hamilton provides the actual data about traffic flows for each link to compare with the simulated results. In this research, only 2006 traffic flows are collected from the City of Hamilton, and the data are only for 91 links in the city, including major roads and rural roads, but not for highways, see Figure 3.2. Since the Lincoln Alexander Highway is not available until 1997, the system does not count it into the simulation for 1991 and 1996. After 1996, this highway is added into the simulation automatically. The same as the calculation of inter-zonal trips, the observed data about the traffic flows for a 3-hour morning period should be multiplied by the coefficient 0.422 in order to estimate the flows in the morning peak hour.

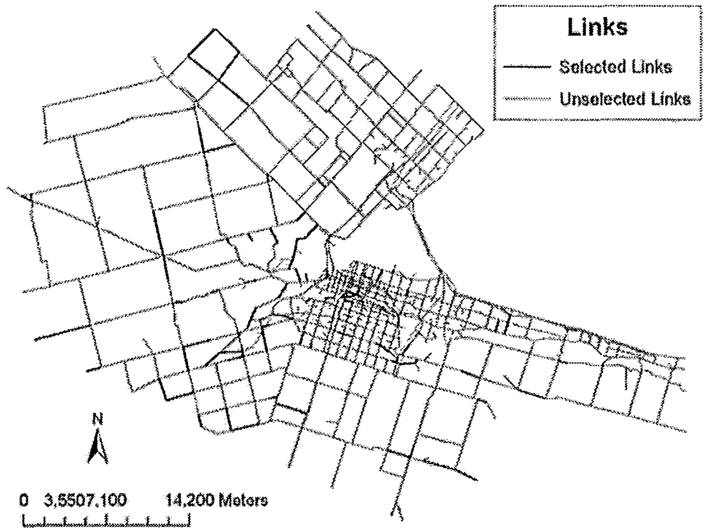



**Figure 3.2 Selected links in the road network**

### 3.1.3 Data for IMPACT

In IMPACT, the TAZs layer makes use of the 1996 census tract divisions, which include 163 zones. The road network layer includes 1,540 traffic links. IMPACT is composed of three main modules: the demographic module, the transportation module and the environmental module. This research only validates the first two modules. The observed data which are used to validate the demographic module and the transportation module are collected from Statistics Canada, TTS and City of Hamilton.

**Demographic module**

This module simulates the development of population across TAZs in the city by gender and age. Statistics Canada provides the observed population data by age and gender to test the estimates produced through simulations with this module. Since the calibration of the demographic module is based on data for the time period 1996-2001, a simulation was performed for the time period 2001-2006. The simulated results for 2006 are then compared with observations reported by Statistics Canada for the year 2006. The 2006 data collected from Statistics Canada are based on 177 census tract divisions, while the simulated data use a system of 163 zones. Thus, prior to the validation, the observed data should be cleaned and transformed into the corresponding 163 TAZs in IMPACT. Actually, only 162 zones' data are used in this research, since the simulated result in the zone with ID 142 is lost during the simulation process. This zone is located in the downtown area of Hamilton and some of its related utility variables are lost in the starting database.

**Transportation module**

The four-stage UTMS model is adopted in this module. The first three stages (trip generation, trip distribution and modal split) generate O-D matrices for different transportation modes, while the last stage (traffic assignment) simulates the distribution of automobile traffic flows across all the roads in the city. TTS provides the data about the auto driver trips between each O-D pair across TAZs for the morning period by three travel purposes: work, school and discretionary. Actually, IMPACT splits the simulated results by work purpose and non-work purpose. Thus, for the purposes of our validation, we should group the observed school and discretionary trips into one group, called non-work trips.

Since IMPACT simulates the development based on 163 TAZs, the observed data should be transformed into the same number of zones as the simulated results. The module predicts the traffic on the transportation network for four periods of a typical day:

morning, day, afternoon and evening. The system estimates the results for the peak hour in each period using the total value of each period multiplied by the corresponding coefficient, as shown in Table 3.2. TTS data is used to validate this module.

**Table 3.2 Coefficients to estimate the peak hour trips for different periods of a typical day**

| Period | Morning | Day | Afternoon | Evening |
|---|---|---|---|---|
| *Time Interval* | *06:00 – 08:59* | *09:00 – 15:59* | *16:00 – 18:59* | *19:00 – 05:59* |
| Coefficient | 0.37 | 0.15 | 0.34 | 0.18 |

The automobile flows for each link of the road network in the city are simulated by the traffic assignment sub-module. The City of Hamilton provides the observed data to compare with the simulated results. The observed data is only for the same 91 road links selected in the validation work of IMULATE, should be multiplied with their corresponding coefficients to estimate the flows in the peak hour of each period.

## 3.2 Methodologies

### 3.2.1 Overview

The validation work of IMULATE and IMPACT tests whether their predictions match actual urban dynamics. The overall process is composed of two basic methodologies with three main techniques. Several parameters are used to assess the statistical analysis results.

**Two methodologies**

Historical validation and sensitivity analysis are two basic methodologies often used to evaluate and assess simulation models. Historical validation entails a comparison of model predictions against observed data. In this research, various techniques are utilized for the historical test. The following section will introduce those techniques.

The usual meaning of sensitivity analysis is to test whether the simulated results would be sensitive to certain parameters in the model or to exogenously observed input to the model. However, the sensitivity analysis method is only used in this thesis in order to improve the exogenous input of the system.

**Three techniques**

In the overall process, the comparison between two groups of data, mostly the simulated and observed data, is essential. There are three basic techniques utilized in the comparison work: statistical, graphical and GIS visualization.

The statistical method is an objective approach for quantitative analysis. The vital element is to use quantitative testing methods to assess whether the simulated results are close to what is observed. We often use the percentage difference between the simulated and observed values to indicate the accuracy. For performance evaluation of models, linear regression, chi-square test and some information theoretic statistics are adopted to determine the association between the simulated and observed data. In this research, linear regression is utilized and combined with a graphical comparison technique. For large data sets, such as a trip matrix, a combination of those three methods is employed. In addition to the standard Pearson chi-squared test, Neyman-Pearson hybrid chi-squared test and Freeman-Tukey chi-squared test are adopted when a large number of zeros are present in the simulated and observed values (Voas et al., 2001). Some information theoretic methods, including the algebraic percentage error (ALPE) and the standardized root mean square error (SRMSE), provide useful indices to test the degree of difference between the simulated and observed values (Rodier, 2005; Fotheringham et al., 1989). In addition, we adopt Wilson's Entropy concept in this research to test the dispersion level for O-D matrices.

Graphical methods are used to compare the simulated results with observations. In particular, scatter plots and line charts are used in this research. Scatter plots display numerical values along the x and y axes, combining them into single points in a graph. We use scatter plots to represent the relationship between the simulated results and its related observed values. Line charts distribute the ordinal data along the x axis and all the numerical values along the y axis. We use line charts to depict the changes over time intervals.

GIS technologies are useful for the analysis of geographical data. Since the outputs of IMULATE and IMPACT have spatial characteristics, they can be displayed in thematic maps and analyzed visually using GIS software, such as ArcGIS. An important advantage of this method is that it integrates the results of statistical analysis with their spatial information, helping us to analyze the spatial distribution of errors between observed and estimated values.

**Some parameters**

The statistical test provides many parameters to determine the performance of the simulation. Table 3.3 lists the main parameters used to explain the testing results.

**Table 3.3 Definitions of statistical parameters**

| Methodology | Parameter | Description |
|---|---|---|
| Linear regression | $\beta_1$ | Slope in linear regression |
| | $\beta_0$ | Intercept in linear regression |
| | $R^2$ | Proportion of variation of dependent variable explained |
| Chi-squared tests | $X_H^2$ | Neyman-Pearson hybrid chi-squared test |
| | $FT^2$ | Freeman-Tukey chi-squared test |
| Other tests | SRMSE | The standardized root mean square error |
| | ALPE | The algebraic percentage error |
| Entropy | ENT | The entropy value of a matrix |

The correlation between simulated results and observed data can be constructed as two variables in a simple regression model, see Equation 3.1.

$$O_i = \beta_1 S_i + \beta_0 \tag{3.1}$$

The independent variable $S_i$ is the simulated result in zone i, while the dependent variable $O_i$ is the observed value in zone i. In a perfect model, the simulated results should be equal to the corresponding actual value, which means the parameter $\beta_1$ should be 1, and the parameter $\beta_0$ should be 0. In other words, if we draw those simulated-observed (S-O) value pair in a scatter plot graph, a perfect simulation would have all the S-O points located on the linear regression line y = x. In addition, the regression generates a parameter $R^2$ determining the strength of relationship between the simulated results and the observed value. The upper limit of $R^2$ is 1, which represents the highest precision of the simulation. Thus, the evaluation of a model depends on how close $\beta_1$ is to 1, how close $\beta_0$ is to 0, and how close $R^2$ is to 1.

Voas and Williamson (2001) introduced various tests to evaluate the model's "fit" based on the $x^2$ distribution. The most commonly used method is the Pearson statistic. The classic Pearson chi-squared test has a weakness that zero cannot exist in simulated and observed values. In this research, chi-squared tests are employed to assess the simulation of origin-destination (O-D) matrices, which typically contain a large number of zeros. Thus, a new hybrid form developed by Neyman is utilized here.

$$X_H^2 = \sum_i \sum_j \frac{(O_{ij} - S_{ij})^2}{AVE_{ij}} \tag{3.2}$$

Where $S_i$ is the simulated trips between origin i and destination j, while $O_i$ is the observed trips between origin i and destination j. $AVE_{ij}$ is the average of $O_{ij}$ and $S_{ij}$, or 1 if both $O_{ij}$ and $S_{ij}$ are 0. Like the classic method, $X_H^2$ can also be tested with the $x^2$ distribution. The other statistic used to validate the O-D matrices is Freeman-Tukey, which can test the data with zeros as well.

$$FT^2 = 4\sum_i \sum_j (\sqrt{O_{ij}} - \sqrt{S_{ij}})^2 \tag{3.3}$$

Besides the statistical methods mentioned above, there are some simple measures to assess errors. The easiest way is to calculate the algebraic percent error (ALPE) for each zone (Rodier, 2005).

$$ALPE_i = \left(\frac{S_i - O_i}{O_i}\right) * 100 \tag{3.4}$$

For the spatial characteristics of data, the distribution of simulated errors can be displayed on the map with ALPEs. This statistic is only adopted at the zone level, and should be used with other methodologies. In order to evaluate the overall simulated errors, a meaningful measure is the standardized root mean square error (SRMSE), calculated as:

$$SRMSE = \frac{\sqrt{\sum_i \sum_j (O_{ij} - S_{ij})^2 / N}}{\sum_i \sum_j O_{ij} / N} \tag{3.4}$$

Where $S_{ij}$ is the simulated trips between origin i and destination j, $O_{ij}$ is the observed trips between origin i and origin j, and N is the total number of O-D pairs. SRMSE has a lower limit of zero, which indicates a perfect accuracy of simulations, and its value is usually below 1.0 (Fotheringham et al., 1989). Smaller SRMSE values correspond to higher simulated quality of the model. If the average simulated error were greater than its actual mean, the corresponding value of SRMSE would be larger than 1.0, meaning the model has poor performance. Sometimes, this standardized statistic is preferred to the non-standardized way, which has high sensitivity to the large variance of the variable involved (Knudsen et al., 1986; Black, 1973).

Examples of the use of this statistic in the test of models include Thorsen et al.'s (1998) use to compare a competing destinations model with the traditional gravity model; the use by Celik (2004) to test the simulated quality of an artificial neural network model combined with $R^2$ method; Hu et al.'s (2002) use to test the performance levels of competing destinations models during the calibration process; the use by Wilmot et al. (2006) to discriminate between two destination choice models, the gravity model and Intervening opportunity model; Mozolin et al.'s (2000) use to evaluate the simulated quality of a doubly-constrained model and a neural network model by the commuter flows.

We use the aforementioned methods to test O-D matrices by assuming every O-D pair is independent. However, each O-D pair has its own spatial relationship with the other pairs. Thus, we adopt the entropy parameter to assess the similarity of trips' distributions over zones. Equation 3.5 shows the calculation of the entropy.

$$ENT = (T \ln T - T) - \sum_i \sum_j (T_{ij} \ln T_{ij} - T_{ij}) \tag{3.5}$$

Where T is total number of trips, and $T_{ij}$ is the number of trips between origin i and destination j. We calculate the entropy values for simulated O-D matrices and their related observed O-D matrices individually. The differences of trips' distributions are determined by the corresponding differences of entropy values.

### 3.2.2 IMULATE validation

IMULATE is composed of three modules, including the land use module, transportation module and environmental module. The accuracy of those modules is tested by various methodologies.

**Land use module**

The land use module consists of the POPMOB sub-module and the EMPLOC sub-module. The main function of POPMOB is to simulate the development of the population's distribution across the city. Since the population is calculated as the product of the number of households and their size, the test to ascertain whether the simulated results can reflect the actual distribution of the households is used to validate this sub-module.

The first step to determine the quality of POPMOB is to compare the simulated total number of households with its corresponding actual value by their percentage difference. In POPMOB, the simulation of households' growth at city level depends on the observed migration and household formation propensities in the city. The distribution of households across TAZs is estimated with the multinomial logit model (MNL). To test the performance of the MNL location choice model, the simulated number of households in each TAZ is compared with its related observed data, determining whether the simulated results can reflect the real distribution of households correctly. The simulated errors of households' distribution can be indicated by ALPE that is used to test the simulated percentage error in each TAZ. Thus, using some visualization GIS software, the spatial distribution of the simulated errors can be displayed according to their associated spatial locations on the map of Hamilton CMA.

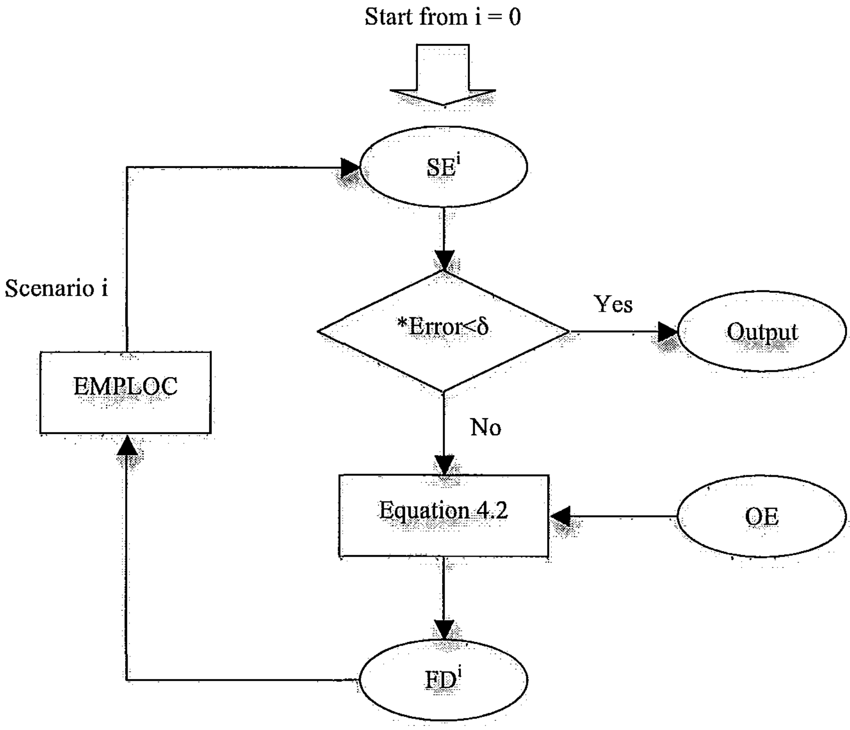
Furthermore, the relationship between the simulated and observed households at TAZs level plays an important role to assess the simulated accuracy of MNL. The integration of linear regression and graphical comparison method is used in this research. The slope ($\beta_1$), intercept ($\beta_0$) and r-squared ($R^2$) value are important parameters to determine the accuracy of MNL location choice model.

An important note regarding the simulation should be mentioned here. The original system uses the projected new dwellings as exogenous input to calculate the utility in the MNL location choice model. In order to provide a fair comparison, the observed data about new dwellings provided by Statistics Canada is utilized as the exogenous input to the model.

The EMPLOC sub-module simulates the distribution of employment at two levels: city level and TAZs level. The input-output (I-O) model is used to estimate the overall employment in the city, which can be assessed by the percentage difference between the simulated and observed values. The I-O model itself is affected by some exogenous factors. A key exogenous input used here is the final demand in each sector of the economy. We projected the original exogenous final demands using 1986 final demands. Thus, we should recalculate a new group of final demands to make the simulated results closer to the observed values. If the new final demands can produce a more accurate prediction of employment' growth in the overall city, the rest of the simulation errors will occur in the simultaneous auto regressive (SAR) model and MNL location choice model. The sensitivity analysis method is used to determine what changes of final demands can produce better results. In this research, we use Equation 3.6 to estimate the new final demands. In this equation, the type of economic sector is identified using xxxx, 4-digital 1980 Standard Industrial Classification (SIC) code. The whole process to estimate the final demands is like the iterations of a loop. The generated results from a loop will be used as the input to the next loop. Each of the iterations is constructed as a scenario owning different exogenous inputs.

$$FD_{xxxx}^{i+1} = FD_{xxxx}^{i} \cdot \frac{OE}{SE^{i}} \tag{3.6}$$

Where $FD^{i}_{xxxx}$ is the final demand in economic sector xxxx used for scenario i, OE is the observed number of employments, and $SE^{i}$ is the simulated number of employments generated in scenario i. Scenario i is the iteration i in Figure 3.3.




* The simulated error is determined by the percentage difference between the simulated and observed employments. δ is defined by the user as the error limit. As δ decreases, the higher accuracy of EMPLOC increases.

**Figure 3.3 The diagram to estimate the new final demands in EMPLOC**

This flowchart is calculated from the base scenario (scenario 0) using the original final demands as the input. After that, two options are generated based on the comparisons between the simulated errors with the user defined error limit (δ). If the simulated error is lower than the error limit, the final demands used in this scenario will be the most appropriate value that will make the input-output function produce optimal results. Otherwise, the iterations continue. The generated results in this scenario will be used in the next scenario, calculating a new group of final demands. The whole process is a loop, and the best results can be obtained when the overall process reaches equilibrium.

In practice, this method is challenging to implement, since it could require hundreds of iterations for the overall process to reach equilibrium. In this research, we only use four scenarios to produce new final demands. Though the best results are not obtained here, the new final demands can produce better results than the original ones.

Since the observed employments in 1996, 2001 and 2006 are available, three new groups of final demands for 1996, 2001 and 2006 can be obtained. Equation 3.7 can be used to estimate the final demands for the other simulated periods.

$$FD^{t}_{xxxx} = \lambda_{xxxx} \cdot FD^{t-1}_{xxxx} + \varepsilon_{xxxx} \tag{3.7}$$

Where $FD^{t}_{xxxx}$ is the final demand in economic sector xxxx for simulation at time t; $\lambda_{xxxx}$ and $\varepsilon_{xxxx}$ are the coefficients used to estimate the final demand in economic sector xxxx. $\lambda_{xxxx}$ and $\varepsilon_{xxxx}$ can be calculated using the new group of final demands in 1996, 2001 and 2006.

On the other hand, the distribution of employment at TAZ level is estimated by a simultaneous auto regressive (SAR) model that predicts the number of firms lost from each zone, and a multinomial logit (MNL) model that predicts the location choice probability of a relocating or a newly established firm. Like POPMOB, the spatial distribution of simulated errors (e.g. ALPE) is displayed visually on the map using GIS software. In order to ascertain whether those two models can predict the distribution of employment correctly, the linear regression method with graphical comparisons can be employed to assess their accuracy.

**Transportation module**

This module includes two main sub-modules: TRANDEM and traffic assignment. TRANDEM predicts the trips occurring between each pair of TAZs during the morning peak time. The trips are generated directly from the aforementioned land use module. Using the Fratar trip distribution model, the origin-destination (O-D) matrices are generated based on the travel purposes of residents, including work, school and discretionary trips. The MNL modal split model divides all the trips into four groups according to their transportation modes: private automobile driver, private automobile passenger, public transit and walk.

Thus, the observed O-D matrices are used to assess TRANDEM's simulated quality by comparison with the forecast results. The simulated error is indicated by the percentage difference between the simulated and observed value. In order to test the accuracy of distributing all the generated trips over all O-D pairs, the linear regression method is an important measure to determine the correlation between simulated and observed trips for each O-D pair. Besides that, $X_H^2$, $FT^2$ and SRMSE are also very important statistics to test whether the predicted results are close to reality.

The traffic assignment sub-module simulates the traffic flows on the road network in the morning peak time, based on the automobile trips generated by TRANDEM sub-module. The linear regression method is used to validate this sub-module combined with graphical comparisons. The overestimated and underestimated links are symbolized on the map using GIS technologies.

### 3.2.3 IMPACT validation

IMPACT is composed of three modules: the demographic module, transportation module and environmental module. Different methods are used to test different modules.

**Demographic module**

The demographic module predicts the population size by gender and age at the municipal and TAZ levels. The overall growth of population for each municipality is estimated by Rogers' multiregional demographic model (RMDM), while the spatial distribution of population across TAZs is estimated by an aggregated spatial multinomial logit (ASMNL) model. The simulated errors are calculated with the percentage difference between the simulated results and their corresponding observed values. The line chart is utilized to depict the population's development over different age intervals. Its corresponding accuracy is verified using the graphical comparisons between the predicted development and the actual situation.

**Transportation module**

The transportation module uses a four-stage UTMS model to simulate the travel behavior between different TAZs. This module generates two main results: generated trips for each O-D pair and automobile flows attributed to the city's road network. The inter-zonal trips are included in the O-D matrices, which are produced according to cohorts (adult and elder), trip purposes (work and non-work) and periods of a typical day (morning, day, afternoon and night). The MNL mode choice model splits those matrices into three transportation modes: auto driver, auto passenger and other modes. The simulated accuracy for O-D matrices is tested using linear regression, chi-squared test and some information theoretic methodologies. The parameters produced by those statistical methods are very useful in determining the simulated ability of trip generation, trip distribution and modal split models.

The traffic assignment model generates the number of automobile flows for each road in the city. It is validated with the traffic data for 91 selected links in the city road network. The similarity between the simulated results and the observed value is assessed with the linear regression method. The assessment process is combined with graphical comparison methods.

# Chapter 4. Results and Discussion

## 4.1 Validation results and discussion for IMULATE

### 4.1.1 Land use module

**POPMOB**

Table 4.1 shows a comparison between the number of households generated by POPMOB and the data from Statistics Canada. Overall the total number of households is overestimated by 1.87% in 1991, which means the simulated total number of households in 1991 is close to the real value. The overestimated rate of the total households in the city increases over time, until it reaches 2.07% in 2001. Relative to the simulation in 2001, the number of households in 2006 is overestimated by 4.04%, twice as much as the rate in 2001. Since the simulation for each year is processed based on the results from the previous simulation, the error propagation causes POPMOB to systematically overestimate the number of households over time.

The test of households' growth prediction over simulated years amplifies the simulated errors, as shown in Table 4.1. The overestimated rates in 1996 and 2001 are within acceptable levels. Much like the comparisons of total households, the forecast number of households' growth in 2006 is larger than its actual number.

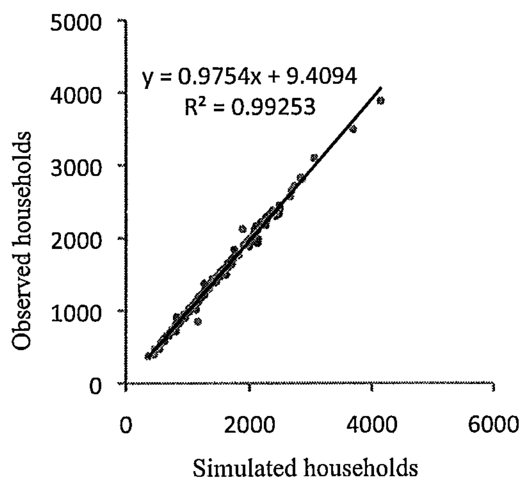
**Table 4.1 Simulated households vs. observed households over years**

| | 1991 | 1996 | 2001 | 2006 |
|---|---|---|---|---|
| *Households* | | | | |
| **Observed Households** | 215,060 | 227,610 | 240,360 | 248,875 |
| **Simulated Households** | 219,088 | 232,188 | 245,334 | 258,931 |
| ***% Difference*** | *1.87* | *2.01* | *2.07* | *4.04* |
| | | **1991-1996** | **1996-2001** | **2001-2006** |
| *Growth of households* | | | | |
| **Observed Change** | | 12,550 | 12,750 | 8,515 |
| **Simulated Change** | | 13,100 | 13,146 | 13,597 |
| ***% Difference*** | | *4.38* | *3.11* | *59.68* |

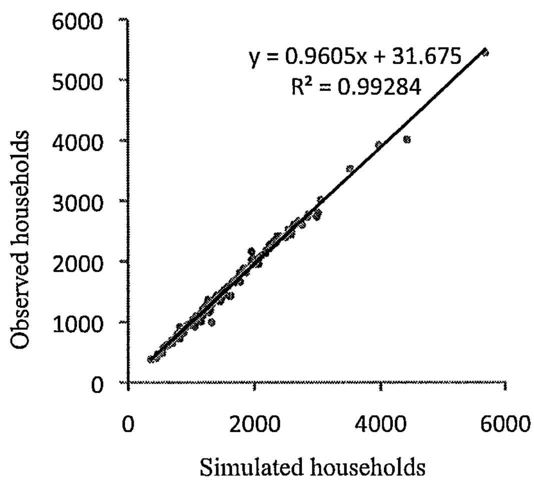
Thus, though the performance of POPMOB gets worse over time, it can predict the change of households at the city level. The reason that the accuracy of the simulated results decline in 2006 relates to the occupied rate of new constructed dwellings in 2006. POPMOB relies on the assumption that the number of vacant dwelling units in each zone remains constant, therefore all newly constructed dwellings are considered occupied. By dividing the number of new occupied dwellings by the total number of new constructed dwellings from Statistics Canada, we found that the occupied rate of new dwellings in 2006 was close to 70%, while the rates in 1996 and 2001 were more than 95%. Thus, the low occupancy rate of new dwelling explains the simulated error in 2006.

At the census tract level, Figure 4.1 shows whether the households generated by POPMOB have the correct distribution over census tracts throughout the city for each forecast year. There are 146 simulated-observed pair points in each figure, each of which represents a TAZ based on 1986 census tract division in IMULATE. For each scatter point, its x coordinate value indicates the simulated number of households in this TAZ, while its y coordinate value indicates the corresponding observed number. The trend line is produced with the linear regression equation y = $\beta_1$x + $\beta_0$, whose independent variable is the simulated number of households in each TAZ and dependent variable is the observed number of households in each TAZ. If the simulation is perfect, all the scatter points should be on the trend line y = x. Thus, the accuracy of the simulation can be tested by determining how close the slope parameter $\beta_1$ is to 1, how close the interception parameter $\beta_0$ is to 0, and how close r-squared ($R^2$) value is to 1. $R^2$ value is an important constraint to test whether the destination choice model can simulate the distribution of households correctly.

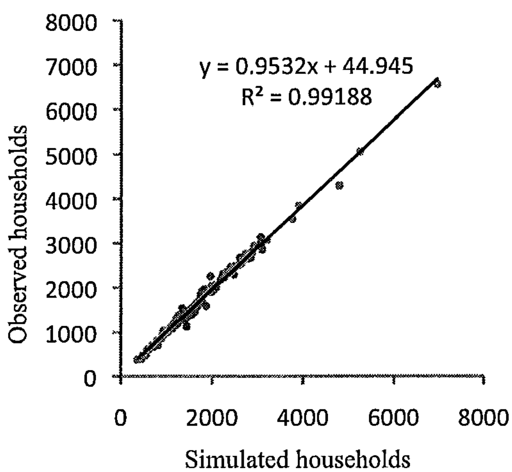
As shown in Figure 4.1, $R^2$ values in those four graphs are quite close to 1, and the slope parameter and the intercept parameter in the linear function are close to their perfect values. This demonstrates that the simulated spatial distribution of households is close to the actual distribution. However, the parameters in the linear equation deviate from their corresponding perfect values over time because of the error propagation.



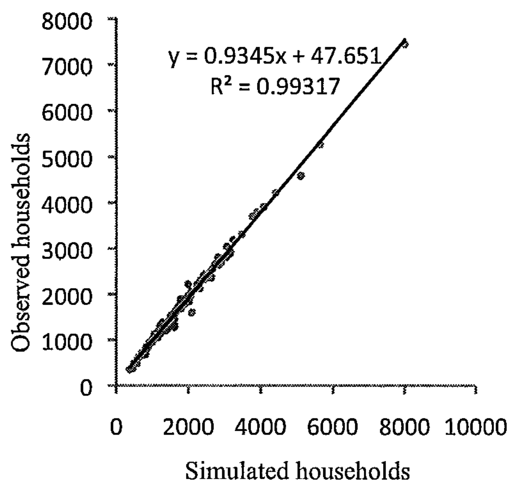
(a)



(b)




(c)




(d)

**Figure 4.1 Simulated households vs. observed households across all TAZs in the city for 1991(a), 1996(b), 2001(c) and 2006(d)**

Figure 4.2 shows the difference between simulated growth and observed growth in each TAZ. Comparisons between the growth of households, rather than the number of households, make the simulation errors more readily apparent. As shown in the three graphs, the $R^2$ values are all very high, which means the simulated growth has high correlation with the observed value. However, the correlation is decreasing over time by the propagated errors from the previous simulated period. The $\beta_1$ and $\beta_0$ in the linear function also indicate the simulation ability of this module has high accuracy. Those graphs also indicate that the system hardly simulates negative growth of households, while the decreasing of households often occurs in some zones in the real world.

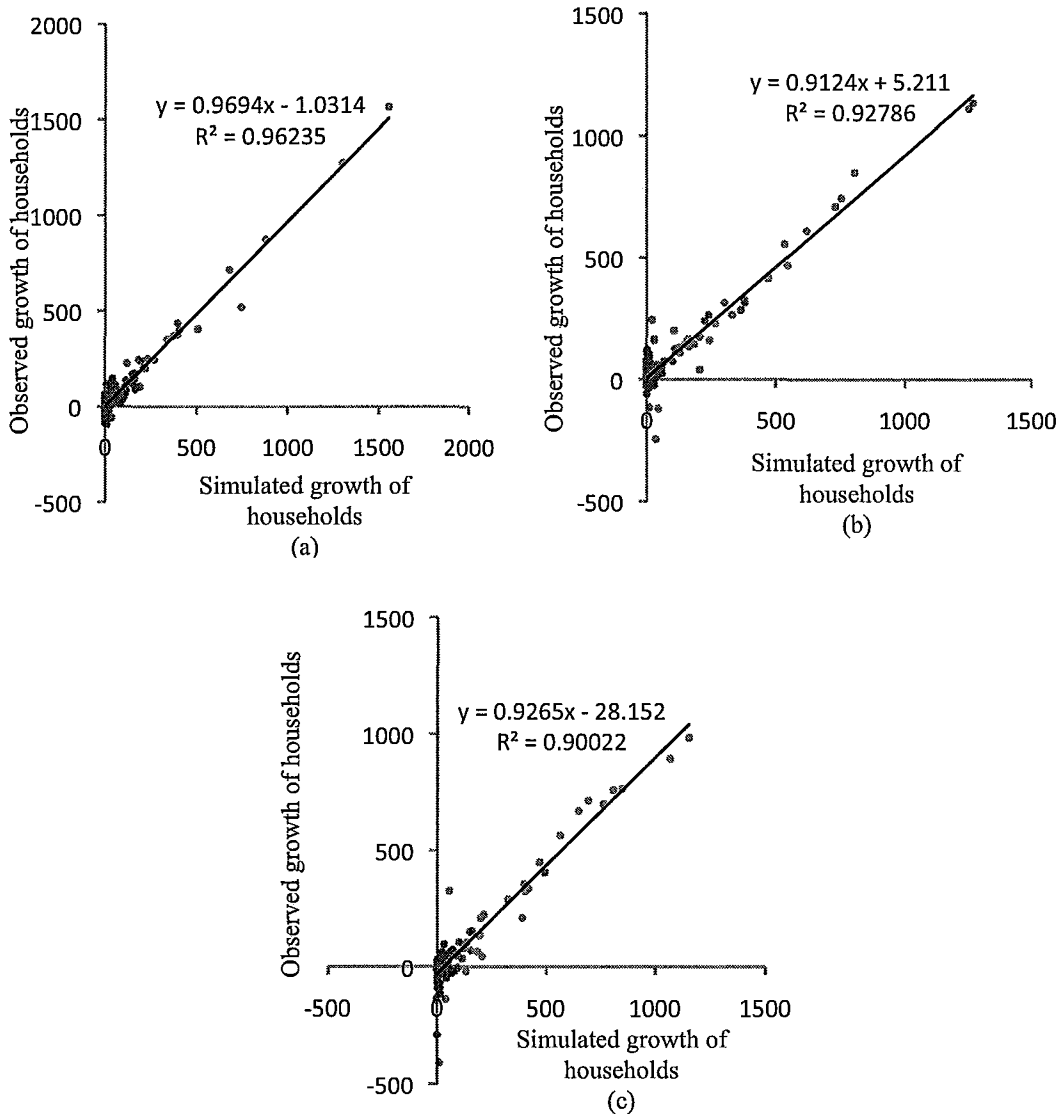



**Figure 4.2 The simulated growth of households vs. the observed growth of households across all TAZs in the city, from 1991 to 1996(a), 1996 to 2001(b) and from 2001 to 2006(c)**

ALPE is employed to determine the simulated error. The spatial distribution of simulated errors over TAZs throughout the city is displayed in Figure 4.3. In the 1991 simulation, though POPMOB overestimates the number of households in most TAZs, the overestimation rates are mostly relatively low. This overestimation rates are similar in

1996 and 2001, but more TAZs have higher overestimation rates than in 1991. In contrast to those three simulated years, the overestimation spreads throughout most of the city in 2006, except several areas in Glanbrook, east Dundas, centre Burlington, south Stoney Creek and center Hamilton downtown. The reason is the aforementioned low occupied rate of new dwellings in 2006.


**Figure 4.3 The spatial distribution of ALPE for household across all TAZs in the city, for 1991(a), 1996(b), 2001(c) and 2006(d)**

As shown in Figure 4.3, the number of households in industrial areas (TAZs with ID 1531 and 1532) and the highway area (TAZ with ID 1536) are consistently overestimated. As mentioned above, the proximity to industrial areas plays an important role in the simulation of households' distribution across the city. Areas with high proximity to industry are not appealing to people. Though the predicted number of households in the industrial areas stays constant over time, it is still much higher than the real number. The critical part of this problem is the estimation of access to industries. The residential area near the Queen Elizabeth Way (QEW) also has consistent overestimation of households, which may be affected by the mobility utility of residents there. The predicted numbers of households in this area increase over time. In contrast, the observed number of households increases some years and decreases in others, varying around an average value.

**EMPLOC**

As mentioned above, the overall number of employments in Hamilton is estimated by the input-output (I-O) model, which relates to its final demands input. Before testing the firms lost model and firms gained model, the I-O model should be improved to make its results closer to the actual data. Table 4.2 shows four groups of new final demands, each of which generates a new group of employment data.

**Table 4.2 Five scenarios used to estimate final demands in dollars**

| | Base | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 4 |
|---|---|---|---|---|---|
| ***1996 period*** | | *1.081*[a] | *1.052* | *1.032* | *1.019* |
| **Manufacturing** | 2673980 | 2890318 | 3039721 | 3137840 | 3198378 |
| **Construction** | 390147 | 421712 | 443510 | 457826 | 466659 |
| **Wholesale Trade** | 263115 | 284402 | 299103 | 308758 | 314715 |
| **Retail Trade** | 592505 | 640442 | 673547 | 695288 | 708702 |
| **All Services** | 2363256 | 2554455 | 2686497 | 2773214 | 2826717 |
| ***2001 period*** | | *1.213* | *1.129* | *1.074* | *1.041* |
| **Manufacturing** | 2673980 | 3243770 | 3661147 | 3931932 | 4094486 |
| **Construction** | 423309 | 513511 | 579584 | 622451 | 648185 |
| **Wholesale Trade** | 285480 | 346312 | 390872 | 419782 | 437136 |
| **Retail Trade** | 642868 | 779855 | 880199 | 945300 | 984381 |
| **All Services** | 2564133 | 3110516 | 3510747 | 3770409 | 3926285 |
| ***2006 period*** | | *1.313* | *1.192* | *1.110* | *1.061* |
| **Manufacturing** | 2319975 | 3045012 | 3628910 | 4027951 | 4272513 |
| **Construction** | 398485 | 523019 | 623311 | 691852 | 733858 |
| **Wholesale Trade** | 268739 | 352725 | 420362 | 466586 | 494915 |
| **Retail Trade** | 605169 | 794296 | 946607 | 1050698 | 1114492 |
| **All Services** | 2413767 | 3168116 | 3775620 | 4190794 | 4445243 |

a. Coefficient to estimate final demand, calculated by observed employment and simulated employment in the previous scenario

Each scenario is generated based on the output of the last scenario. As shown in Table 4.3, four new groups of simulated results are compared with their corresponding observed values. The simulation using the original final demands predicts 7.85% fewer jobs than the actual progress in 1996. Furthermore, the underestimation error grows over time, until it reaches 24.42% in 2006. Through the enhancements to final demands, the simulated results from the scenarios improve over the base case. In the scenario 4, the model only underestimates the employments by 1.27% in 1996, 2.64% in 2001 and 3.78% in 2006, which means the final demands used in scenario 4 increase the quality of the results obtained from the model.

**Table 4.3 Simulated employment vs. observed employment over years**

| | Total Employment | | | Growth of Employment | |
|---|---|---|---|---|---|
| | **1996** | **2001** | **2006** | **1996-2001** | **2001-2006** |
| **Observed** | 211,970 | 251,725 | 275,145 | 39,755 | 23,420 |
| **Base** | 195,910 | 206,615 | 207,951 | 10,705 | 1,336 |
| *% Difference* | *-7.58* | *-17.92* | *-24.42* | *-73.07* | *-94.30* |
| **Scenario 1** | 201,365 | 222,198 | 229,240 | 20,833 | 7,042 |
| *% Difference* | *-5.00* | *-11.73* | *-16.68* | *-47.60* | *-69.93* |
| **Scenario 2** | 205,161 | 233,599 | 246,275 | 28,438 | 12,676 |
| *% Difference* | *-3.21* | *-7.20* | *-10.49* | *-28.47* | *-45.88* |
| **Scenario 3** | 207,797 | 240,896 | 257,725 | 33,099 | 16,829 |
| *% Difference* | *-1.97* | *-4.30* | *-6.33* | *-16.74* | *-28.14* |
| **Scenario 4** | 209,277 | 245,068 | 264,745 | 35,791 | 19,677 |
| *% Difference* | *-1.27* | *-2.64* | *-3.78* | *-9.97* | *-15.98* |

The simulated growth of employment from 1996 to 2001 and from 2001 to 2006 is compared with their related observed values in Table 4.2. The comparison of employment growth is used to amplify the simulated error. The growth comparisons show significant improvement in the model from the base scenario to the last scenario. Though the base scenario shows poor comparison results of 73.07% and 94.30% underestimation, this rate is reduced to less than 16% in scenario 4.

Since scenario 4 generates the best results, the comparison between simulated distribution of employment across the city in scenario 4 and its corresponding observed value is used to test the simulation ability of the firms lost model and firms gained model. As shown in Figure 4.4(a), 1996 simulated results are highly correlated with the related observed values, since the $R^2$ value is 0.87, which is close to 1. The slope parameter in the linear equation is about 0.82, also close to its perfect value of 1. Compared with the large employment in most areas, the intercept in the linear equation is acceptable. Thus, EMPLOC can simulate the distribution of employment in 1996 very well.

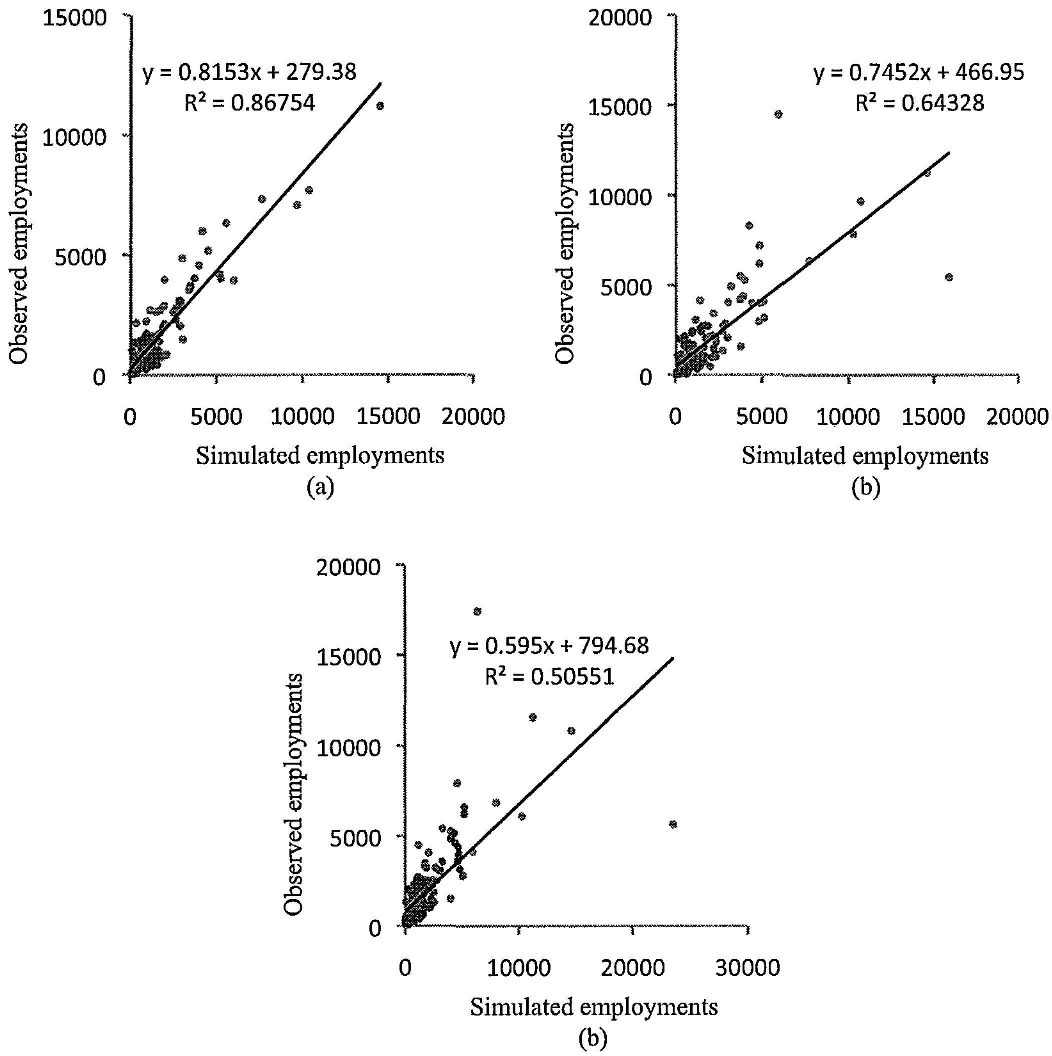



**Figure 4.4 Simulated employment vs. observed employment across all TAZs in the city for 1996(a), 2001(b) and 2006(c)**

In contrast to the 1996 simulation, the simulated accuracy of the employment distribution decreases over time. As shown in Figure 4.4(b) and 4.4(c), the $R^2$ value between the simulated and observed distribution is reduced to 0.64 in 2001 and 0.51 in 2006. In the linear equation, the slope value moves further from its perfect value of 1, and the intercept value is increasing, much higher than 0. The changes of those parameters indicate poor simulation quality of employment distribution in 2001 and 2006.

There are two excepted points found in those graphs. One is extremely underestimated, while the other is extremely overestimated. The underestimated point is the TAZ with ID 4126, which is located in the residential areas near the downtown of Burlington. Its special location and high density of population attracts more firms moving into this area than the model expects. The overestimated point is the TAZ with ID 1500, which is in the west of Hamilton downtown and near St. Josephs Hospital. Though St. Josephs Hospital is a large employer, it does not provide as many job opportunities as the model estimates. The misestimates of the employment in those two TAZs have great impact on the comparison results. As shown in Figure 4.5, without those two incorrectly estimated points, the comparison results become better.

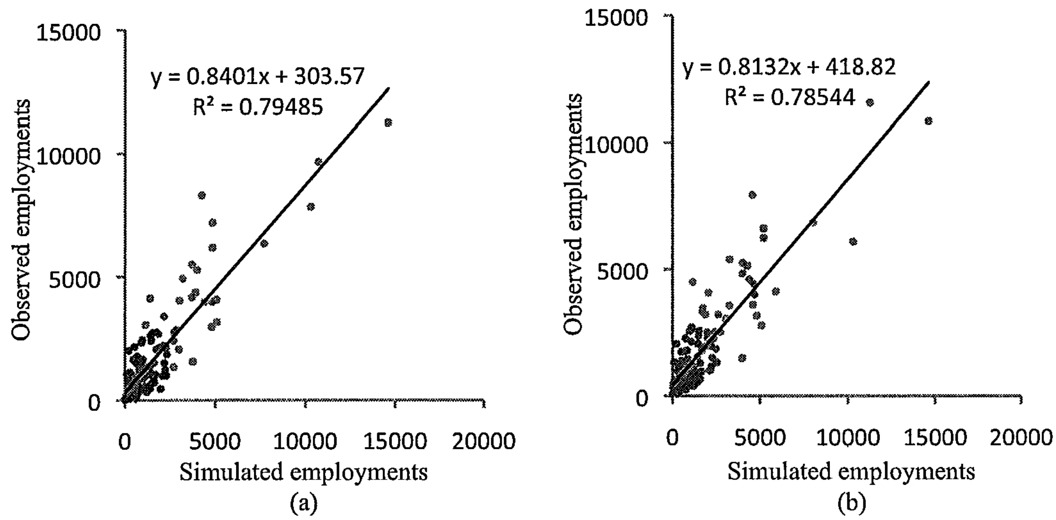



**Figure 4.5 Simulated employment vs. observed employment in the city except the TAZs with ID 4126 and 1500, for 2001(a) and 2006(b)**

The $R^2$ values increase to 0.79 in 2001 and 2006 by deleting the simulated results in TAZs with ID 4126 and 1500. The slope values in the linear equations have a corresponding increase, moving closer to 1, while the intercepts are slightly reduced. Though the results improve in this case, the simulation accuracy still decreases over the simulated years. The trend line of those simulated-observed (S-O) pair points deviate further from the perfect line y=x over time. The simulated results are losing correlation with the observed value by the decline of $R^2$ value. This kind of error propagation exists in many models.

ALPE is also employed to determine the simulated error. Figure 4.6 shows that all the TAZs are divided into three groups: underestimated rate below -20%, estimated rate within -20% and 20% and the overestimated rate above 20%. From those maps, it is evident that the major manufacturing areas are consistently overestimated over time. Some steel factories locate in these manufacturing areas and they could not provide as many job opportunities as the model expects. The bankruptcy of some factories makes a lot of workers lose their jobs.

EMPLOC does not simulate the employments in Burlington and Grimsby with the Input-output model. The employment in those two municipalities are simulated with the exogenous input. Thus, employment' simulation can be improved by changing the exogenous employments in those areas.


**Figure 4.6 The spatial distribution of ALPE for employment across all TAZs in the city, for 1996(a), 2001(b) and 2006(c)**

### 4.1.2 Transportation module

**TRANDEM**

The main outputs of the TRANDEM sub-module are the origin-destination (OD) matrices for the morning period of a typical day. The simulation is split by three trip purposes (work, school and discretionary) and four transportation modes (auto-driver, auto-passenger, transit and walk/bicycle).

In this research, the observed morning trips by auto drivers are compared with the simulated results, in order to test the simulated quality of TRANDEM. As shown in Table 4.4, the system far overestimates the morning work trips, and the overestimation is growing over time. The simulated trips are nearly three times more than the observed data. The system also overestimates the school trips and discretionary trips. The simulated growth of school trips over years is faster than the observed value, while the simulated growth of discretionary trips is slower than the observed value. Thus, the difference between simulated school trips and the corresponding observed value is increasing over time, while the difference between simulated discretionary trips and the related observed value is reduced.

**Table 4.4 Simulated trips vs. observed trips over years**

| | | | Work | School | Discretionary |
|---|---|---|---|---|---|
| **1991** | **Observed** | | 23,910 | 1,552 | 8,103 |
| | **Simulated** | | 70,714 | 2,413 | 17,982 |
| | | ***% Difference*** | *195.75* | *55.48* | *121.92* |
| **1996** | **Observed** | | 23,748 | 1,604 | 9,685 |
| | **Simulated** | | 77,450 | 2,345 | 20,120 |
| | | ***% Difference*** | *226.13* | *46.20* | *107.74* |
| **2001** | **Observed** | | 24,413 | 1,501 | 13,035 |
| | **Simulated** | | 82,970 | 2,372 | 22,280 |
| | | ***% Difference*** | *239.86* | *58.03* | *70.92* |
| **2006** | **Observed** | | 21,566 | 1,182 | 14,509 |
| | **Simulated** | | 87,344 | 2,321 | 24,158 |
| | | ***% Difference*** | *305.01* | *96.36* | *66.50* |

The total trips are estimated by the trip generation model. In IMULATE, we use the average number of workers in different type of household to estimate the total workers. The overestimates of the average number of workers in each type of household generate the overestimates of the total workers, which finally produce the overestimates of total trips. Since the simulated total number of trips is far larger than the actual data, the TRANDEM sub-module has poor simulation quality. Thus, we do not need to test the distribution of trips over O-D pairs in the O-D matrices.

### Traffic assignment

The generated automobile trips from TRANDEM are distributed onto each link of Hamilton's road network with the traffic assignment sub-module. The observed traffic flows in 93 selected links are compared with the simulated results, validating the simulation quality of this sub-module.

As shown in Table 4.5, the system highly overestimates the overall traffic flows in those 91 links, about three times more than the actual number. The estimation of traffic flows for each link is also poor.

**Table 4.5 Simulated traffic flows vs. observed traffic flows in 91 selected links, for 2006**

| | Observed | Simulated | Overestimated | % Overestimated |
|---|---|---|---|---|
| **2006 Traffic Flows** | 63,795 | 213,500 | 149,704 | 234.66 |

The $R^2$ value in Figure 4.7 indicates low correlation between the simulated and observed flows for each link, and the low slope value in the linear equation indicates that the predicted number of traffic flows for each link has a huge difference from the observed number. This means that the simulated distribution of traffic flows is almost totally different from the actual situation.

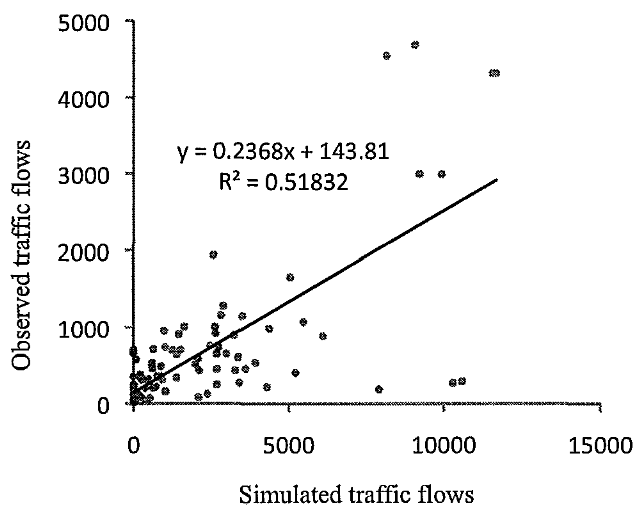
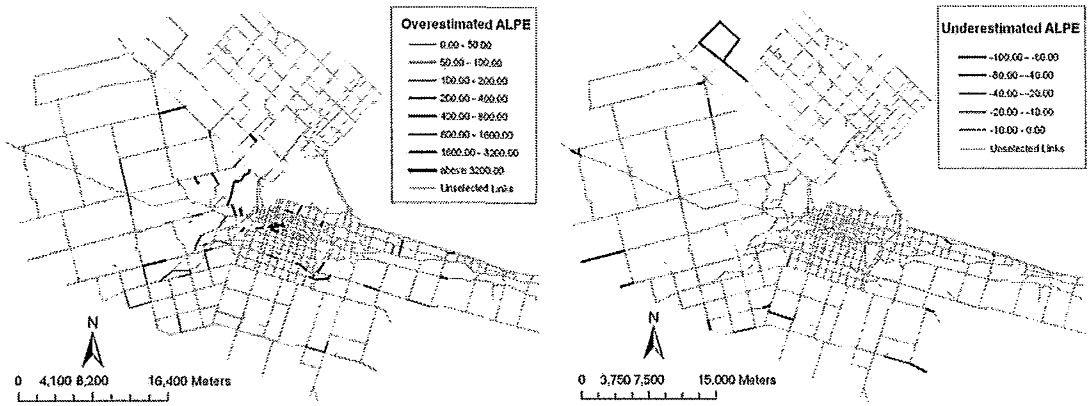



**Figure 4.7 Simulated traffic flows vs. observed traffic flows over all TAZs in the city, for 2006**

Thus, the system overestimates the number of traffic flows for some links, and underestimates for others. The overestimated links and underestimated links are displayed as maps in Figure 4.8. Obviously, most roads in Hamilton are overestimated, and the other underestimated roads are located in rural areas.




(a) (b)

**Figure 4.8 Overestimated links and underestimated links for 2006**

There are two sources generating this simulation error. One is the inaccurate input of this sub-module that is generated by TRANDEM, and the other is the traffic assignment sub-module itself. Since many other researchers have tested the SUE traffic assignment model, the main reason affecting the simulated accuracy should be from the error propagation of TRANDEM. Since IMPACT also uses SUE traffic assignment model, we will validate this model again in the validation work of IMPACT.

## 4.2 Validation results and discussion for IMPACT

### 4.2.1 Demographic module

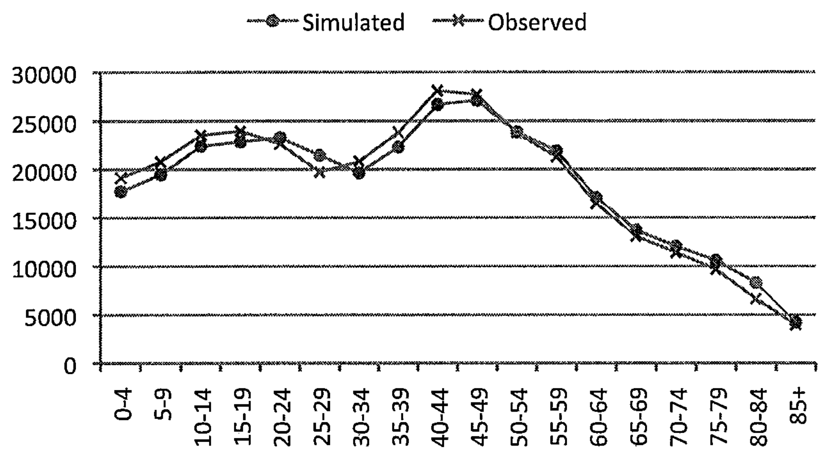
The demographic module is designed to simulate the development of population across all the TAZs in Hamilton CMA. The forecast population is divided by gender and follows a 5-age incremental simulation., The simulated population appears to be close to the observed. Table 4.6 shows that the system only underestimates the total population by 1.52%.

**Table 4.6 Simulated population vs. observed population for 2006**

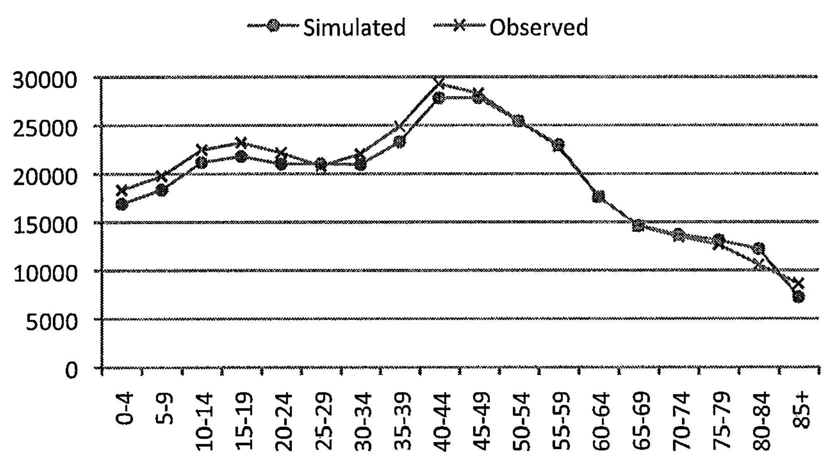
| | Male | Female | Total |
|---|---|---|---|
| **Observed** | 336,335 | 356,545 | 692,880 |
| **Simulated** | 334,554 | 346,774 | 682,315 |
| *% Difference* | *-0.53* | *-2.74* | *-1.52* |

For the male population, this module generates accurate results, which is only 0.53% lower than the real data. The simulated results for female seem worse than male, whose underestimated rate reaches 2.74%. Though the overall simulation has high accuracy, the simulation for female is a key reason for reducing the overall simulated accuracy. It means that, beside the factors that have been considered in RMDM, there are more personal attributes affecting the growth of female population than male population.

Figure 4.9 shows the distribution of population over 5-year age intervals, for both males and females. As shown in these two charts, though the system incorrectly estimates the population slightly for each 5-year interval, the overall development trend is similar to the actual distribution, which means the system can simulate the population development by age very well.




(a)




(b)

**Figure 4.9 The distributions of population over 5-year age intervals by male (a) and female (b), for 2006**

In order to test the simulation quality of ASMNL, the simulated spatial distribution of population across TAZs in the city is compared with the corresponding observed value. As shown in Figure 4.10, the simulated qualities are nearly the same for male and female.

The $R^2$ value used to measure the correlation between the simulated population and observed population for each TAZ approaches 0.8, which implies high correlation. The regression parameters in the linear equation are not as good as the $R^2$ value. The slope value is higher than its perfect value of 1, and the interception is much lower than its perfect value of 0. Thus, though the simulated results are highly correlated with the actual data, the inaccurate slope and interception values in the linear regression equation indicates that some simulation errors exist in the system.

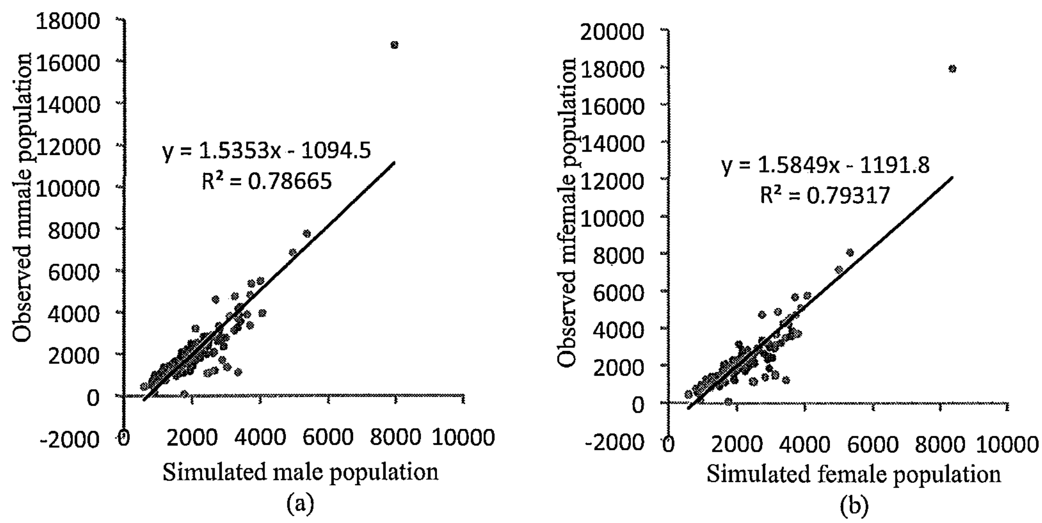



**Figure 4.10 Simulated population vs. observed population by male (a) and female (b), across all TAZs in the city, for 2006**

In Figure 4.10, there is an estimated point extremely far from the trend line, which is the TAZ with ID 16. This TAZ's population is highly underestimated by the system. Without this TAZ, the overall measurement is altered, as shown in Figure 4.11. The r-squared value is reduced a bit, but the parameters in the linear equation show great improvement which increases the quality of the simulation overall. This TAZ is located near the downtown of Burlington. This area has good living conditions, including several parks and a golf center, but the high rent makes this area lose some attractiveness to the residents. The misestimate of the rent in this location would be a reason for this underestimated error.

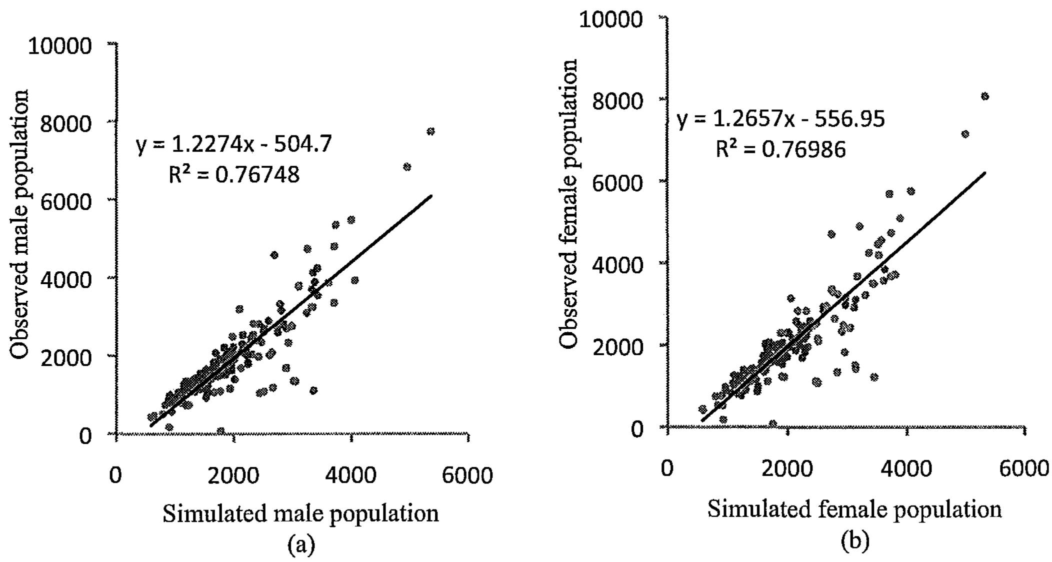



**Figure 4.11 Simulated population vs. observed population by male (a) and female (b), across all TAZs except the TAZ with ID 16, for 2006**

### 4.2.2 Transportation module

The transportation module simulates the travel behavior of all population and adults with the four-stage UTMS model. With trip generation, trip distribution and modal split models, the system generates O-D matrices by trip purposes and transportation modes, for four periods of a typical day.

In order to ascertain whether the distribution of simulated trips can reflect the reality in the city, the observed O-D matrices are used to compare with the simulated results. Sixteen observed auto driver O-D matrices, divided by work and non-work purposes for morning, day, afternoon and night, are utilized in this research. As shown in Table 4.7, the total number of daytime trips predicted by the system is the closest to the actual value, while the simulation for the afternoon is the worst.

**Table 4.7 Total simulated auto-driver trips vs. total observed auto-driver trips for 2001 and 2006**

| | | Morning | Day | Afternoon | Night |
|---|---|---|---|---|---|
| | **2001** | | | | |
| **Work** | Observed | 21,405 | 4,873 | 15,716 | 4,240 |
| | Simulated | 18,884 | 4,630 | 3,148 | 2,188 |
| | *% Difference* | *-11.78* | *-4.99* | *-79.97* | *-48.40* |
| **Non-work** | Observed | 12,745 | 19,263 | 23,297 | 13,732 |
| | Simulated | 13,682 | 19,030 | 32,598 | 11,565 |
| | *% Difference* | *7.35* | *-1.21* | *39.92* | *-15.79* |
| | **2006** | | | | |
| **Work** | Observed | 18,908 | 4,615 | 13,651 | 3,688 |
| | Simulated | 19,639 | 4,461 | 2,904 | 2,309 |
| | *% Difference* | *3.86* | *-3.34* | *-78.73* | *-37.39* |
| **Non-work** | Observed | 13,757 | 19,367 | 22,271 | 11,904 |
| | Simulated | 14,096 | 20,101 | 34,590 | 12,150 |
| | *% Difference* | *2.47* | *3.79* | *55.31* | *2.07* |

The trip generation sub-module generates home-based trips; people who leave their homes produce this kind of trip. If their destinations are working places, they are work trips, otherwise they are non-work trips. Since morning is the peak time for workers leaving their homes to their jobs, work trips should be greater than non-work trips for the morning period and the morning work trips should be greater than the work trips for the other three periods. Both observed and simulated values show this fact.

The simulated day trips are close to the observed data, and the maximum percentage difference is 4.99%. TTS data indicate that the non-work trips are more than work trips for the day period. The simulated results show the same phenomenon. The simulated non-work trips are more than work trips for the day period. The simulated night non-work trips are quite close to the observed data, while the corresponding work trips have a large difference from the observed data. Both observed and simulated data show that non-work trips predominate the total trips for the night period.

For the afternoon period, the simulated results have a huge difference from observed data. The simulated work trips are much fewer than observed data, while the simulated non-work trips are more than the observed data. In order to show the differences between simulated and observed values clearly, Figure 4.12 depicts the comparisons using line graphs.

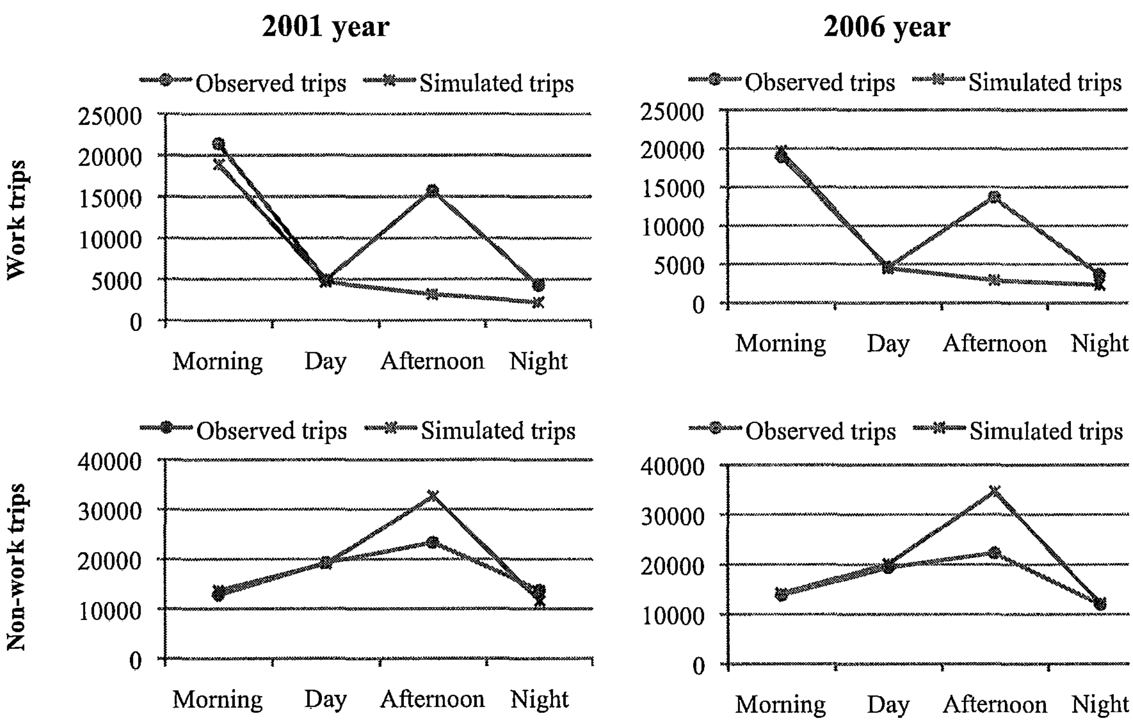



**Figure 4.12 Total simulated auto-driver trips vs. total observed auto-driver trips for 2001 and 2006**

Two factors cause the misestimates of afternoon trips. One is that the observed data includes some errors, and the other is that the trip generation model has poor accuracy. In this research, we used TTS data as the observed values. The observed data provided by TTS is estimated with a 5% sample of the households in the study area. Some mistakes may be produced during the estimation process. In the real world, only the morning is the peak time for home-based work trips. However, the 2001 and 2006 TTS data shows the afternoon is the other peak time for work trips. It indicates that the work trips are more than the non-work trips for the afternoon period. This phenomenon does not follow what actually occurs. We found that this model was simulated using 1996 TTS data as the starting database. In the 1996 TTS database, the work trips in the afternoon are far less than the non-work trips. Based on this starting database, the simulated afternoon non-work trips are more than the work trips. That can explain why the simulated afternoon trips are greatly different from the observed trips. Thus, we cannot determine whether the errors for the afternoon simulations are from TTS database or from the trip generation sub-module.

Though the total number of simulated trips is close to reality for morning, day and night, the distribution of trips over O-D pairs is a key element affecting simulated results of the traffic assignment sub-module. As mentioned above, linear regression, chi-squared test and information-theoretic methodologies are employed to assess the simulations of trips' distribution. Table 4.10 shows the values of the parameters, which are used to compare between the simulated and observed data. Since all the O-D pairs in the simulated night work O-D matrix are 0, we cannot validate the simulations of night work trips in this research.

**Table 4.8 The simulated distribution of auto-driver trips vs. the observed distribution of auto-driver trips over O-D pairs for 2001 and 2006**

| | | Morning | Day | Afternoon | Night |
|---|---|---|---|---|---|
| *2001* | | | | | |
| | | | *Work Trips* | | |
| | $\beta_1$ | 0.71 | 0.18 | 11.76 | N/A[c] |
| | | *(-18.37)*[a] | *(-19.01)* | *(8.45)* | |
| | $\beta_0$ | 0.75 | 0.35 | 1.16 | N/A |
| | | *(17.99)*[a] | *(33.22)* | *(37.41)* | |
| | $R^2$ | 0.13 | 0.00 | 0.01 | N/A |
| | $X_H^2$ | 48085.66[b] | 10342.66 | 31410.27 | N/A |
| | $FT^2$ | 82292.05[b] | 20461.83 | 62719.36 | N/A |
| | **SRMSE** | 2.76 | 3.53 | 3.25 | N/A |
| | | | *Non-work Trips* | | |
| | $\beta_1$ | 1.05 | 0.31 | 0.24 | 0.38 |
| | | *(3.26)* | *(-192.83)* | *(-294.74)* | *(-139.54)* |
| | $\beta_0$ | 0.10 | 1.02 | 1.20 | 0.74 |
| | | *(2.66)* | *(30.34)* | *(27.00)* | *(29.06)* |
| | $R^2$ | 0.26 | 0.37 | 0.38 | 0.35 |
| | $X_H^2$ | 29437.95 | 33208.02 | 53975.15 | 25061.87 |
| | $FT^2$ | 50562.65 | 55270.45 | 89990.83 | 43373.32 |
| | **SRMSE** | 4.52 | 5.26 | 8.07 | 4.52 |
| *2006* | | | | | |
| | | | *Work Trips* | | |
| | $\beta_1$ | 0.76 | 1.48 | 3.18 | N/A |
| | | *(-16.49)* | *(12.16)* | *(4.88)* | |
| | $\beta_0$ | 0.46 | 0.26 | 1.00 | N/A |
| | | *(11.20)* | *(22.82)* | *(32.67)* | |
| | $R^2$ | 0.17 | 0.09 | 0.00 | N/A |
| | $X_H^2$ | 46188.34 | 9173.58 | 27300.71 | N/A |
| | $FT^2$ | 80335.05 | 17610.00 | 54358.58 | N/A |
| | **SRMSE** | 3.07 | 3.87 | 3.63 | N/A |
| | | | *Non-work Trips* | | |
| | $\beta_1$ | 1.27 | 0.31 | 0.23 | 0.33 |
| | | *(17.87)* | *(-222.79)* | *(-334.49)* | *(-178.38)* |
| | $\beta_0$ | -0.04 | 1.02 | 1.12 | 0.63 |
| | | *(-1.07)* | *(30.83)* | *(26.11)* | *(27.49)* |
| | $R^2$ | 0.35 | 0.42 | 0.41 | 0.36 |
| | $X_H^2$ | 30761.73 | 34977.37 | 56751.46 | 23956.67 |
| | $FT^2$ | 52513.67 | 58080.19 | 95808.09 | 41574.23 |
| | **SRMSE** | 4.26 | 5.73 | 9.08 | 5.49 |

a. Two tailed t-test value of the hypothesis that the slope ($\beta_0$) is 1 and the intercept ($\beta_1$) is 0. The critical value is 1.96 for the t distribution with 13454 degrees of freedom at 0.05 level of significance

b. Values of Neyman-Pearson hybrid chi-squared test ($X_H^2$) and Freeman-Tukey chi-squared test ($FT^2$). The critical value is 13726.97 for the $x^2$ distribution with 13456 degrees of freedom at 0.05 level of significance

c. No data is available

Using linear regression methodology, the slope ($\beta_1$), intercept ($\beta_0$) and r-squared ($R^2$) are important parameters used to test if the system can accurately simulate reality. The small $R^2$s show weak correlations within all the observed-simulated (O-D) pairs. The coefficients $\beta_1$ and $\beta_0$ vary over periods per day. We used two tailed t-test of the hypothesis that $\beta_1$ is 1 and $\beta_0$ is 0 assessing if the simulated results are the same as the actual values. The t values show most estimated coefficients have poor fit. Thus, the system is demonstrated to misestimate trips over O-D pairs in the city with the linear regression method. However, when we use this methodology, we do not take account of the effect of spatial autocorrelation in the distribution of O-D trips (Sayer, 1977). The spatial autocorrelation will influence the coefficients in the regression model and change the $R^2$ value (Cressie, 1991; Griffith et al., 1999; Griffith, 2000). In order to reduce the effect of spatial autocorrelation, we sampled random O-D pairs from each O-D matrix, and tested those samples with a linear regression method, since the random pattern exhibits no autocorrelation (Lembo, 2003). The process to sample observations is described in Figure 4.12.

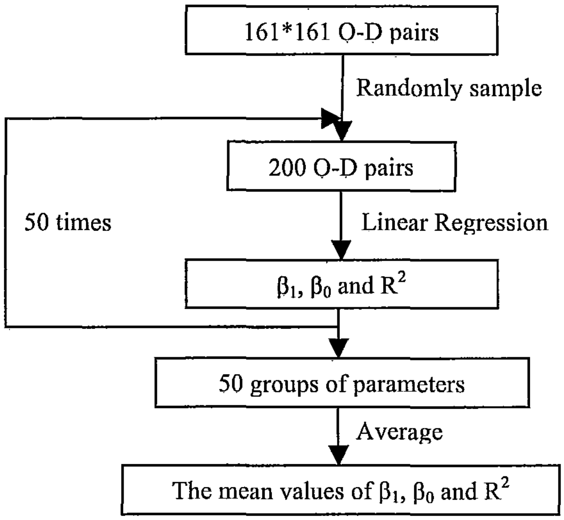



**Figure 4.13 Diagram to estimate the parameters for random samples**

We sample 200 O-D pairs each time, and repeat this sampling work 50 times. Then we can get 50 groups of random samples, which can be used to estimate 50 $\beta_1$s, 50 $\beta_0$s and 50 $R^2$s. The mean value of each kind of parameters is shown in Table 4.9. Those values are close to the values shown in Table 4.10. That means the results generated by sampling method reach the same conclusion as the original linear regression results.

**Table 4.9 The parameters generated by linear regression method using random samples**

| | | Morning | Day | Afternoon | Night |
|---|---|---|---|---|---|
| *2001* | | | | | |
| **Work** | $\beta_1$ | 0.74 | 0.23 | N/A | N/A |
| | $\beta_0$ | 0.71 | 0.35 | N/A | N/A |
| | $R^2$ | 0.14 | 0.02 | N/A | N/A |
| **Non-work** | | | | | |
| | $\beta_1$ | 1.40 | 0.51 | 0.44 | 0.64 |
| | $\beta_0$ | -0.16 | 0.83 | 0.90 | 0.59 |
| | $R^2$ | 0.31 | 0.42 | 0.29 | 0.28 |
| *2006* | | | | | |
| **Work** | $\beta_1$ | 0.71 | 1.10 | N/A | N/A |
| | $\beta_0$ | 0.49 | 0.27 | N/A | N/A |
| | $R^2$ | 0.15 | 0.10 | N/A | N/A |
| **Non-work** | | | | | |
| | $\beta_1$ | 1.26 | 0.44 | 0.38 | 0.54 |
| | $\beta_0$ | -0.09 | 0.86 | 0.85 | 0.52 |
| | $R^2$ | 0.30 | 0.40 | 0.31 | 0.25 |

Besides the linear regression method, $XH^2$ and $FT^2$ are two useful quantitative methods to test the fit of the models using $x^2$ distribution. As shown in Table 4.10, most values of $XH^2$ and $FT^2$ exceed the critical value, except $XH^2$s of day work trips are acceptable. That means the system simulates the work trips for the day period correctly, while the other simulations have low accuracy.

We can compare the predicted qualities between different simulations for different purposes and periods with SRMSE. Relative to the simulations of non-work trips, the system estimates the distribution of work trips more accurately, since the simulations of work trips have lower SRMSE. Among the four periods of a typical day, the trips for the afternoon are simulated with the lowest accuracy, since all SRMSEs for afternoon work trips are higher than 3 and for afternoon non-work trips are higher than 8.

When we used the aforementioned methods, we assumed all the O-D pairs were independent observations. We applied the Entropy statistic to test the level of the dispersion for each O-D matrix. The test of the dispersion is to assess whether the trips are distributed evenly into all O-D pairs. As shown in Table 4.10, the morning simulated trips almost have the same level of dispersion as the corresponding observed values. For the day period, the dispersion situation of simulated non-work trips is close to the observed value. The work trips are not simulated as well as the non-work trips. The observed data shows that the trips are spread across all the O-D pairs in the city, while the simulated data indicates that the trips are only generated for some O-D pairs. In the real world, a zone is always related to its neighborhoods. The model reduces the relationship between a zone and its neighborhoods. As a result, the simulated trips have lower dispersion than the observed trips across the city. The differences for the afternoon trips are produced by the TTS errors.

**Table 4.10 The entropy of simulated O-D matrices vs. the entropy of observed O-D matrices**

| | Morning | Day | Afternoon | Night |
|---|---|---|---|---|
| ***2001*** | | | | |
| | | *Work Trips* | | |
| **Observed** | 160938.01 | 34845.90 | 115942.09 | |
| **Simulated** | 138105.40 | 3333.61 | 16.64 | |
| ***%Difference*** | *-14.19* | *-90.43* | *-99.99* | |
| | | *Non-Work Trips* | | |
| **Observed** | 86270.92 | 146858.55 | 171863.29 | 102731.49 |
| **Simulated** | 88953.11 | 120671.10 | 218158.32 | 68995.27 |
| ***%Difference*** | *3.11* | *-17.83* | *26.94* | *-32.84* |
| ***2006*** | | | | |
| | | *Work Trips* | | |
| **Observed** | 138463.98 | 32382.74 | 97888.39 | |
| **Simulated** | 144592.15 | 4902.12 | 249.38 | |
| ***%Difference*** | *4.43* | *-84.86* | *-99.75* | |
| | | *Non-Work Trips* | | |
| **Observed** | 93012.05 | 146930.93 | 162372.18 | 87443.94 |
| **Simulated** | 92960.59 | 126768.36 | 232072.21 | 72723.47 |
| ***%Difference*** | *-0.06* | *-13.72* | *42.93* | *-16.83* |

The final stage of transportation module is traffic assignment. We use the observed traffic flows in 91 selected links compare with the simulated results for morning, day, afternoon and night. As shown in Table 4.11, the system highly underestimates the overall traffic flows in those links.

**Table 4.11 Simulated traffic flows vs. observed traffic flows in 91 links, 2006**

| | Morning | Day | Afternoon | Evening |
|---|---|---|---|---|
| **Observed Flows** | 56,019.23 | 59,730.36 | 72,226.07 | 35,257.2 |
| **Simulated Flows** | 34,783.18 | 7,709.33 | 18,176.72 | 4,968.43 |
| *% Difference* | *-37.91* | *-87.09* | *-74.83* | *-85.91* |

The spatial distribution of traffic flows across the selected links in the city is compared with the observed value in Figure 4.11. The $R^2$ values in those four charts are all between 0.42 and 0.46, which show low correlation between the simulated and observed flows. The parameters in the linear regression equation show the inaccurate simulation of traffic flows for different periods of a typical day.

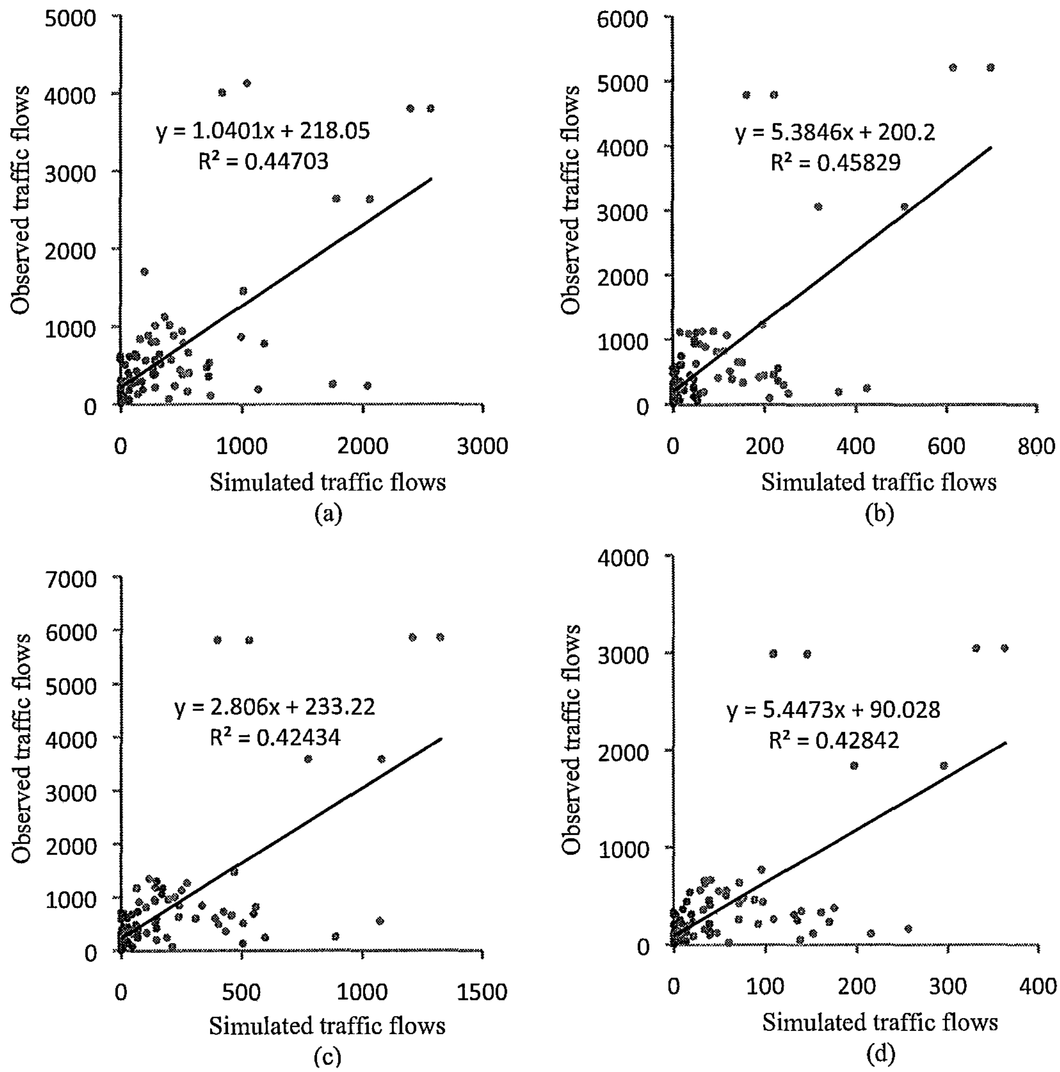



**Figure 4.14 Simulated traffic flows vs. observed traffic flows across all TAZs, for morning (a), day (b), afternoon (c) and evening (d), 2006**

# Chapter 5. Conclusion

IMULATE and IMPACT were both developed to simulate the urban dynamic in the Hamilton CMA. Although they are both constructed based on a four-stage UTMS, they use different sub-models to implement different simulations. Both systems use a three-module framework. The first module is the land use/demographic module, which simulates the distribution of population. The results generated by this module are used by the trip generation sub-model, the first stage of UTMS. In this module, IMULATE can simulate the distribution of households and employments over time, while IMPACT only simulates the distribution of population. Compared to IMPACT, IMULATE considers more factors in the land use/demographic simulations. IMULATE can simulate the distribution of employment, but IMPACT cannot.

The second sub-module is the transportation module, which estimates the inter-zonal trips over TAZs and traffic flows over the road network in the city. The two models use different algorithms to simulate trip generation and trip distribution. Compared to IMULATE, the trip generation model in IMPACT has higher accuracy. IMULATE generates more trips than the actual number. IMULATE's transportation module has lower accuracy due to these overestimates. Both models use MNL and SUE algorithms for modal split and traffic assignment. For these two stages, the only difference between IMULATE and IMPACT is that different variables are considered in the modal split models.

The last module is the environmental module, which estimates the emissions for each road link. Both IMULATE and IMPACT use MOBILE 6C to predict the emissions.

Since this module has no interaction with the other two modules, we did not test it in this research.

In our research, we analyzed the framework of IMULATE and IMPACT. We reported many issues that would cause the misestimates of models by testing the sub-models one by one. Some of these issues can be resolved, while some cannot. This is our first contribution. Another contribution is the data collection work. We collected data observed from the real world and integrated them with the models. These data are not only useful in this work, but they can also be used for other research in the future. Our last contribution is the combination of various methods in validation. Each method can only show validity in one aspect. We used different methods to test the models, and the tested results illustrated the simulated accuracy in several aspects.

Many IUMs are composed of existing sub-models. Our research pointed out the problems that can occur in using some of these sub-models. This provides a good reference point on the use of those sub-models for future development. More accurate results will help the users make the correct decision. This research also provides a systematic procedure to validate IUMs. It will help us to validate other models in the future. Furthermore, based on our findings, we hope to improve IMULATE and IMPACT in the future.

## Bibliography

Abraham, J. (2000). *Parameter Estimation in Urban Models: Theory and Application to a Land Use Transport Interaction Model of the Sacramento, California region.* University of Calgary.

Abraham, J., & Hunt, J. (1999a). Firm location in the MEPLAN model of Sacramento. *Transportation Research Record, 1685*, 187-198.

Abraham, J., & Hunt, J. (1999b). Policy Analysis Using the Sacramento MEPLAN Land Use-Transportation Interaction Model. *Transportation Research Record, 1685*, 199-208.

Alonso, W. (1964). *Location and Land Use.* Cambridge: Harvard University Press.

Anas, A. (1986). From physical to economic models: the Lowry frame-work revisited. In B. Hutchinson, & M. Batty, *Advances in Urban Systems Modelling.* Amsterdam: North Holland.

Anderson, W., Kanaroglou, P., & Miller, E. (1996). *Integrated Land Use and Transportation Model for Energy and Environmental Analysis: A Report on Design and Implementation.* Center of Spatial Analysis. Hamilton: School of Geography and Earth Sciences, McMaster University.

Baxer, M., & Ewing, G. (1981). Models of recreational trip distribution. *Regional Studies , 15*, 327-344.

Ben-Akiva, M. (1974). Structure of passenger travel demand models. *Transportation Research Record, 526.*

Black, W. (1973). An analysis of gravity model distance exponents. *Transportation , 2*, 299-312.

Borning, A., Waddell, P., & Forster, R. (2006). UrbanSim: using simulation to inform public deliberation and decision-making. In H. Chen, *Digital Government: Advanced Research and Case Studies.* Druck: Springer-Verlag.

Brokke, G. (1958). Evaluating trip forecasting methods with an electronic computer. *Highway Research Board Bullentin , 203.*

Celik, M. (2004). Forecasting interregional commodity flows using artificial neural networks: an evaluation. *Transportation Planning and Technology , 27*, 449-467.

Cressie, N. (1991). *Statistics for Spatial Data.* New York: Wiley.

Daganzo, C., & Sheffi, Y. (1977). On stochastic models of traffic assignment. *Transportation Science*, *11*, 253-274.

Echenique, M., & Owers, J. (1994). Research into practice: the work of the Martin Center in urban and regional modeling. *Environment and Planning B: Planning and Design*, *21*, 513-650.

Ettema, D., & Timmermans, H. (2006). Multi-agent modelling of urban systems: a progress report of PUMA System. *Stadt Region Land*, *81*, 165-171.

Ettema, D., Borgers, A., & Timmermans, H. (1996). SMASH (Simulation Model of Activity Scheduling Heuristics): Some simulations. *Transportation Research Record*, *1551*, 88-94.

Federal Highway Administrator. (1975). *Trip Generation Analysis.* Washington, D.C: Department of Transportation.

Ferguson, R. (1999). Snowmelt runoff models. *Progress in Physical Geography*, *23*, 205-227.

Fotheringham, A. (1983). A new set of spatial interaction models: the theory of competing destination. *Environment and Planning A*, *15*, 15-36.

Fotheringham, A., & Knudsen, D. (1987). *Goodness-of-fit Statistics.* Norwish: Geo Books.

Fotheringham, A., & O'Kelly, M. (1989). *Spatial Interaction Models: Formulations and Applications.* Netherlands: Luwer Academic Publishers.

Franklin, J., Waddell, P., & Britting, J. (2002). Sensitivity Analysis Approach for an Integrated Land Development an Travel Demand Modeling System. *Proceedings of the 2003 Agile Programming Conference*, (pp. 21-24). Salt Lake City.

Fratar, T., Voorhees, A., & Raff, M. (1954). Forecasting distribution of inter-zonal vehicular trips by successive approximations. *Highway Research Board Proceedings*, *33*, 376-384.

Golledge, R., Kwan, M., & Garling, T. (1994). Computational Process Modeling of Household Travel Decisions Using a Geographical Information System. *Papers in Regional Science*, *73*, 99-117.

Griffith, D. (2000). A linear regression solution to the spatial autocorrelation problem. *Journal of Geographical Systems*, *2*, 141-156.

Griffith, D., & Layne, L. (1999). *A Casebook for Spatial Statistical Data Analysis.* New York: Oxford.

Ham, H., Kim, T., & Boyce, D. (2005). Assessment of economic impacts from unexpected events with an interregional commodity flow and multimodal transportation. *Transportation Research A , 39*, 849-860.

Hathaway, P. (1975). Trip distribution and disaggregation. *Environment and Planning A , 7*, 71-97.

Hu, P., & Pooler, J. (2002). An empirical test of the competing destinations model. *Journal of Geographical Systems , 4*, 301-323.

Hunt, J., & Simmonds, D. (1993). Theory and application of an integrated land use and transport modeling framework. *Environment and Planning B: Planning and Design , 20*, 221-224.

Hutchinson, B. (1974). *Principles of Urban Transport Systems Planning.* New York: McGraw-Hill.

Jones, P., Koppelman, F., & Orfeuil, j. (1990). Activity analysis: state of the art and future directions. In P. Jones, *Developments in Dynamic and Activity Based Approaches to Travel Analysis* (pp. 34-55). Avebury: Oxford Studies in Transport.

Kanaroglou, P., & Scott, D. (2002). Integrated urban transportation and land-use models for policy analysis. In M. Dijst, *Governing Cities on the Move: Functional and Management Perspectives on Transformations of Urban Infrastructure in European Agglomerations* (pp. 43-75). Ashgate.

Kang, H., Scott, D., Kanaroglou, P., & Maoh, H. (2009). An exploration of issues related to the study of generated traffic and other impacts arising from highway improvements . *Environment and Planning B: Planning and Design , 36*, 67-85.

Kawakami, S., & Isobe, T. (1990). Development of a one-day travel-activity scheduling model for workers. In P. Jones, *Developments in Dynamic and Activity-Based Approaches to Travel Analysis* (pp. 184-205). Avebury: Oxford Studies in Transportation.

Kitamura, R. (1988). An evaluation of activity based travel analysis. *Transportation , 15*, 9-34.

Kitamura, R., Pas, E., Lula, C., Lawton, T., & Benson, P. (1996). The sequenced activity mobility simulator (SAMS): an integrated approach to modelling transportation, land use and air quality. *Transportation* , *23*, 267-291.

Knudsen, D., & Fortheringham, A. (1986). Matrix comparison, goodness-of-fit, and spatial interaction modeling. *International Regional Science Review* , *10*, 127-147.

Knudsen, D., & Fotheringham, A. (1986). Matrix comparison, goodness-of-fit and spatial interaction modeling. *International Regional Science Review* , *10*, 127-147.

la de Barra, T. (1989). *Integrated Land Use and Transport Modelling: Decision Chains and Hierarchies.* Cambridge: Cambridge University Press.

la de Barra, T., Perez, B., & Vera, N. (1984). TRANUS-J: putting large models into small computers. *Environment and Planning B: Planning and Design* , *11*, 87-101.

Landis, J. (1995). Imagining Land Use Futures: Applying the California Futures Model. *Journal of the American Planning Association* , *61*, 438-457.

Landis, J., & Zhang, M. (1998). The second generation of the California futures model. *Environment and Planning B* , *25*, 795-824.

LeBlanc, L., Morlok, E., & Pierskalla, W. (1976). An efficient approach to solving the road network equilibrium traffic assignment problem. *Transportation Research* , *9*, 34-55.

Lee, Z., Carder, K., Steward, R., Peacock, T., Davis, C., & Patch, J. (1998). An empirical algorithm for light absorption by ocean water based on color. *Journal of Geophysical Research* , *103*, 27967-27978.

Legates, D., & McCabe, G. (1999). Evaluating the use of ``goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research* , 233-242.

Lembo, J. (2003). *Spatial Autocorrelation.* Salisbury University, Salisbury.

Leontief, W. (1941). *The Structure of the American Economy 1919-1939.* New York: Oxford University Press.

Lewis, P. (1975). An empirical test of alternative theories of trade. *Annals of Regional Science* , *9*, 102-111.

Lowry, I. (1964). *A Model of Metropolis.* Santa Monica: The Rand Corporation.

Macgill, S. (1977). The Lowry model as an input-output model and its extension to incorporate full intersectoral relations. *Regional Studies* , *11*, 337-354.

Macgill, S., & Wilson, A. (1979). Equivalences and Similarities between some Alternative Urban and Regional Models. *Systemi Urbani* , *1*, 9-40.

Maoh, H., Kanaroglou, P., Scott, D., Paez, A., & Newbold, B. (2008). IMPACT: An integrated GIS-based model for simulating the consequences of demographic changes and population ageing on transportation. *Computers, Environment and Urban Systems* , *33*, 200-210.

Martinez, F. (1996). MUSSA: land use model for Santiage city. *Transportation Research Record* , *1552*, 126-134.

Meyer, M., & Miller, E. (1984). Demand analysis. In M. Meyer, & E. Miller, *Urban Transportation Planning* (pp. 225-289).

Miller, E. (1996). Microsimulation and activity-based forecasting. In T. T. Institute (Ed.), *Activity-based Travel Forecasting Conference, June 2-5, 1996: Summary, Recommendations, and Compendium of Papers* (pp. 151-172). Washington, D.C.: Travel Model Improvement Program, U.S. Department of Transportation and U.S. Environmental Protection Agency.

Miller, E., & Roorda, M. (2003). A prototype model of 24-h household activity scheduling for the Toronto Area. *Transportation Research Record* , *1831*, 114-121.

Miller, E., Noehammer, P., & Ross, D. (1987). A micro-simulation model of residential mobility. *Proceedings of the International Symposium on Transport, Communication and Urban Form. 2*, pp. 217-234. Victoria: Monash University.

Moeckel, R., Schurmann, C., & Wegener, M. (2002). Microsimulation of Urban Land Use. *Paper presented at the 42nd European Congress of the Regional Science Association* .

Mohammadian, A., & Miller, E. (2003). Empirical investigation of household vehicle type choice decisions. *Transportation Research Record* , *1854*, 99-106.

Mohammed, A., Shalaby, A., & Miller, E. (2008). Empirical analysis of transit network evolution : Case study of Mississauga, Ontario, Canada, bus network. *Transportation Research Record* , *1971*, 51-58.

Mozolin, M., Thill, J., & Usery, E. (2000). Trip distribution forecasting with multilayer perceptron neural networks: A critical evaluation. *Transportation Research Part B , 34,* 53-73.

Nguyen, S. (1974). An algorithm for the traffic assignemnt problem. *Transportation Science , 8,* 203-216.

Paktunc, A. (1998). MODAN: an interactive computer program for estimating mineral quantities based on buik composition. *Computers and Geosciences , 24,* 425-431.

Pas, E. (1990). Is travel demand analysis and modelling in the doldrums? In P. Jones, *Developments in Dynamic and Activity Based Approaches to Travel Analysis* (pp. 3-27). Avebury: Oxford Studies in Transport.

Putman, S. (1983). *Integrated Urban Models: Policy Analysis of Transportation and Land Use.* London: Pion.

Rodier, C. (2005). *Verifying the Accuracy of Land Use Models Used in Transportation and Air Quality: A Case Study in the Sacramento, California Region.* Davis: Institute of Transportation Studies.

Roorda, M., Miller, E., & Habib, K. (2008). Vlidation of TASHA: A24-h activity scheduling microsimulation model. *Transportation Research A , 42,* 360-375.

Salvini, P., & Miller, E. (2005). ITLUTE: an operational prototype of a comprehensive microsimulation model of urban systems. *Networks and Spatial Economics , 5,* 217-234.

Sayer, R. (1977). Gravity models and spatial autocorrelation, or atrophy in urban and regional modelling. *Area , 9,* 183-189.

Smith, D., & Hutchinson, B. (1981). Goodness of fit statistics for trip distribution models. *Transportation Research , 15A,* 295-303.

Tait, J., & Revenaugh, J. (1998). Source-transport inversion: an application of geophysical inverse theory to sediment transort in Monterey Bay, California. *Journal of Geophysical Research , 103,* 1275-1283.

Thorsen, I., & Gitlesen, J. (1998). Empirical evaluation of alternative model specifications to predic commuting flows. *Journal of Regional Science , 38,* 273-292.

Voas, D., & Williamson, P. (2001). Evaluating goodness-of-fit measures for synthetic microdata. *Geographical and Environmental Modelling , 5,* 177-200.

Waddell, P. (2000). A behavioral simulation model for metropolitan policy analysis and planning: residential location and housing market components of UrbanSim. *Environment and Planning B: Planning and Design , 27*, 247-263.

Waddell, P. (1998). An urban simulation model for integrated policy analysis and planning: residential location and housing market components of UrbanSim. *8th World Conference on Transport Research.* Antwerp, Belgium.

Waddell, P., & Ulfarsson, G. (2004). Introduction to urban simulation: design and development of operational models. In K. Haynes, P. Stopher, K. Button, & D. Hensher, *Handbook: Transport Geograhy and Spatial Systems* (Vol. 5). Seattle: Pergamon Press.

Waddell, P., Franklin, J., & Britting, J. (2003). *UrbanSim: development, application and integration with the Wasatch Front Regional Travel Model.* Technical Report, Center for Urban Simulation and Policy Analysis, University of Washington.

Waddle, P. (2002). UrbanSim: modeling urban development for land use, transportation, and environmental planning. *Journal of the American Planning Association , 68*, 297-314.

Wagner, P., & Wegener, M. (2007). Urban land use, transport and environmental models: experiences with an integrated microscopic approach. *disP - The Planning Review , 170*, 45-56.

Wegener, M. (1982). Modeling urban decline: a multilevel economic-demographic model for the Dortmund region. *International Regional Science Review , 7*, 217-241.

Wegener, M. *The IRPUD Model: Overview.* Institut für Raumplanung, Universität Dortmund, Dortmund.

Wilmot, C., Modali, N., & Chen, B. (2006). *Modeling Hurricane Evacuation Traffic: Testing the Gravity and Intervening Opportunity Models as Models of Destination Choice in Hurricane Evacuation.* Louisiana Transportation Research Center, Department of Civil and Environmental Engineering. Baton Rouge: Louisiana State University.

Wilson, A. (1970). *Entropy in Urban and Regional Modelling.* London: Pion.

Wilson, A., Coelho, C., Macgill, S., & Williams, H. (1981). *Optimisation in Locational and Transport Analysis.* Chichester: John Wiley and Sons.