2010

INTEGRATION OF THE REGRESSION-BASED LAND USE MODEL AND THE COMBINED TRIP DISTRIBUTION-ASSIGNMENT TRANSPORTATION MODEL

Meiwu An
University of Kentucky, anmeiwu@engr.uky.edu

ABSTRACT OF DISSERTATION

Meiwu An

The Graduate School
University of Kentucky

2009

INTEGRATION OF THE REGRESSION-BASED LAND USE MODEL AND THE COMBINED TRIP DISTRIBUTION-ASSIGNMENT TRANSPORTATION MODEL

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Engineering
at the University of Kentucky

By
Meiwu An
Lexington, Kentucky

Co-Director: Dr. Mei Chen, Professor of Civil Engineering
and Dr. Nick Stamatiadis, Professor of Civil Engineering
Lexington, Kentucky
2009

ABSTRACT OF DISSERTATION

INTEGRATION OF THE REGRESSION-BASED LAND USE MODEL AND THE COMBINED TRIP DISTRIBUTION-ASSIGNMENT TRANSPORTATION MODEL

Regional growth caused the emergence of traffic congestion and pollution in the past few decades, which have started to affect small urban areas. These problems are not only related to transportation system design but also to land use planning. There has been growing recognition that the relationship between land use and transportation needs to be understood and analyzed in a consistent and systematic way. Integrated urban models have recently been introduced and implemented in several metropolitan areas to systematically examine the relationship between land use and transportation. The general consensus in the field of integrated urban models is that each model has its own limitations and assumptions because they are each designed for different application purposes. This dissertation proposes a new type of methodology to integrate the regression-based land use model and the combined trip distribution-assignment transportation model that can be applied to both metropolitan areas and small urban areas.

The proposed integrated land use and transportation model framework has three components: the regression-based land use model, the combined trip distribution-assignment transportation model, and the interaction between these two models. The combined trip distribution-assignment model framework provides the platform to simultaneously integrate the transportation model with the land use model. The land use model is developed using an easy-to-implement method in terms of correlation and regression analysis.

The interaction between the land use model and the transportation model is examined by two model frameworks: feedback model framework and simultaneous model framework. The feedback model framework solves the land use model and the transportation model iteratively. The simultaneous model framework brings the land use model and the transportation models into one optimization program after introducing the used path set. Both the feedback model and the simultaneous model can be solved to estimate link flow, origin-destination (OD) trips, and household distribution with the results satisfying network equilibrium conditions.

The proposed integrated model framework has an "affordable and easy-to-implement" land use model; it can be performed in small urban areas with limited resources. The model applications show that using the proposed integrated model framework can help decision-makers and planners in preparing for the future of their communities.

KEYWORDS: Integrated Land Use and Transportation Model, Travel Demand Model, Combined Trip Distribution and Traffic Assignment, User Equilibrium, Entropy

Meiwu An

12-21-2009

INTEGRATION OF THE REGRESSION-BASED LAND USE MODEL AND THE COMBINED TRIP DISTRIBUTION-ASSIGNMENT TRANSPORTATION MODEL

By

Meiwu An

Dr. Mei Chen

Co-Director of Dissertation

Dr. Nick Stamatiadis

Co-Director of Dissertation

Dr. Kamyar C. Mahboub

Director of Graduate Studies

12-21-2009

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name	Date

DISSERTATION

Meiwu An

The Graduate School
University of Kentucky
2009

INTEGRATION OF THE REGRESSION-BASED LAND USE MODEL AND THE COMBINED TRIP DISTRIBUTION-ASSIGNMENT TRANSPORTATION MODEL

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Engineering
at the University of Kentucky

By
Meiwu An
Lexington, Kentucky

Co-Director: Dr. Mei Chen, Professor of Civil Engineering
and Dr. Nick Stamatiadis, Professor of Civil Engineering
Lexington, Kentucky
2009

For my well-beloved family

ACKNOWLEDGMENTS

I would like to acknowledge and thank my advisor, Dr.Mei Chen who offered this great opportunity for me to complete this dissertation. This research would never have been accomplished without your continual guidance, support, inspiration in time of doubt and tribulation. Your insightful and significant challenges have been playing an extremely important role in improving this study throughout many starts and stops along the way.

I would like to express my sincere gratitude to Dr. Nick Stamatiadis for serving co-director of my dissertation. Your thoughtful support and advice have greatly improved the quality of this dissertation. Your genial friendship and encouragement are always helpful in sustaining me through the dissertation procedure.

I truly appreciate Dr.Carl Lee, who is willing to give up precious time to help improving math formulation for this dissertation. Your valuable suggestion and discussion is indispensable in developing this dissertation. Thanks for providing me with math knowledge as well as inspiration and encouragement.

I greatly appreciate Dr. Jerry G. Rose for serving as my committee member. Both your insightful advice, comments and your encouragement, support are essential to making solid progress on this dissertation.

My special thanks go to Dr.Ted Grossardt and John Ripy (Kentucky Transportation Center), David Hamilton and Kong Ee (Kentucky Transportation Cabinet), Nicole Lefever (Woodford County Planning Commission) for your generous sharing of your data resource, technical assistance, encouragement and expertise.

I must earnestly be grateful to my parents and my friends whose love and support made this work possible. Although my parents in China are thousands of miles away from me, their endless hours of child care, unrelenting optimism made it possible for me to pursue this study at ease.

Above all I want to thank my wife Fengxia Wang and my son Fredy An who are always bringing me the love, laughter, encouragement and inspiration which stimulates me to complete my goals and accomplish this work in greater perspective.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 BACKGROUND

1.1.1 Problem Statement

The emergence of explicit transportation problems such as congestion and pollution has caused increasing governmental and public concerns about urban development. These problems are not only caused by transportation system design, but also relate to land use planning (urban sprawl) to some degree. There has been growing recognition that the relationship between land use and transportation needs to be understood in a consistent and systematic way (Miller, 2004).

Land use has been interacting with transportation systems during the course of urban development. For example, a new urban road will encourage the development of adjacent land. As the land is developed, travel demand will increase, leading to the congestion on this new road. As the traffic increases, the road need to be improved or a new road will have to be built. The new highway will then encourage additional land development, and the cycle continues. Considering environmental and financial constraints, it is implausible to “build our way out of congestion” by continuing new highway construction. This cycle has resulted in reshaping policy for metropolitan planning (Downs, 1992). The interaction between land use and transportation needs to be better understood as we strive to resolve urban problems and work toward sustainable development.

Recognizing the interaction between transportation systems and land use development, legislators have attempted to coordinate transportation and land use plans. Recent legislation includes the Inter-modal Surface Transportation Efficiency Act of 1991 (ISTEA), the Transportation Equity Act for the 21st Century (TEA 21), and Safe, Accountable, Flexible, Efficient Transportation Equity Act: a Legacy for Users

(SAFETEA-LU). ISTEA mandates that Metropolitan Planning Organizations (MPOs) integrate land use and transportation planning as stated in section 134:

"In developing transportation planning plans and programs pursuant to this section, each metropolitan planning organization shall, at a minimum consider the following

The likely effect of transportation policy decisions on land use and development and the consistency of transportation plans and programs with the provision of all applicable short- and long-term land use and development plans"

Although these laws explicitly require the coordination between land use plans and transportation plans, none of them specify the methods to be used to achieve this integration. In general, land use plans are developed by local government; regional land use plans are rare, and only a few states have established statewide frameworks for land use planning (Parsons Brinckerhoff, 1998). However, transportation plans are developed by state Departments of Transportation (DOT), MPOs and transit agencies with significant funding support and regulation from the federal government. Since different organizations make public decisions on land use planning and transportation planning, coordination is difficult. Even in the same jurisdiction such as city governments, land use and transportation are often handled by different departments, with engineers responsible for transportation decisions and planners responsible for land use planning. As a result, land use planners, transportation engineers, and decision makers could have different or even conflicting goals and objectives (Parsons Brinckerhoff, 1998).

Not only are land use plans and transportation plans developed by separate organizations, but these plans are implemented by different sectors. In the United States, transportation infrastructure investment and construction are made by multiple levels of governments, including federal, state and local. Local government is responsible for maintaining local roads and new local roads are usually constructed by land developers through contribution or impact fees. Federal and state governments are mainly in charge of maintaining, rehabilitating and constructing federal and state highway networks in order to accomplish state or regional goals. In contrast, most land use development and investment is made by individuals and firms within the context of local land-use plans.

Governments have limited control on land-use development, since land owners are permitted to develop their property to its highest use. Governments only intervene when development causes damage to protected species or is incompatible with land use plans (Parsons Brinckerhoff, 1998).

1.1.2 Why Integrated Land Use and Transportation Model

In response to both federal legislation and increasing concerns regarding transportation system problems, the integrated land use and transportation (urban) model has been proposed as a tool to strengthen the coordination between land use plans and transportation plans at different level of governments. An integrated model can also be used to investigate the interrelationship between land use and transportation in a systematic way. Decision makers are able to use model outputs to assess the impact of land use development on transportation systems, and the impact of transportation policies on land use development. Public policies can be developed that include transportation demand management, congestion pricing, parking pricing and management, and etc. Thus, the integrated model can play an important role in shaping the long-term future of a community.

1.2 INTEGRATED LAND USE AND TRANSPORTATION MODEL

1.2.1 Connection between Land Use and Transportation

Land use can be defined as the way in which land is used; it not only includes the buildings on the land such as houses, factories, offices, stores, etc., but also includes the activities occurring in these buildings such as working, shopping, education, etc. (Miller, 2004). The participation of out-of-house activities such as working and shopping gives rise to the need for travel on transportation networks. For example, movement of people from home to workplace and goods from one factory to another for production cannot be achieved without the support of transportation systems.

It is increasingly recognized that there is a significant interactive relationship between land use and transportation. Transportation demand is engendered from land use development; but also transportation systems have an important impact on land use development by providing accessibility. Land-use development configuration is highly related to transportation system design, and vice versa. For example, if a transportation system is built differently, people will use it differently, and they will spatially organize themselves differently. Conversely, if a city is built differently, transportation systems and needs will be different (Miller et al, 1999).

Accessibility is the nexus between land use and transportation systems. The interaction between land use and transportation can be measured by accessibility, which reflects both attractiveness and ease of reaching destinations (Handy, 1993). The accessibility can be defined as a function of land use development (urban activity) distribution and transportation system configuration. The pattern of land-use development has a significant impact on accessibility since it determines the distribution of attractiveness in terms of urban activities. The structure and capacity of the transportation system affects accessibility as significantly as land-use pattern since it determines the ease of reaching urban activities. For example, decreasing transportation cost in terms of money or time between any two places will result in increasing interaction between them. The relationship between land use and transportation can be illustrated in Figure 1.1 (Parsons Brinckerhoff, 1998).

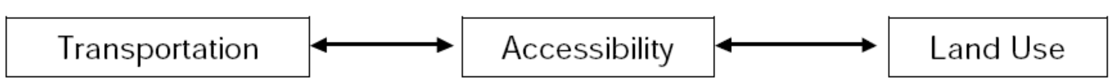

Figure 1.1 Accessibility Links Land Use and Transportation

1.2.2 Integrated Land Use and Transportation Model

A model is a pattern or representation designed to simulate the structure or working of an object. In the context of an urban activity system, a model consists of mathematic equations that can be used to simulate human activities such as demographic distribution and travel patterns (Miller et al, 1999). As a consequence, land use models are designed to scientifically simulate demographic and economic distribution in urban or regional areas. In this dissertation, land use models focus on the estimation of the household, employment distribution. The output of land use models will provide key inputs for travel demand models. A transportation model is primarily devised to forecast the spatial movement of people and goods and convert these movements into traffic volume over a transportation system.

Historically, land use models and transportation models are developed separately and applied in urban and transportation planning. The innate and indispensable connection between land use and transportation systems brings about the demand to comprehensively unite transportation models and land use models. An integrated model is designed to capture the interrelationship between land use and transportation as much as possible (Miller, 2004). Therefore, creating an integrated model includes not only developing land use and transportation models but also investigating the interactions between them.

1.3 RESEARCH OBJECTIVES AND CONTRIBUTION

There have been substantial research developments in the field of integrated urban modeling. Several operational integrated urban models have been developed and applied to some metropolitan areas, which will be discussed in Chapter 2. However, each model has its own limitations because of different application purposes. For example, existing integrated urban models cannot be applied to small urban areas since they target large metropolitan areas.

This dissertation aims to develop a new type of integrated land use and transportation model framework that can be used on both small urban areas and large metropolitan areas. This new integrated model framework is composed of a regression-based land use model, a combined trip distribution-assignment transportation model, and the interaction between these two models. The new model integrates the regression-based land use model and the combined trip distribution-assignment transportation model. This model framework is capable of estimating urban activity distribution and traffic flow distribution in a consistent and comprehensive way.

Compared to existing integrated models, this model framework possesses several clear-cut advantages. This framework presents the first instance of integration of the regression-based land use model and the combined trip distribution-assignment transportation model. It is designed to be compatible with modern transportation modeling and to be affordable to implement not only by metropolitan areas with adequate resources but also by small urban areas with limited resources.

The combined trip distribution-assignment model will serve as the transportation model in this framework, which has rarely been examined in existing integrated model frameworks. Existing integrated models mostly adopt a traditional four-step travel demand model as the transportation model. This dissertation will contribute to exploring the combined trip distribution-assignment transportation model within the context of an integrated model framework, including formulation, calibration, and application.

The land use model is developed using an easy-to-implement method in terms of correlation and regression analysis. This easy-to-implement method has not been seen in the current literature of land use models. This land use model can be reasonably achieved with a limited budget and with limited professional crew in small urban areas. In contrast, existing land use models require extensive data and a large budget, which typical small urban areas cannot afford.

The interaction between land use and transportation models will be investigated by two methods in this framework. One is to build a feedback loop between these two models through intermediate variables of accessibility; the other is to formulate the land use model and the transportation model as a united optimization program after introducing the used path set. Under the method, the land use model and the transportation model can be simultaneously solved. In the second method, household and employment distribution, as well as traffic flow over the network are regarded as endogenously-determined variables and can be simultaneously estimated. This method utilizes the used path set to simultaneously formulate/solve the land use model and the transportation model, which has not been seen in the existing integrated models.

1.4 RESEARCH APPROACHES

The proposed new type of integrated model framework consists of three components: a transportation model, a land use model, and the interaction between these two models. After gathering necessary transportation data such as socio-economic data and roadway characteristics, origin-destination (OD) trip tables, and land use data such as the area of each type of land use from multiple data sources, it starts to develop the proposed model framework.

The combined trip distribution-assignment model is formulated based on entropy concepts and calibrated by base year data. The gravity model form is built in this transportation model for trip distribution. The output of trip assignment is able to satisfy the conditions of user equilibrium (or network equilibrium), which will be discussed in detail in Chapter 4. This transportation model can help identify the deficiency in transportation networks using the measure of operational level of service and system-wide measures of effectiveness over time throughout the study area. It can also generate transportation performance measures for each analysis zone in terms of accessibility, which are critical inputs to the land use model.

The land use model is then established after developing two categories of factors associated with land use structure and transportation measures. The factors with statistically significant influence on the urban activity distribution are identified through correlation analysis. The appropriate model forms are developed by combining some of these factors, which provide better explanation and estimation for urban activity distribution. These models can help in understanding not only the impact of certain transportation measures on urban activity distribution, but also the impacts of neighborhood design such as zone characteristics on urban activity distribution. More importantly, the output of the land use models in terms of household and employment distribution will serve as a major input to the transportation model.

The interaction between these two models is then investigated using two solution procedures. These two solution procedures are able to produce the consistency between the land use model output and the transportation model output. The consistency is such a condition that the transportation model outputs as input to the land use model will be able to produce the same land use model outputs as those initially put into the transportation model, and vice versa. These two solution procedures are formulated by two types of model frameworks respectively. The first is a feedback loop configuration between the land use model and the transportation model through the intermediate variables of accessibility. The iteration between the land use and the transportation model continues until pre-defined convergence criteria are reached (discussed in Chapter 6). The second is the simultaneous model framework, which formulates the land use model and the transportation model together as an optimization program after introducing the used path set. Therefore, the land use model and the transportation model can be solved at the same time instead of the iterations between these two models in the feedback loop configuration.

1.5 ORGANIZATION OF THIS DISSERTATION

This dissertation is divided into seven chapters. Chapter 1 discusses the background, problem statement, research objective and methodology, and dissertation organization. The literature review of integrated land use and transportation models based on different theory is provided in Chapter 2. Chapter 3 introduces the procedure to prepare the data for developing the new type of integrated land use and transportation model. Chapter 4 discusses the formulation and calibration of the combined trip distribution-assignment transportation model. Chapter 5 describes the correlation and regression analysis for developing the regression-based land use model. Chapter 6 discusses the interaction between the land use and transportation models using a feedback loop model framework and a simultaneous model framework. Conclusions and recommendations for future research are presented in Chapter 7.

CHAPTER 2
LITERATURE REVIEW

2.1 INTRODUCTION

An urban system can be divided into several subsystems such as transportation network, travel, goods transport, land use, employment, population, housing, workplace, and environment (Wegener, 1994). An urban model simulates the structure or function of one subsystem in an urban system. An integrated urban model simulates two or more subsystems.

Before 1960, integrated urban models did not truly exist. After 1960, demand for preparing highway impact statements and government concern about urban problems such as environment and energy greatly stimulated the development of integrated urban models. A few integrated urban models have been developed, which are designed to model land use and transportation subsystems and other subsystems as well. Several comprehensive studies about these integrated urban models have been conducted by researchers for different purposes. Wegener (1994) reviewed the state-of-the-art of integrated urban models according to subsystem modeled, model structure, and theory. The Oak Ridge National Laboratory compiled a report that discussed the ability of current integrated models to develop vehicle travel reduction strategies associated with energy concerns (Southworth, 1995). NCHRP (National Cooperative Highway Research Program) report 423 summarized the pros and cons of different operational land use models associated with its application (Parsons, 1998). Hunt, et al (2005) discussed six integrated models that are regarded as operational, comprehensive and integrated.

As discussed in Chapter 1, this dissertation is devoted to developing a new type of integrated land use and transportation model that models two subsystems in terms of land use and transportation in the context of urban system. Therefore, a land use model and a transportation model are two core components. The transportation model simulates travel behavior of trip-makers as well as traffic flow on physical road networks. The land use model simulates the location choice of household and employer, which leads to the distribution of household and employment. The land use model can further be decomposed into household distribution and employment distribution models in this dissertation. The household distribution model is mainly devised to estimate the number of households located in each planning zone. The employment distribution model is primarily used to forecast the number of employments aggregated in each planning zone.

When household and employer choose their location, transportation cost plays an important role in their decision-making. Consumer equilibrium theory is proposed to describe how a household choose its residential location (Alonso, 1964). It assumes that household income is equal to the summation of house cost, transportation cost, and all other expenditures.

Transportation cost is a general term; in real application, it needs to be combined with urban activities. For example, the household with a fixed employment location considers commuting cost when searching for a location to reside, while an industry prefers to choose a location with high access to supplier and customers. Accessibility as a function of transportation cost and urban activity distribution has been commonly used in modeling location choice. “Accessibility is the raison d’être of transportation system, to provide the ability for people and goods to be able to move efficiently and effectively from point to point in space in as unconstrained a fashion as possible” (Miller, 2004).

This chapter reviews existing integrated urban models in a comprehensive way. Based on model theory and model structure primarily associated with land use models, existing integrated urban models can fall into one of three major categories: gravity-based model, input-output based models, and discrete response simulation models referring to categories of land use models (Lemp et al, 2007). The representative integrated urban models under each category are discussed in detail.

2.2 GRAVITY-BASED INTEGRATED URBAN MODELS

Gravity-based integrated urban models originate from the Lowry model, which was developed for the city of Pittsburgh (Lowry, 1964). The most widely used successor to Lowy's model is integrated transportation-land use package (ITLUP).

2.2.1 Lowry Model

The Lowy model estimates spatial distribution of household and employment based on the concept of gravity. The original Newton's law states that "any two bodies attract one another with a force that is proportional to the product of their masses and inversely proportional to the square of the distance between them." By taking the concept of distance into consideration, the Lowy model assumes that the probability of making a trip is inversely proportional to the travel time (trip length) between origin and destination. It indicates that the longer the travel time between two zones, the less likely a person will make a trip between them. According to this assumption, the probability for a worker to choose a residential location is inversely proportional to the travel time between working place and residential location.

The framework of the Lowry model can be briefly described as follows. Employment is categorized as basic and service sectors. The magnitude and location of basic employment are exogenously determined by macro, regional factors such as land use plan and policy. The workers in basic sectors generate dependent households according to a regional activity ratio (the ratio of total regional households to total regional employment). These households will choose their residential locations based on the model assumption that probability of choosing a residential location is inversely proportional to the travel time between workplace and residential location. These households will be allocated into each planning zone after choosing their location. Planning zone is the geographic area dividing the planning region into relatively small areas during land use planning.

The households generated from basic employment demand services to satisfy their living needs. The number of service job created will be estimated by household-to-service multipliers in the region. The employment distribution model is then used to allocate these service jobs into planning zones based on the same model assumption. Since the Lowry model is almost the same as the land use models used in ITLUP, a detailed description of model theory and structure is given in section 2.2.2.1.

2.2.2 Integrated Transportation-Land Use Package

ITLUP is the most widely used integrated urban model currently used. It was developed under contract with the U.S. Department of Transportation (Putman, 1983). It was designed to improve long-range forecasting results by establishing the linkage between land use and transportation. This integrated urban model has been applied to nearly four dozen cities in the United States and abroad for policy analysis and planning (Wegener, 1994). For example, it is used by the Mid-America Council of Governments (Kansas City) and the Puget Sound Council of Governments (Seattle), and in Dallas-Fort Worth, Detroit, Houston, Los Angeles, Phoenix, and by the Florida DOT. ITLUP consists of two major model components: a land use model and a transportation model. The land use model is composed of disaggregate residential allocation model (DRAM) and employment allocation model (EMPAL), which were developed by Putman and colleagues (Putman, 1983, 1988; Putman, S. H. Associates, 2001). The traditional four-step travel demand model serves as the transportation model. The feedback mechanism between DRAM/EMPAL and the transportation model is built to the ITLUP framework.

2.2.2.1 Land Use Model

The logic behind DRAM lies in that the probability for a worker to choose a zone as a residential location is proportional to this zone's attractiveness and accessibility. The number of workers in zone j who are willing to choose zone i in which to reside are determined by the ratio of attractiveness and accessibility of zone i to all other zones' attractiveness and accessibility. Accessibility is a function of congested travel time between zones. The attractiveness function is expressed as the product of land area and distribution of different levels of household income groups. The original equation of DRAM can be seen in Appendix 2-A1.

The assumption in EMPAL is that the probability for a household to choose a zone in which to work is proportional to this zone's attractiveness and accessibility. EMPAL allocates employment across each planning zone not only based on attractiveness and accessibility but also on the number of jobs in the previous period, since employment distribution has a strong historic trend. EMAPL allocates employment in the future (period t+1) based on employment in base year (period t), accessibility and attractiveness in base year (period t). Accessibility function is a function of congested travel time between zones. Attractiveness is a function of land area and employment in the previous period. The original EMAPL model equations can be found in Appendix 2-A1.

The DRAM and EMPAL equations have undergone changes over time associated with data availability and application purposes (Putman, 1983, 1991, 1995; Cambridge Systematics, 2004). For example, Krishnamurthy and Kockelman (2003) used different DRAM and EMPAL model forms to investigate the propagation of uncertainty in this integrated land use and transportation model; peak travel time and off-peak travel time are utilized in the accessibility function.

2.2.2.2 Transportation Model

The traditional four-step travel demand model is implemented in ITLUP; this transportation model can be implemented using different professional software package. It consists of four sub-models for each step: trip generation, trip distribution, mode choice and traffic assignment.

During transportation model development, a study area is divided into smaller geographic areas called traffic analysis zones (TAZs). TAZs represent origins and destinations of travel activity. The trip generation model includes trip production and trip attraction; trip production estimates the number of trips produced in each TAZ; trip attraction predicts the number of trips attracted to each TAZ. The model is generally developed according to trip purpose, which typically includes home-based work (HBW: work trips that begin or end at home), home-based others (HBO: other home-based trips such as to shop or attend school that begin or end at home), non-home-based trips (NHB: trips that neither begin nor end at home).

Trip distribution is the second sub-model, which estimates the number of trips between each two TAZs based on travel time (cost) and trip generation. Mode choice, the third sub-model, is used to predict the choices that individuals or groups make in selecting transportation modes such as auto or transit to achieve their travel purposes.

Traffic assignment is the fourth sub-model in the four-step transportation model. It is the process of assigning interzonal trips to physical roadway networks using different mathematic methodologies such as user equilibrium. Its output includes traffic flow, travel time on each road segment, etc.

2.2.2.3 ITLUP Configuration

ITLUP can be a sequential procedure or an iterative procedure. In the sequential procedure, land use models (DRAM/EMPAL) estimate future household/employment distribution using base-year congested travel time or future free-flow travel time; future household/employment distribution (output of land use models) is then added to the transportation model to forecast future traffic flow. The sequential procedure does not put future congested travel time (output of transportation model) back into land use models to re-estimate future household/employment distribution (Putman, S. H. Associates, 2001). Therefore, it lacks the consistency between land use model outputs and transportation model outputs.

The iterative procedure strengthens the consistency between land use model outputs and transportation model outputs. The iterative procedure is shown in Figure 2.1. The iterative procedure has a loop between land use model and transportation model. In iterative procedure, after running the land use model and the transportation model, future congested travel time (outputs of the transportation model) is then put back into the land use model to re-estimate future household/employment distribution, which is then put into the transportation model to re-forecast future traffic flow and travel time. Iteration between the land use model and the transportation model continues until pre-defined criteria (link flow variation) are reached between the two successive iterations, or until the maximum number of iterations has been reached (Putman, S. H. Associates, 2001).

It is worth noting that the size of planning zones in the land use model is not the same as the size of TAZs in the transportation model. A planning zone is typically composed of several TAZs. After DRAM/EMAPL estimates household/employment distribution, these outputs have to be disaggregated into TAZs from larger planning zones. Once the transportation model generates the congested travel-time matrix, it has to be aggregated or "squeezed" into a larger spatial level of planning zones from TAZs. Both the aggregating and disaggregating process inevitably result in losing information about accessibility, household and employment distribution (Parsons Brinckerhoff, 1998). However, no study has been conducted to assess the errors produced in these aggregating/disaggregating processes (Tayman, 1996).

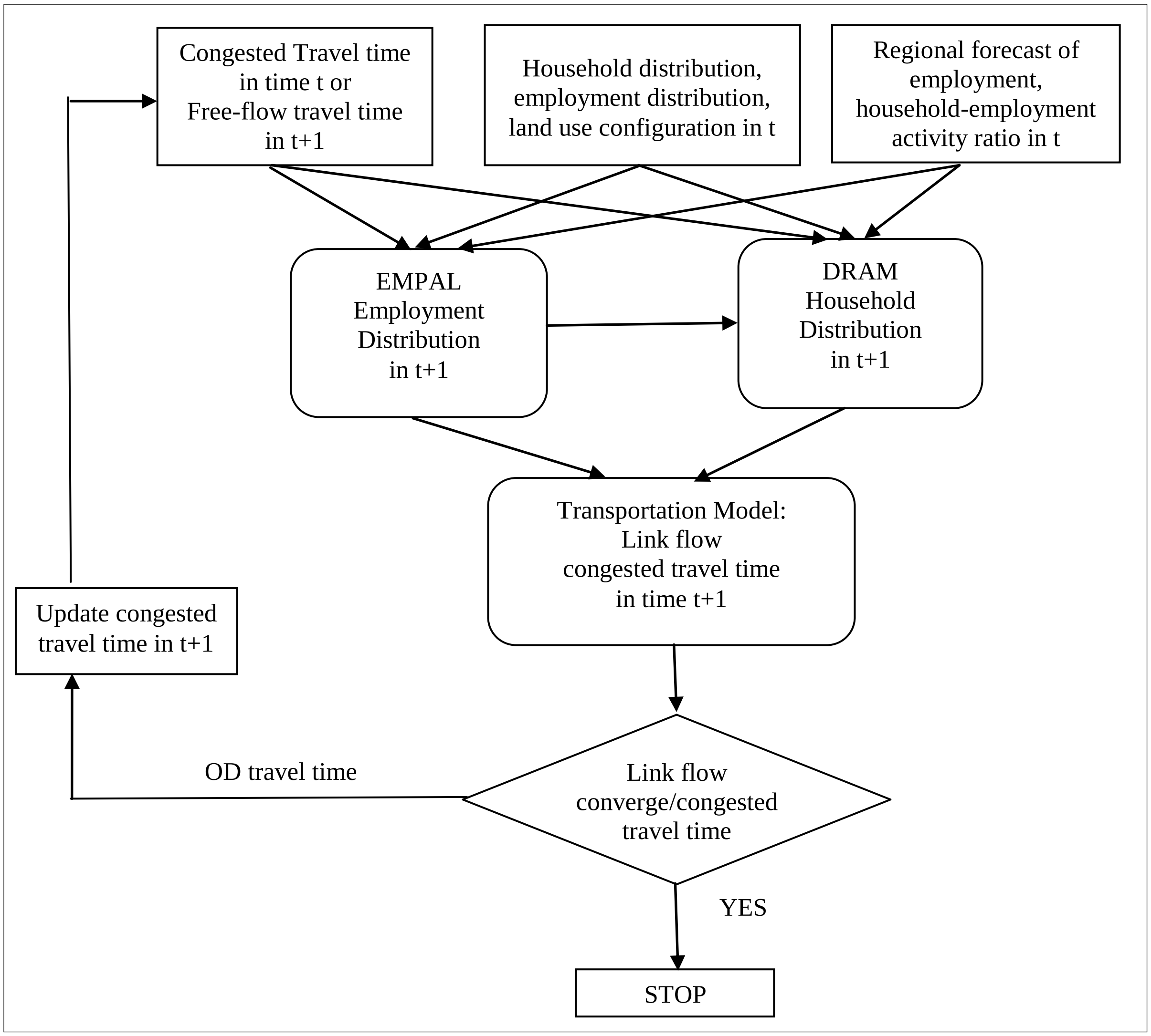

Figure 2.1: ITLUP Iterative Model Structure

It can be seen that ITLUP can assess the impact of changes in accessibility associated with transportation projects on land use. However, EMPAL/DRAM models do not incorporate land use policy variables in attractiveness function such as residential/industrial land use (as shown in Appendix 2-A). Thus, they cannot be used to evaluate the impact of land use policy on household/employment distribution (Parsons Brinckerhoff, 1998). Also, ITLUP is designed for regional planning rather than local/small community planning. This model assumes that model area is a closed system, which is suitable for a region. In a closed system, 95 percent of jobs are filled by the residents from the model area, and 95 percent workers in the model area work in the model area. Therefore, ITLUP cannot be applied on those open system areas such as the study area in this dissertation, in which a significant number of jobs are filled by workers from outside the study area.

2.2.3 Other Gravity-Based Models

Other gravity-based models are briefly introduced in this section since they are only formulated by their own authors and are not get used by other urban planning agencies. Miyagi (1989) followed the principles of the Lowry model to develop an integrated urban model that combines the residential location choice model and the transportation network equilibrium model. Mackett (1983) summarized the properties of the Leeds Integrated Land-Use Transport (LILT) model; the land use model is Lowy-based.

An interesting integrated urban model is formulated by linking the Lowy-based land use model and the combined trip distribution-assignment transportation model (Meng et al, 2000). In this integrated model, only journey-to-work trips are considered; both the land use model and the combined trip distribution-assignment mode are only theoretically formulated without discussion of calibration and implementation concerns. Assumed parameters were used to test this framework.

A quasi-gravity-based model is developed by Boyce and colleagues for home-to-work trip purpose trips (Boyce, 1978, 1980, 1986; Boyce et al 1978; Boyce et al, 1983; Boyce et al, 1988). The reason its name is quasi-gravity is that it has the same model assumption/theory as the Lowy-based model but has a different model formulation. By reinterpreting the home-to-work trip variable between each OD pair, the traditional traffic assignment model can be associated with household location choice. As for home-to-work trips, origins are related to residential location and destinations are associated with employment location. By adding entropy constraints into the traffic assignment model formulation, this integrated model is then theoretically formulated. The household distribution model emerges in the optimal conditions of this integrated urban model, which has the same model form as the DRAM in ITLUP. The original equations of quasi-gravity model are listed in Appendix 2-A2.

The entropy constraints in the traffic assignment model were discussed by Erlander and colleagues (Erlander, 1974, 1977, 1980, 1981; Erlander et al, 1979). The entropy is denoted by as “S” and $S = -\sum_i \sum_j p_{ij} \ln p_{ij}$ where p_{ij} is defined as t_{ij} / T ; t_{ij} is the number of trips from zone i to j ; and T is total tips in a model area. Erlander noted that entropy S could be explained as “a measure of the spread of distribution of journeys over the cells of trip matrix” (Erlander 1977). Boyce and Southworth (1979) explained this term as a measure of the level of spatial interaction among zones in a region. The value of S for the home-to-work trip indicates the level of interaction between residential zones and employment zones. High value of S implies that households in a residential zone are working in most of the employment zones. Low value of S indicates that households in a residential zone are working in only a few employment zones.

In the quasi-gravity-based model, the congested travel time is endogenously determined since the transportation model and household location model are simultaneously solved, while the DRAM model takes congested travel time from the output of the transportation model as exogenously determined variables. The simultaneous formulation of the household location model and transportation model provided the inspiration for developing the simultaneous model configuration used to solve the proposed integrated model in this dissertation. The combined trip distribution-assignment model provides the feasible platform to simultaneously formulate the proposed land use and transportation models.

However, this quasi-gravity integrated urban model is only theoretically formulated and not operational yet. Several areas need to be investigated before this model can be put into real application, including the method used to calibrate, how to consider trips with purposes other than home to work, and identifying the physical meaning of entropy constraints.

2.3 INPUT-OUTPUT-BASED MODEL

The input-output-based integrated urban model is based on relationship among different economic sectors. The relationship or interaction between these economic sectors is used to forecast the distribution of urban activity, person trips and commodity flows.

2.3.1 Input-Output Framework

The input-output framework has been used for urban model development since its introduction by Leontief (1967). This framework formulates an urban system as a system of equations using different economic sectors. Household and employment are divided into different economic sectors based on industry classification and household income. The input-output framework is briefly described below.

There are n economic sectors in a region: $1 \cdots n$: each economic sector has to consume some products from other economic sectors, including itself, in order to produce its own product. Let m_{ij} denote the number of units from sector i that is required to produce one unit of sector j. If production level (or total product) of each economic sector in this region is known as P_j, then $m_{ij} \times P_j$ is equal to the number of units from sector i that are used to produce P_j. It is assumed that the total product of sector i is consumed by all other sectors to meet their production levels. Therefore, the total product of sector i is equal to the summation of all consumption by other sectors:

$$m_{i1} \times P_1 + m_{i2} \times P_2 + ... + m_{in} \times P_n = P_i$$

It is assumed that the economy of this region is in balance and that the total product of each economic sector will be consumed by all other sectors. The economy of this region can be formulated as a linear system to represent the relationship between different economic sectors.

$$
\begin{aligned}
&m_{11} \times P_1 + m_{12} \times P_2 + \ldots + m_{1n} \times P_n = P_1 \\
&m_{21} \times P_1 + m_{22} \times P_2 + \ldots + m_{2n} \times P_n = P_2 \\
&. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\
&m_{n1} \times P_1 + m_{n2} \times P_2 + \ldots + m_{nn} \times P_n = P_n
\end{aligned}
$$

$$
\text{Let} \quad A = \begin{pmatrix} m_{11} & m_{12} & \ldots & m_{1n} \\ m_{21} & m_{22} & \ldots & m_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ m_{n1} & m_{n2} & \ldots & m_{nn} \end{pmatrix}
$$

"A" is called the input-output matrix or technical coefficient in the MEPLAN model (Hunt, 1994), which indicates the relationship between different economic sectors. The MEPLAN model will be discussed in the next section. "A" is the core of the input-output-based urban models. Then the linear system above can be transformed into:

$$
AP = P, \text{ where } P = \begin{pmatrix} P_1 \\ P_2 \\ \ldots \\ P_n \end{pmatrix}
$$

This is the basic framework of the Leontief input-output model. The following section will introduce the MEPLAN model framework, which is the most widely used input-output-based urban model.

2.3.2 MEPLAN Model

The MEPLAN model framework has been in use for 25 years and is the second most widely used integrated urban model (Wegener, 1994). It has been applied to more than a dozen urban regions, including Greater London, the United Kingdom; Naples, Italy; and Sacramento, California (Wegener, 1994; Hunt, 1994; Abraham and Hunt, 1999). This model was developed under the leadership of Marcial Eschenique and is the property of Marcial Echenique and Partners (MEP) firm (Echenique et al, 1990).

The core feature of the MEPLAN model is the relationship/interaction between different economic sectors in a region. The product of each sector in a zone will be transported into all zones for consumption; this generates economic interactions among different economic sectors in zones. Input-output framework and random utility choice modeling are two major components in this model framework. First, the input-output framework is utilized to estimate the number of products of each sector which will be consumed in a zone in order to meet the production level of each economic sector in this zone. This estimation is achieved by using the technical coefficient. After the consumption of each sector in a zone is obtained, a random utility choice model is developed to allocate this consumption into all other zones for the purpose of production. Two major variables in this utility function are transportation cost and production cost. A detailed description of the original MEPLAN equations is shown in Appendix 2-B.

The economic sectors can be categorized based on industry classification and household income. For example, in the Sacramento model, industries are divided into: agriculture, manufacturing, service and office, retail, health, education, government, private education, commercial transportation, and wholesale; households are divided into low-income, mid-income, and high-income households. The industries and households will occupy lands at different rates and prices. Household sectors provide the labor force to other industry sectors and also generate person trips.

The interactions among different sectors in various zones bring about freight flow and person-trip flow, which generates the demand for transportation. Freight flow and person-trip flow are distributed across different transportation modes and are then assigned to the physical transportation network, which produces transportation costs between zones for each sector.

Clay and Johnston (2006) discussed the error and uncertainty propagation in every step of the MEPLAN model. They concluded that commercial trip generation rates have the most effect on model outputs after comparison with other socio-economic input. Zhao and Kockelman (2004) investigated the existence and uniqueness of input-output-based model solution. The models were proved to have a unique solution and to be solved with convergence by fixed point algorithm.

By incorporating different industry and household factors, MEPLAN is able to evaluate a variety of land use policies (Abraham, et al, 1999). However, this model requires a large amount of data such as production costs and land use prices in each economic sector, which are not normally collected by typical urban areas and MPOs. The structure of MEPLAN focuses on the economic interaction between different economic factors; it is suitable for regional or intercity modeling rather than for the typical urban area.

2.3.3 Other Input-Output Based Models

Other input-output-based urban models will be briefly discussed since they either have a similar model framework as MEPLAN or they are only formulated by their own authors without calibration and application in urban planning. TRANUS is an integrated urban model based on input-output framework (de la Barra et al, 1984; de la Barra, 1989). It is similar to the MEPLAN model framework; both have the same structure in model framework and concept (Hunt et al, 2005). Most of the descriptions in MEPLAN apply to TRANUS.

Another integrated urban model was developed by Kim based on input-output framework (Kim, 1989 and 1990). This model aims to find general equilibrium between demand and supply associated with transportation infrastructures and activity locations in a strict economic sense. The model is formulated as a standard linear programming with an objective function and four constraints. The objective is to minimize total cost, including production costs and transportation costs under resource constraints and production-consumption equilibrium constraints. Resource constraints include:

Export constraints: the production of an economic sector should at least satisfy export needs of this sector.

Land constraints: all lands used by different economic sectors and transportation systems cannot exceed the amount of available land in the study region.

Transportation constraints: the transportation supply (including mode and capacity) should satisfy the needs of transporting freight flows and person-trip flows.

Production-consumption equilibrium constraints use the input-output framework is utilized to describe the relationship among different economic sectors in zones. In this model, the transportation network is converted into units of input for carrying each economic sector, such as operating cost per mile to move a unit of sector. This model requires extensive data sources that are not regularly gathered by a typical urban area such as operating cost per mile and export need. To date, the model has not been applied to any region for the purpose of urban planning or policy analysis.

Jun (1999) used the input-output framework to develop an integrated metropolitan model, which examines the interspatial relationship between the demographic-economic system and the transportation system. However, this model is only theoretically formulated with no discussion of calibration and implementation.

2.4 DISCRETE RESPONSE SIMULATION MODEL

This category of urban models aims to simulate household and employment location choice. It mainly takes into consideration transportation measures generated from the transportation model along with other variables in modeling household and employment location choice. The most widely used methodologies are discrete choice theory and bid rent theory.

2.4.1 Discrete Choice Model

Discrete choice theory (random utility maximization) can be utilized to develop the location choice model (Mcfadden, 1974, 1978, 2001; Domencich and McFadden 1975). This approach is to estimate the choice between mutually exclusive alternatives on the basis of attributes of these alternatives. The attributes of these alternatives are described by a utility function. When this approach is applied in the context of urban models, it is used to forecast which zone among all zones in a city will be chosen by a household or employer to reside in.

The most popular urban model based on discrete choice theory is UrbanSim, which was developed under the program of travel model improvement program (TMIP) (Waddle, 2002; Waddle et al, 2007). In UrbanSim, a city is divided into many grid cells so that households and employers can make location choices among these cells. The utility function is used to describe attributes of these grid cells associated with different categories of variables. For example, in the employment location choice model, variables include real estate attributes, land use composition, land value in the immediate neighborhood, land use mix, the number of jobs in the neighborhood, and transportation measures such as accessibility to labor and consumers (Waddle and Ulfarsson, 2003a). In the household location choice model, variables include house price, development types, neighborhood employment, neighborhood land use mix, transportation measures such as accessibility to jobs and travel time to central business district (CBD) (Waddle, et al. 2003b).

Transportation measures are generated from the transportation model outputs. However, the transportation model is usually not performed in the same time frame as UrbanSim. For example, in the Wasatch Front Regional Council integrated urban model, UrbanSim is performed every year to estimate household and employment distribution; the transportation model (traditional four-step travel demand model) is only performed every five years. Therefore, transportation measures used in the UrbanSim are not updated until the transportation model is performed (Waddle et al, 2007).

POLIS (Projective Optimization Land Use Information System) is developed based on this theory to allocate urban activities subject to planning constraints such as land supply (Prastacos, 1986a, 1986b); CUF-I/CUF-II (California Urban Future models) is also developed from this theory to allocate new development into grid cells (Landis and Zhang, 1998a, 1998b).

2.4.2 Bid-Rent Model

The concept lying the bid-rent urban model is that consumers (households and developers) choose their location based on the lowest price, and land owners want to sell their land (location) for maximum profit. Market equilibrium can be reached after balancing the interaction between consumers and land owners. In the land bid-rent process, consumer surplus can be used to describe consumer behavior, which is defined as the difference between the price that consumers are willing to pay and the actual paid price.

Martinez (1991, 1992a, 1992b) developed a five-stage urban model based on bid-rent theory. In this model, accessibility measures produced from the transportation model are used to represent the transportation characteristics of specific land lots. The willingness to pay (WP) function is utilized to describe the attributes of land lots or zones. In its household location choice model, variables the in WP function consist of proportion of residential land, accessibility measures, neighborhood characteristics such as neighborhood land use mix, and land price. In its employment location choice model, variables in WP function include proportion of industry land, employment accessibility measures, neighborhood employment, and land price. Appendix 2-C shows the original equations of the household location choice model.

This model uses transportation measures produced from the transportation model to estimate household/employment location choice. There is no feedback between this urban model and the transportation model. Other urban models based on the bid-rent concept include the RURBAN model developed by Miyamoto and Kitazume (1989), and the bi-level transport and residential location model (Chang and Mackett, 2006).

2.5 LIMITATIONS OF EXISTING MODELS

Integrated urban models have recently been introduced and implemented in several metropolitan areas and regions for different application purposes. Since these models target large metropolitan areas, they lack the flexibility to be applied to small urban areas. For example, the most widely used integrated model, ITLUP, assumes that the model area is a closed system, i.e., there are no significant percentage of households who are employed outside of the model area, and no significant percentage of jobs are occupied by the workers from outside of the model area (Putman, 1983, 1988; Putman, S. H. Associates, 2001). But in some small urban areas such as the study area in this dissertation, a considerable number of households are employed outside of the area (in the neighborhood city); a substantial number of jobs in the model area are occupied by people who live outside the model area (in the neighborhood city). Therefore, it is appropriated to consider such small urban areas as open systems, which is not in accord with the assumption of ITLUP.

The second widely used urban model (MEPLAN) based on an input-output framework is suitable for regional or intercity modeling (Hunt and Simmonds, 1993; Hunt, 1994). This model is more appropriate for interurban areas (Parsons, 1998). It is questionable whether this model is appropriate for small urban areas (Lemp, et al, 2007).

In both ITLUP and MEPLAN, the planning zone in the land use model typically consists of several TAZs associated with the transportation model. Information exchange such as travel time and household distribution between planning zones and TAZs will cause substantial information loss for small urban areas. Also, by aggregating several TAZs into one planning zone, these models produce only a few planning zones in a small urban area. It is difficult to develop the land use model with few planning zones because of the small sample size.

The recently developed urban model (UrbanSim) requires intense data at disaggregate spatial level such as parcels. It requires over a thousand parameters and tens of thousands of variable values to develop and calibrate the model. Very few regions or metropolitan areas routinely collect all of this data, not to mention small urban areas.

Most integrated urban models use the traditional four-step travel demand model as the transportation model. Few efforts have been made to integrate the combined trip distribution-assignment transportation model and the land use model. The combined trip distribution-assignment model has been successfully formulated by researchers. It has not been developed and calibrated for a real application within the context of integrated urban models.

Existing integrated urban models require a substantial budget and professional personnel to perform the land use model. Small urban areas are typically not able to afford the budget and professional crew needed to develop these land use models. The integrated land use and transportation model framework described in this dissertation is designed to develop an affordable and easy-to-implement land use model based on available data, which can be applied to both metropolitan areas and small urban areas. Chapter 5 compares the proposed land use model to the existing land use models. The three components of this framework are the transportation model, the land use model, and the interaction between these two models, which will be discussed in the following chapters.

CHAPTER 3
DATA COLLECTION AND PREPARATION

The development of land use and transportation models demands extensive data collection and preparation, including land-use structure data, household and employment distribution data, and transportation network and performance data. This chapter will discuss data collection and data preparation. The data collection section describes the data sources used for the model development. The data preparation section focuses on the development of transportation performance data.

3.1 DATA COLLECTION

3.1.1 Land Use Data Description

Woodford County in Kentucky was selected to test the proposed model framework. This county is located between Lexington (the second largest city in Kentucky) and Frankfort (Kentucky's capital). There are two towns, Versailles and Midway, located on the major highways of US60 and US62. The rest of the county consists of rolling farmland and timber stands.

In the process of land use and transportation planning, an urban area is spatially divided into small zones for the purpose of analysis. In the proposed integrated land use and transportation model framework, both land use and transportation models are established at the same spatial level: TAZ. The TAZ configuration of Woodford County is illustrated in Figure 3.1. The county contains 78 TAZs.

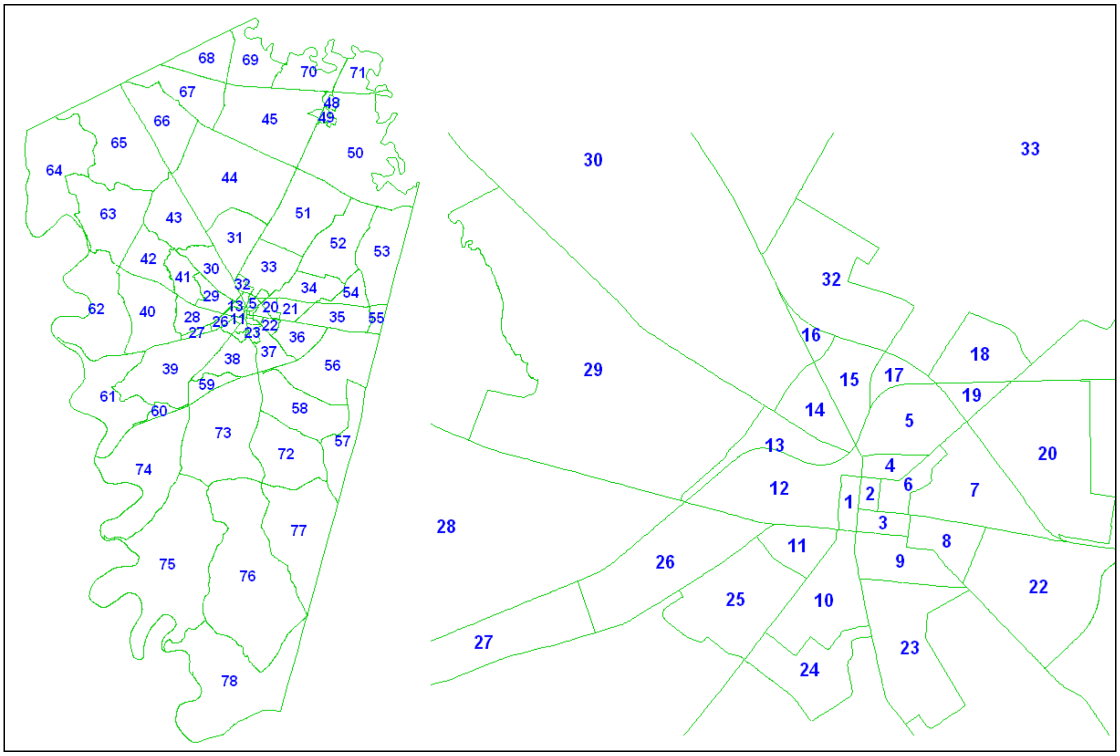

Figure 3.1: Woodford County TAZ Configuration

The study area has a population of 23,208, with 8,893 households and 9,486 jobs for the base year of 2000. Multiple data sources exist for developing the proposed integrated model. 2000 U.S. Census survey data provides the household and population distribution at census block level and the spatial commuting pattern. Dun & Bradstreet (D&B) employment survey data provides the number of employments/jobs in each census block. A TAZ is typically composed of several census blocks. Therefore, the data at census block level are aggregated into TAZ level for the integrated model development.

Land use parcel data is obtained from the Woodford County Planning Commission. It provides area coverage by different land use type in each TAZ, such as the area/footprint of residential land use, mobile home and multi-family residential land use, agriculture and preserved agriculture land use, industrial land use, commercial land use, professional office and institutional land use, vacant land use, and other land use. These data will be used for the land use model development.

3.1.2 Transportation Network Data

One of the most important goals of developing a transportation model is to evaluate transportation system performance. In doing so, detailed transportation networks need to be represented in the model. The detailed transportation network for this study area is based on a road network map downloaded from the Kentucky Transportation Cabinet's (KYTC) website. The required network attributes for the transportation model development consist of functional classification, number of lanes, speed limit, and traffic count. The road network for Woodford County by functional classification is shown in Figure 3.2.

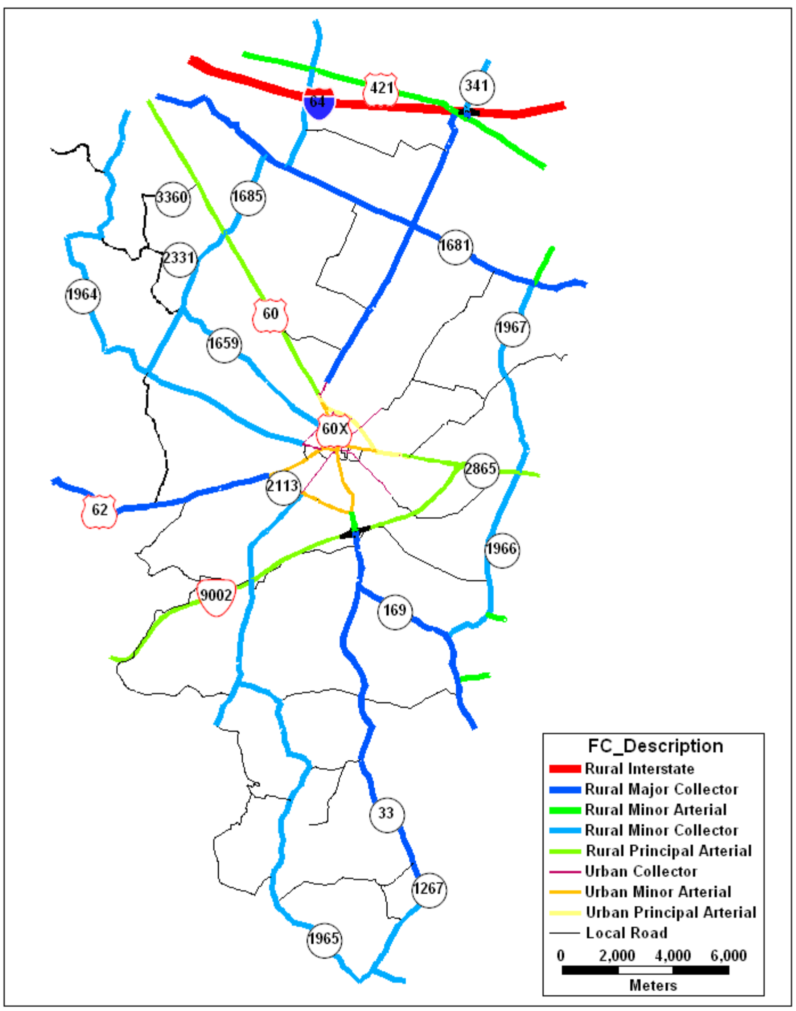

Figure 3.2: Woodford County Roadway Network by Functional Classification

US Highway 60 linking Franklin, Woodford County, and Fayette County provides the mobility for workers to commute within reasonable travel time among these counties. With scenic horse farms and gently rolling hills, it has become the prominent corridor serving Woodford County. Interstate 64 and Bluegrass Parkway (Route 9002 in the map) has one interchange in Woodford County, providing easy access to major interstate and parkway systems in Kentucky for local residents.

The transportation network considered in this study consists of approximately 251 miles of road. The mileage of each functional class is illustrated in Table 3.1 with the percentage of each functional class over total mileage of all roads.

Table 3.1: Functional System Mileage

Functional System	Mileage	Percent of Total Mileage
Rural Interstate	17.0	6.8%
Rural Principal Arterial	39.0	15.5%
Rural Minor Arterial	9.6	3.8%
Rural Major Collector	41.9	16.6%
Rural Minor Collector	52.8	21.0%
Urban Principal Arterial	5.4	2.1%
Urban Minor Arterial	9.3	3.7%
Urban Collector	5.4	2.1%
Local Road	71.3	28.3%
Total	251.7	100%

3.2 TRANSPORTATION SYSTEM DATA

One of the three components in the proposed integrated land use and transportation model is the combined trip distribution-assignment transportation model. A crucial step in developing this model is to calibrate the model. The calibration requires an OD trip table and an OD travel time table. The OD trip table is usually obtained by conducting a household travel survey; unfortunately, an OD survey is not conducted by most small urban areas such as this study area because it is difficult and expensive. Since an observed OD trip table is unavailable, this study uses an estimated OD trip table and a travel time table. These tables can be generated from the output of the Woodford travel demand model, which is developed using the traditional four-step method and default parameters recommended by NCHRP Report 365 (Martin et al, 1998). The procedure for developing this Woodford travel demand model (WTDM) is shown in Appendix 3-A. The estimated OD trip and travel time tables are regarded as being reasonable since WTDM has the satisfactory error bound between the observed traffic volume and the modeled traffic volume. Percent Root Mean Squared Error (PRMSE), as defined in equation 3.1, is used to measure the difference between the observed and the modeled traffic volumes.

$$PRMSE = \frac{\sqrt{\sum_{i}^{N}(\hat{v}_n - v_n)^2 / N}}{\sum_{n}^{N} v_n / N} \qquad 3.1$$

Where

$\hat{v}_n$: Estimated traffic volume on traffic count station n

v_n : Observed traffic volume on traffic count station n

N : Total number of traffic count stations in the area

The acceptable PRMSE in common practice is under 30% (Kentucky Transportation Cabinet, 2004). The PRMSE of WTDM is 25.7%. The comparison between the modeled and the observed traffic volumes on each traffic count station is shown in Appendix 3-B. The estimated base-year OD trip table and OD travel-time table are shown in Table 3.2 and 3.3. These tables will be regarded as the observed OD trip and travel time tables for the development of the combined trip distribution-assignment model. It is important to note that trips in Table 3.2 are vehicle trips per day. In addition, the OD trip table is symmetrical because WTDM is a daily model with the assumption that trips originating from a TAZ will return to this TAZ in the same day. In these two tables and other tables in this dissertation, TAZ is abbreviated as Z; for example, TAZ 1 is expressed as Z1. Since there are 78 TAZs in this study area, it is difficult to show all

OD trips between each two OD pairs, so only 21 TAZs are selected to demonstrate the OD trip table for illustration purpose.

OD trips between two TAZs are internal trips. Also, a number of trips on the transportation network are associated with external TAZs located outside the study area. These trips are regarded as external trips that have at least one end outside the study area. The external trips are also estimated in WTDM in Appendix 3-A. Estimated external trips will be regarded as given variables or background traffic in the combined trip distribution-assignment transportation model development.

After the required data is generated, the transportation model development is then discussed in Chapter 4. The land use model is discussed in Chapter 5, and the interaction between these two models is presented in Chapter 6.

Table 3.2: Estimated OD Trip Table

Trips	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9	Z10	Z11	Z12	Z13	Z14	Z15	Z16	Z17	Z18	Z19	Z20	…	Z78
Z1		2.5	7.0	2.9	15.0	17.5	5.1	9.4	20.4	33.8	8.5	12.0	5.2	9.1	8.8	0.2	0.3	0.0	2.8	1.2	…	3.8
Z2	2.5		2.6	1.0	3.1	8.4	3.1	2.9	5.1	7.0	1.5	3.4	1.1	2.3	1.8	0.1	0.3	0.0	0.8	1.3	…	0.9
Z3	7.0	2.6		2.0	7.2	10.2	8.5	14.6	11.0	18.7	3.8	9.7	2.5	6.3	4.2	0.4	0.9	0.0	1.7	4.6	…	2.2
Z4	2.9	1.0	2.0		3.4	1.8	15.5	2.3	1.7	5.2	0.8	4.7	0.7	2.8	1.3	0.4	1.3	0.0	0.8	5.4	…	0.6
Z5	15.0	3.1	7.2	3.4		3.6	56.2	8.0	4.3	21.7	2.8	26.2	2.7	16.2	6.0	2.6	5.9	0.1	2.0	26.2	…	1.2
Z6	17.5	8.4	10.2	1.8	3.6		20.0	8.8	6.7	16.1	2.2	12.0	1.1	6.8	2.4	1.0	2.4	0.0	0.7	11.4	…	0.6
Z7	5.1	3.1	8.5	15.5	56.2	20.0		13.1	18.8	40.6	8.5	28.0	10.6	20.4	17.5	0.6	2.6	0.0	19.4	9.9	…	8.3
Z8	9.4	2.9	14.6	2.3	8.0	8.8	13.1		15.0	26.6	4.9	13.9	2.7	8.5	4.8	0.7	1.8	0.0	2.2	9.0	…	3.0
Z9	20.4	5.1	11.0	1.7	4.3	6.7	18.8	15.0		25.8	4.0	14.2	1.3	7.9	2.8	1.2	2.4	0.0	0.8	12.1	…	0.9
Z10	33.8	7.0	18.7	5.2	21.7	16.1	40.6	26.6	25.8		18.3	46.7	7.3	28.0	13.3	3.0	6.3	0.1	4.3	30.4	…	7.9
Z11	8.5	1.5	3.8	0.8	2.8	2.2	8.5	4.9	4.0	18.3		8.4	0.9	4.8	1.8	0.7	1.3	0.0	0.5	6.3	…	0.8
Z12	12.0	3.4	9.7	4.7	26.2	12.0	28.0	13.9	14.2	46.7	8.4		16.8	34.8	15.6	2.3	4.5	0.1	5.9	17.7	…	10.4
Z13	5.2	1.1	2.5	0.7	2.7	1.1	10.6	2.7	1.3	7.3	0.9	16.8		8.5	1.8	1.1	2.0	0.0	0.4	7.3	…	0.4
Z14	9.1	2.3	6.3	2.8	16.2	6.8	20.4	8.5	7.9	28.0	4.8	34.8	8.5		17.6	4.3	6.0	0.1	4.1	16.5	…	4.5
Z15	8.8	1.8	4.2	1.3	6.0	2.4	17.5	4.8	2.8	13.3	1.8	15.6	1.8	17.6		3.2	4.2	0.0	0.9	11.7	…	0.9
Z16	0.2	0.1	0.4	0.4	2.6	1.0	0.6	0.7	1.2	3.0	0.7	2.3	1.1	4.3	3.2		0.0	0.0	0.7	0.0	…	0.6
Z17	0.3	0.3	0.9	1.3	5.9	2.4	2.6	1.8	2.4	6.3	1.3	4.5	2.0	6.0	4.2	0.0		0.0	3.5	0.0	…	1.7
Z18	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.0	0.0		0.0	0.0	…	0.0
Z19	2.8	0.8	1.7	0.8	2.0	0.7	19.4	2.2	0.8	4.3	0.5	5.9	0.4	4.1	0.9	0.7	3.5	0.0		18.0	…	0.3
Z20	1.2	1.3	4.6	5.4	26.2	11.4	9.9	9.0	12.1	30.4	6.3	17.7	7.3	16.5	11.7	0.0	0.0	0.0	18.0		…	10.4
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
Z78	3.8	0.9	2.2	0.6	1.2	0.6	8.3	3.0	0.9	7.9	0.8	10.4	0.4	4.5	0.9	0.6	1.7	0.0	0.3	10.4	…	

Table 3.3: Estimated OD Travel Time

Travel Time	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9	Z10	Z11	Z12	Z13	Z14	Z15	Z16	Z17	Z18	Z19	Z20	…	Z78
Z1		0.5	0.7	1.2	1.2	0.6	1.4	0.9	0.6	1.1	0.9	1.8	1.4	1.7	1.2	1.3	2.1	2.3	1.9	2.8	…	17.4
Z2	0.5		0.6	0.9	1.5	0.3	1.1	0.8	0.5	1.4	1.1	2.0	1.6	1.9	1.4	1.6	2.1	2.0	1.7	2.5	…	17.6
Z3	0.7	0.6		1.2	1.6	0.7	1.5	0.6	0.7	1.5	1.3	2.2	1.8	2.1	1.6	1.8	2.4	2.3	2.0	2.9	…	17.8
Z4	1.2	0.9	1.2		1.0	0.9	0.7	1.5	1.1	2.0	1.7	2.2	1.8	2.1	1.6	1.8	1.6	1.6	1.2	2.1	…	18.2
Z5	1.2	1.3	1.6	1.0		1.4	0.9	1.8	1.5	2.1	1.8	1.9	1.5	1.8	1.3	1.5	1.9	1.8	1.5	2.3	…	18.3
Z6	0.6	0.3	0.7	0.9	1.5		1.2	0.9	0.6	1.4	1.2	2.1	1.7	2.0	1.5	1.6	2.1	2.0	1.7	2.6	…	17.6
Z7	1.4	1.1	1.5	0.7	0.9	1.2		1.7	1.3	2.2	1.9	2.1	1.7	2.0	1.5	1.7	1.3	1.2	0.9	1.7	…	18.4
Z8	0.9	0.8	0.6	1.5	1.8	0.9	1.7		0.6	1.5	1.2	2.4	2.0	2.3	1.8	2.0	2.5	2.4	2.1	2.9	…	17.7
Z9	0.6	0.5	0.7	1.1	1.5	0.6	1.3	0.6		1.1	0.9	2.1	1.6	1.9	1.4	1.6	2.3	2.2	1.9	2.7	…	17.3
Z10	1.1	1.4	1.5	2.0	2.1	1.4	2.2	1.5	1.1		1.2	2.6	2.2	2.5	2.0	2.2	3.0	3.1	2.8	3.6	…	17.9
Z11	0.9	1.1	1.3	1.7	1.8	1.2	1.9	1.2	0.9	1.2		2.4	1.9	2.2	1.8	1.9	2.7	2.8	2.5	3.3	…	17.6
Z12	1.8	1.9	2.2	2.1	1.9	1.9	2.1	2.4	2.1	2.6	2.4		1.1	1.8	1.8	1.8	2.6	2.7	2.7	3.5	…	18.9
Z13	1.4	1.5	1.8	1.7	1.5	1.5	1.7	2.0	1.6	2.2	1.9	1.1		1.4	1.4	1.4	2.2	2.3	2.3	3.1	…	18.4
Z14	1.7	1.8	2.1	2.0	1.8	1.8	2.0	2.3	1.9	2.5	2.2	1.8	1.4		1.1	0.8	1.6	1.8	2.1	2.9	…	18.7
Z15	1.2	1.3	1.6	1.5	1.3	1.3	1.5	1.8	1.4	2.0	1.8	1.8	1.4	1.1		0.8	1.6	1.7	2.0	2.8	…	18.2
Z16	1.3	1.5	1.8	1.7	1.5	1.5	1.7	2.0	1.6	2.2	1.9	1.8	1.4	0.8	0.8		1.3	1.5	1.7	2.6	…	18.4
Z17	2.2	2.0	2.4	1.6	1.8	2.1	1.2	2.4	2.3	3.0	2.8	2.6	2.2	1.7	1.6	1.4		0.9	1.2	2.0	…	19.2
Z18	2.3	2.0	2.3	1.5	1.8	2.0	1.2	2.4	2.2	3.1	2.8	2.8	2.3	1.8	1.7	1.5	0.9		1.2	2.0	…	19.3
Z19	1.9	1.7	2.0	1.2	1.5	1.7	0.9	2.1	1.9	2.8	2.5	2.7	2.3	2.1	2.0	1.8	1.2	1.1		1.3	…	19.0
Z20	2.8	2.5	2.8	2.1	2.3	2.6	1.7	2.9	2.7	3.6	3.3	3.5	3.1	2.9	2.8	2.6	2.0	1.9	1.3		…	19.8
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
Z78	17.4	17.6	17.8	18.2	18.3	17.6	18.4	17.7	17.3	17.9	17.6	18.9	18.4	18.7	18.2	18.4	19.2	19.3	19.0	19.8	…	

CHAPTER 4
TRANSPORTATION MODEL DEVELOPMENT

The proposed integrated model includes three components: a transportation model, a land use model, and the interaction between these two models. Transportation demand is driven by human social activities, which involves making trips to satisfy needs. Transportation models assign these trips generated by households and employments to between TAZs to the road network. The resulting traffic flow patterns are then used to estimate transportation performance measures. Taking transportation performance measures and land use structure variables into consideration, the regression-based land use model is developed to identify those variables that have a significant impact on household and employment distribution, and to find the appropriate combination of these significant variables to estimate household and employment distribution. The interaction between the land use model and the transportation model is then investigated by two different methodologies: feedback model configuration and simultaneous model configuration. The interaction between these two models will strengthen the consistency between the land use model output and the transportation model output by showing that household/employment distribution is in accord with transportation system performance.

This chapter outlines the steps to develop the combined trip distribution-assignment model. In state-of-the-art transportation modeling, three categories of transportation models have been formulated: the traditional four-step travel demand model, the combined trip distribution-assignment model, and the activity-based model. The traditional four-step travel demand model has been widely put into practice by planning agencies; this model was introduced in ITLUP framework in Chapter 2 and is discussed in Appendix 3-A. The activity-based transportation model is designed to simulate second-by-second movements of each individual person and each vehicle across transportation networks, which requires very detailed activity location data and household travel survey data. The activity-based transportation model is still in development and is not broadly accepted by planning agencies. The combined trip distribution and assignment model has been successfully formulated (Wilson, 1970; Sheffi, 1992), but it has rarely been put into practice in the field of integrated urban models. The transportation model developed in this dissertation will adopt the combined trip distribution-assignment model. This model is the first to utilize the combined transportation model in the operational integrated urban model field. Also the combined trip distribution-assignment model framework provides a feasible platform to simultaneously formulate the land use model and the transportation model, which will be discussed in simultaneous model configuration in Chapter 6.

4.1 COMBINED TRIP DISTRIBUTION-ASSIGNMENT MODEL

4.1.1 Model Formulation

As mentioned in Chapter 3, the study area contains 78 TAZs, which are connected to each other through the transportation network. The combined trip distribution-assignment model is designed to estimate the number of trips between each two TAZs (trips refer to vehicle trips per day in the following discussion) and assign these trips to the transportation network. The variables as well as the set associated with this transportation model are defined below.

Each TAZ serves as both origin and destination for trips; a trip has two ends, origin and destination. Let I denote the set of origin TAZs, and J denote the set of destination TAZs. The components of both set I and set J are the 78 TAZs in the study area. Let O_i denote the trips originating from TAZ i, ($\forall i \in I$); let D_j denote the trips destined to TAZ j, ($\forall j \in J$); let t_{ij} denote the number of trips between origin i and destination j, ($\forall i \in I$ and $\forall j \in J$). Each OD is connected by road segments and there are multiple paths that travelers can use.

Consider the road network; let A denote the set of links (road segments) in the study area. tt_a^f denotes free-flow travel time on link a, ($\forall a \in A$), which can be calculated by link length divided by free-flow speed. CP_a denotes the capacity of link a, ($\forall a \in A$). v_a denotes the traffic volume on link a, ($\forall a \in A$); tt_a denotes the congested travel time on link a with traffic flow v_a on this link, ($\forall a \in A$). The widely used Bureau of Public Road (BPR) function is adopted as link impedance function for relating travel time to traffic volume: $s_a(v_a) = tt_a = tt_a^f \left[1 + 0.15(\frac{v_a}{CP_a})^4 \right]$. There are multiple paths connecting each OD pair; R_{ij} denotes the set of paths from origin i to destination j, ($\forall i \in I$ and $\forall j \in J$). R denotes the set of complete paths which connect all OD pairs in the study area, therefore $R = \cup_i \cup_j R_{ij}$, ($\forall i \in I$ and $\forall j \in J$). h_r denotes the number of trips on path r, ($\forall r \in R$). Since a path between an OD pair usually consists of several links, the relationship between links and paths needs to be defined. δ^{ar} denotes the incidence coefficient to describe the relationship between path and link; $\delta^{ar} = 1$ if link a is on path r; otherwise $\delta^{ar} = 0$; therefore $v_a = \sum_{r \in R} h_r \delta^{ar}$, ($\forall a \in A$ and $\forall r \in R$).

Researchers have conducted extensive studies in the formulation of a combined trip distribution-assignment model (Sheffi, 1992; Boyce et al, 1988, Evan, 1976). The combined trip distribution-assignment model has a computational advantage since results can be obtained relatively faster than by using the traditional four-step model. Also, its performance is significantly better than the traditional four-step travel demand model in comparing model results (Zargari et al, 2008). Meng et al (2000) described an integrated urban model by linking a combined trip distribution-assignment model with a Lowry-based land use model for home to work trips. However, the combined trip distribution-assignment model in their study is only tested using assumed parameters and is not calibrated at all.

This chapter presents the complete calibration and application of the combined trip distribution-assignment model based on a real network. A combined trip distribution-assignment model can be formulated as the following optimization program (Sheffi, 1992).

$$Min\sum_a \int_0^{v_a} s_a(w)dw + \frac{1}{\beta}\sum_i \sum_j (t_{ij}\ln t_{ij} - t_{ij}) \quad 4.1a$$

$S.T.$

$$\sum_j t_{ij} = O_i \qquad (i \in I) \quad 4.1b$$

$$\sum_i t_{ij} = D_j \qquad (j \in J) \quad 4.1c$$

$$\sum_{r\in R_{ij}} h_r = t_{ij} \qquad (i \in I\,,\, j \in J) \quad 4.1d$$

$$v_a = \sum_{r\in R} h_r \delta^{ar} \quad (a \in A) \quad 4.1e$$

$$h_r \geq 0 \qquad (i \in I\,,\, j \in J\,, r \in R_{ij}) \quad 4.1f$$

In this formulation (equation 4.1a-f), O_i and D_j are given variables; t_{ij} and v_a are unknown variables that can be obtained after solving this optimization problem. Link impedance function ($s_a(v_a) = tt_a = tt_a^f\left[1 + 0.15(\frac{v_a}{CP_a})^4\right]$) is used in this original formulation. However, in this study area, only internal trips between each OD pair inside the study area are considered; external trips are considered as given variables and discussed in the last chapter. The assignment of these external trips onto networks produces background traffic volume (or preloaded traffic) on each link. The original link impedance function has to take this background traffic volume into consideration. Let BG_a denote the background traffic volume resulting from external trips on link a, ($\forall a \in A$). The original link impedance function will be transformed into $s_a(v_a) = tt_a = tt_a^f\left[1 + 0.15(\frac{v_a + BG_a}{CP_a})^4\right]$ as the modified BPR function, which will be used in this study.

The first item in the objective function is the sum of integrals of link impedance function. This item does not have any physical meaning; it is only constructed to satisfy user equilibrium conditions. User equilibrium conditions state that all used paths between each OD pair must have equal travel time and no road user can improve his/her travel time by switching paths (Wardrop, 1952).

The second item in the objective function does not have any physical meaning either. This item ensures that trip distribution has greatest number of states during the procedure of allocating trip generation in each TAZ (O_i, $\forall i \in I$ and D_j, $\forall j \in J$) into each OD pair (t_{ij}, $\forall i \in I$ and $\forall j \in J$) (Wilson, 1970). In this item, β is an empirically determined parameter based on the observed data whose value is greater than 0. The procedure of obtaining the value of β will be discussed in the section on model calibration.

Constraint 4.1b denotes that trips originating from i are the summation of trips from i to all other destination TAZs. Constraint 4.1c represents that the trips terminating in destination j are the summation of trips from all other origins to this destination. Constraint 4.1d refers to flow conservation, which states that the summation of all path flows that connect an OD pair is equal to trip exchange between this OD pair. Constraint 4.1e refers to definitional constraint, which describes network structure by formulating the relationship between path flow and link flow. The non-negativity constraint 4.1f ensures that path flow is always greater than or equal to zero, which makes the solution of this program physically meaningful since there is no negative trip in the real world.

The combined trip distribution-assignment model with the original BPR function in the objective function can meet user equilibrium conditions, has trip distribution functions in the gravity form. These have been discussed in several studies (Sheffi, 1992; Boyce et al, 1988, Evan, 1976). The purpose of the following section is to prove the combined trip distribution-assignment model with the modified BPR function can still arrive at the same results as those with the original BPR function.

4.1.2 User Equilibrium Conditions and Trip Distribution Function

The purpose of this section is to prove that the optimality condition of this program can satisfy user equilibrium conditions, and to derive the trip distribution function in the proposed combined trip distribution-assignment model. The trip distribution function can be used to determine the value of β through the model calibration process. To derive the optimality conditions of this optimization program, the Lagrangian of this minimization problem can be formulated as equation 4.2.

$$
\begin{aligned}
L = & \sum_a \int_0^{v_a} s_a(w)dw + \frac{1}{\beta}\sum_i \sum_j (t_{ij} \ln t_{ij} - t_{ij}) + \sum_i \lambda_i (O_i - \sum_j t_{ij}) \\
& + \sum_j \sigma_j (D_j - \sum_i t_{ij}) + \sum_i \sum_j u_{ij}(t_{ij} - \sum_{r\in R_{ij}} h_r) + \sum_{r\in R} -\theta_r h_r
\end{aligned}
\qquad 4.2
$$

λ_i, σ_j, u_{ij}, θ_r are the Lagrange multipliers, denoting the dual variables associated with each corresponding constraint.

In definitional constraint 4.1e (incidence relationship between link flow and path flow), the derivative of link flow with respect to a particular path flow equals to an incidence (0 or 1).

$$
\frac{\partial v_a}{\partial h_r} = \frac{\partial}{\partial h_r}\sum_{r\in R} h_r \delta^{ar} = \begin{cases} 1 & \textit{if link } a \textit{ is on path } r \\ 0 & \textit{otherwise} \end{cases}
$$

The first item in the objective function is the sum of integrals of the link impedance function. The derivative of this item with respect to the link flow can be written as follows:

$$
\frac{\partial}{\partial v_a}\sum_a \int_0^{v_a} s_a(w)d_w = s_a(v_a) = tt_a = tt_a^f \left[1 + 0.15(\frac{v_a + BG_a}{CP_a})^4\right]
$$

So the first-order condition for this minimization program governing h_r, t_{ij} can now be explicitly expressed as:

$$
\frac{\partial L}{\partial h_r} = \sum_a s_a(v_a)\delta^{ar} - u_{ij} - \theta_r = 0 \qquad 4.3a
$$

$$
\theta_r h_r = 0 \qquad (\theta_r \geq 0) \qquad 4.3b
$$

$$
\frac{\partial L}{\partial t_{ij}} = (\ln t_{ij})/\beta + u_{ij} - \lambda_i - \sigma_j = 0 \qquad 4.3c
$$

The derivative of objective function with respect to the path flow comes to $\sum_a s_a(v_a)\delta^{ar} = \sum_a tt_a\delta^{ar} = u_{ij} + \theta_r$; $\sum_a tt_a\delta^{ar}$ is the congested travel time on path r according the definition. Equations (4.3a-b) indicate that the optimality condition of this program can meet user equilibrium conditions as below:

If $h_r > 0$, then $\theta_r = 0$ the travel time on path r, $\sum_a s_a(v_a)\delta^{ar} = u_{ij}$

If $h_r = 0$, then $\theta_r \geq 0$ the travel time on path r, $\sum_a s_a(v_a)\delta^{ar} \geq u_{ij}$

Since path r belongs to OD pair from i to j, it indicates all used paths (path flow greater than zero) from i to j have the equal travel time of u_{ij}, and all unused path (path flow equal to zero) have higher travel time than u_{ij} or the same travel time as u_{ij}. The minimal travel time between an OD pair is equal to the Lagrange multiplier u_{ij}. These optimal conditions indicate that no road user can improve his/her travel time by switching paths.

Equation 4.3c is used to derive OD trips, which can be transformed into equation 4.4.

$$t_{ij} = e^{\beta(\lambda_i + \sigma_j - u_{ij})} \qquad 4.4$$

This expression is then substituted into constraints 4.1b and 4.1c respectively; the following results are derived:

$$\sum_j e^{\beta(\lambda_i + \sigma_j - u_{ij})} = O_i \rightarrow e^{\beta\lambda_i} = \frac{O_i}{\sum_j e^{\beta(\sigma_j - u_{ij})}} \qquad 4.5a$$

$$\sum_i e^{\beta(\lambda_i + \sigma_j - u_{ij})} = D_j \rightarrow e^{\beta\sigma_j} = \frac{D_j}{\sum_i e^{\beta(\lambda_i - u_{ij})}} \qquad 4.5b$$

To further simplify, set:

$$A_i = \frac{e^{\beta\lambda_i}}{O_i} = \frac{1}{\sum_j e^{\beta(\sigma_j - u_{ij})}} = \frac{1}{\sum_j \frac{D_j e^{\beta(-u_{ij})}}{\sum_i e^{\beta(\lambda_i - u_{ij})}}} = \frac{1}{\sum_j D_j B_j e^{\beta(-u_{ij})}}$$

$$B_j = \frac{e^{\beta\sigma_j}}{D_j} = \frac{1}{\sum_i e^{\beta(\lambda_i - u_{ij})}} = \frac{1}{\sum_i \frac{O_i e^{\beta(-u_{ij})}}{\sum_j e^{\beta(\sigma_j - u_{ij})}}} = \frac{1}{\sum_i O_i A_i e^{-\beta u_{ij}}}$$

Substitute $e^{\beta\lambda_i}$, $e^{\beta\sigma_j}$ into equation 4.4, and the OD trips from i and to j can be expressed as:

$$t_{ij} = A_i B_j O_i D_j e^{-\beta u_{ij}} \qquad 4.6$$

The travel impedance function between OD pairs is characterized by exponential decay form of $f(u_{ij}) = e^{-\beta u_{ij}}$, in which the β is a parameter determined during the calibration procedure. Trip distribution function 4.6 will be used in the calibration procedure to determine β value based on the observed OD trip and travel time tables provided in Chapter 3. After the calibration, the determined β value will be put into the objective function of this combined transportation model for future forecasting.

It is noted that both $A_i = \frac{1}{\sum_j e^{\beta(\sigma_j - u_{ij})}}$ and $B_j = \frac{1}{\sum_i e^{\beta(\lambda_i - u_{ij})}}$ are balancing factors without any physical meaning (Wilson, 1977). Equation 4.6 indicates that the trips between an OD pair from i to j is directly proportional to the number of trips originating from i and the number of trips destined to j, and inversely proportional to travel time between this OD pair.

4.1.3 Model Solution

The convexity and the unique solution with respect to link flow and OD trip of the combined transportation model have been discussed in some studies, where the original

BPR function is in the objective function (Sheffi, 1992; Boyce et al, 1988, Evan, 1976). The purpose of the following section is to prove that the proposed combined trip distribution-assignment model with modified BPR function is still a convex optimization program and has a unique solution with respect to link flow and OD trip. In order to prove this program has one and only one optimal solution, it is sufficient to prove that the feasible region by the constraints is convex and the objective function is strictly convex. The constraints are constituted by linear equality equations; therefore, the feasible region is convex. The non-negative path-flow constraint does not affect the property of convexity, it only needs to prove strict convexity of the objective function. The variables in the objective function include link flow (v_a) and OD trips (t_{ij}); convexity can be proven with respect to these two variables. We can set

$$Z(v_a, t_{ij}) = \sum_a \int_0^{v_a} s_a(w)dw + \frac{1}{\beta}\sum_i \sum_j (t_{ij} \ln t_{ij} - t_{ij}) \qquad 4.7$$

It can be seen that t_{ij} is greater than zero, and that v_a will be greater than zero based on the relationship between path flow and link flow. The strict convexity of the objective function can be proven by demonstrating the Hessian matrix (the matrix of the second derivative of the objective function with respect to these two variables) is positive definite. The derivative of $Z(v_a, t_{ij})$ with respect of link flow on link a and b can be demonstrated as below.

$$\frac{\partial}{\partial v_a} \sum_a \int_0^{v_a} s_a(w)d_w = s_a(v_a) = tt_a \qquad 4.8a$$

$$\frac{\partial^2 Z}{\partial v_a \partial v_b} = \frac{\partial tt_a(v_a)}{\partial v_b} = \begin{cases} \dfrac{\partial tt_b(v_b)}{\partial v_b} & for\ a = b \\ 0 & otherwise \end{cases} \qquad 4.8b$$

$$\frac{\partial tt_b(v_b)}{\partial v_b} = \frac{\partial tt_b^f \left[1 + 0.15(\frac{v_b + BG_b}{CP_b})^4\right]}{\partial v_b} = \frac{0.6tt_b^f}{CP_b}(\frac{v_b + BG_b}{CP_b})^3 > 0 \qquad 4.8c$$

The derivative of $Z(v_a, t_{ij})$ with respect to the OD trips from i to i and the OD trips from r to s can be demonstrated in equation 4.9. Since t_{ij} and β are both greater than 0, $\frac{1}{\beta t_{rs}}$ is greater than 0.

$$\frac{\partial^2 Z}{\partial t_{ij} \partial t_{rs}} = \frac{\frac{1}{\beta} \partial \ln t_{ij}}{\partial t_{rs}} = \begin{cases} \frac{1}{\beta t_{rs}} > 0 & for\ i = r\ and\ j = s \\ 0 & otherwise \end{cases} \qquad 4.9$$

It is noted that $\frac{\partial^2 Z}{\partial v_a \partial t_{ij}} = 0$ and $\frac{\partial^2 Z}{\partial t_{ij} \partial v_a} = 0$ since they are different categories of variables (Sheffi, 1992). Therefore, all the off-diagonal elements in the Hessian matrix are zero. All the diagonal elements in the Hessian matrix can be obtained from equations (4.8a-c and 4.9). So the Hessian matrix can be explicitly written as equation 4.10.

$$\partial^2 Z(v_a, t_{ij}) = \begin{bmatrix} \frac{\partial t_a(v_a)}{\partial v_a} & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & \frac{\partial t_b(v_b)}{\partial v_b} & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & \ddots & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \frac{\partial t_n(v_n)}{\partial v_n} & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \frac{1}{\beta t_{ij}} & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{\beta t_{rs}} & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \ddots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \frac{1}{\beta t_{nn}} \end{bmatrix} \qquad 4.10$$

Since all the diagonal elements in this matrix are strictly great than zero and other elements in the matrix are zero, the Hessian matrix is strictly positive definite. Thus the objective function is strictly convex. Also, the feasible region by the constraints is convex as well. As a consequence, this minimization program has a unique optimal solution with respect to the link flow and OD trip.

The results illustrate that there is only one flow pattern and OD trip pattern associated with solving this minimization problem. It is worth noting that the strict convexity of the objective function is set up with respect to the link flow and OD distribution instead of path flow. The convexity with respect to path flow is not guaranteed, which shows that path flows are not unique, which was discussed in detail in the literature (Sheffi, 1992).

One of the major goals in transportation model development is to apply the model to the study area and forecast the traffic volume on the road network. To achieve this goal, the given parameters including O_i, D_j and β need to be determined based on socioeconomic data and the observed OD trip and travel time tables in the base year. This can be accomplished by model calibration procedure, which is discussed in the following section.

4.2 Model Calibration

After the proposed combined trip distribution-assignment model is formulated, the next step is to calibrate the model to obtain the appropriate value of the parameters. After calibration, the model can be used for future forecasting with available data input such as socio-economic data.

4.2.1 Trip Generation Model Calibration

As discussed in the model formulation, O_i and D_j are given parameters and are the row totals and the column totals in Table 3-2. It can be seen that the row totals are equal to the column totals or $O_i = D_i$ because this transportation model is a daily model in this study. The daily model assumes that the trips that originated from a TAZ eventually come back to this TAZ, which results in $O_i = D_i$. It is worth mentioning that $O_i = D_i$ is not related to the closed/open system at all. It does not matter if the model area is a closed or open system; the daily model always has $O_i = D_i$ with trips that originated from a TAZ eventually coming back to this TAZ. For example, the study area is an open system, which will be discussed in the next chapter; substantial jobs in the study area are occupied by workers outside the study area, and a significant number of workers in the study area are employed outside the study area. However, the trips originating from a TAZ, whether they are going to the study area or outside the study area, will come back to this TAZ at the end of the day since it is a daily transportation model.

Since trips (O_i / D_j) are related to urban activities such as working and shopping, all these activities are generated by household and employment. Let H_i denote the number of households in TAZ i, ($\forall i \in I$); let E_j denote the number of employments in TAZ j, ($\forall j \in J$). O_i and D_j are highly correlated with the number of households and employments. The trip generation calibration attempts to quantify the relationship between O_i / D_j and household and employment, which will be used to estimate the values of O_i / D_j in the future. The scatter plot between trip generation (O_i / D_i) and household/employment (H_i / E_i) is illustrated in Figure 4.1 and Figure 4.2.

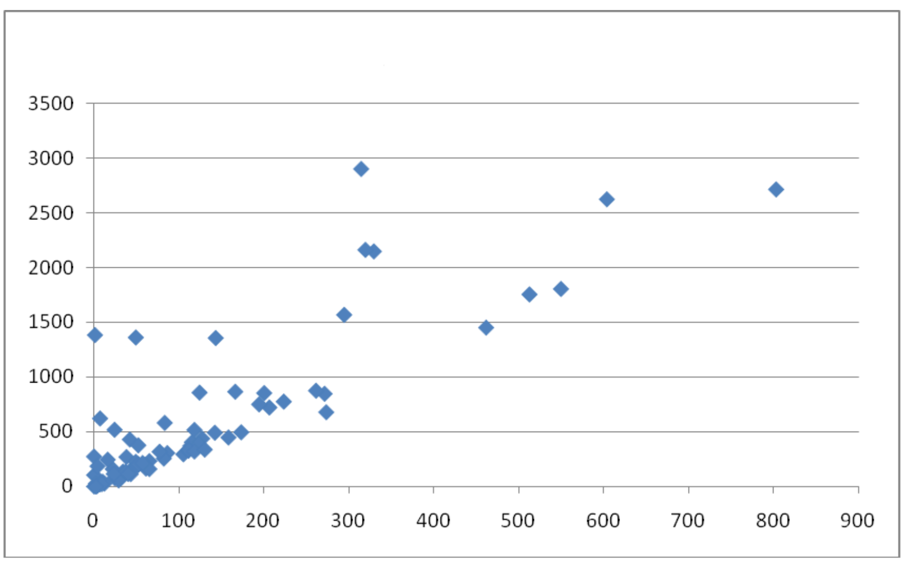

Figure 4.1: Relationship between Trip Generation and Household

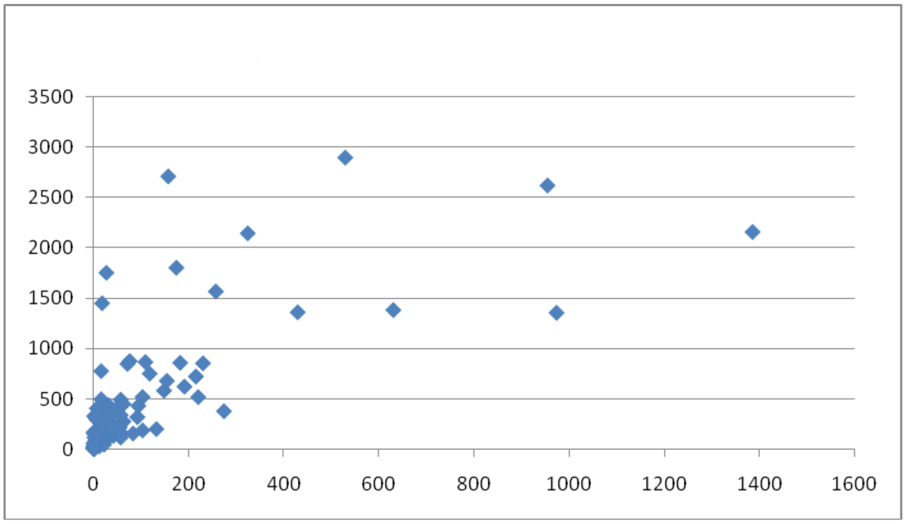

Figure 4.2: Relationship between Trip Generation and Employment

The number of trips generated in each zone in the base year is shown in Table 4.1 along with the number of household and employment, where O_i / D_i are from the observed trip table, and H_i / E_i .are from the census data.

The multiple linear regression technique is used to establish the statistical relationship between trip generation and urban activities in terms of household and employment (Rosner, 2005). The regression equation can be expressed as equation 4.11.

$$\hat{O}_i = \alpha_1 H_i + \alpha_2 E_i \qquad 4.11$$

Where

$\hat{O}_i$: The estimated number of trips originated from zone i

α_1, α_2 : The empirically determined parameters using the observed base year data

It is assumed that the intercept of this equation equals zero since a zone without any household or employment could not generate any trips. For regression development, the regression coefficients of α_1 and α_2 are also known as partial slope coefficients. It indicates the change of response variable of $\hat{O}_i$ corresponding to one-unit change in H_i / E_i respectively, with other explanatory variables fixed. The estimation of the regression coefficients can be achieved by the least square method, which is carried out with statistical software. The least square refers to making the square of difference between the observed value of O_i and the estimated value of $\hat{O}_i$ as small as possible. This estimation can be clearly expressed as the minimization of $\sum_{i=1}^{n}\{O_i - \alpha_1 H_i + \alpha_2 E_i\}^2$; n denotes the total number of TAZs (78).

The regression coefficients fitting the linear model do not guarantee the results are suitable for the purpose of explanation. Whether the regression model is able to significantly explain the variability in the response variable is the first statistical test pursued, which tests if the model is significant at a certain confidence level. The degree of association between the response variable and the explanatory variables is the second statistical test to be pursued, which determines how accurately the model is able to make predictions. It also reflects the degree to which the regression model explains the variability of the response variable. The third statistical test determines whether the regression coefficients are significant at certain confident levels such as 95 percent.

"Total sum of square (Tot SS)" is defined as gross measure of variability of the response variables, which can be decomposed into "regression sum of squares" and "residual sum of squares". "Regression sum of squares (Reg SS)" refers to the variability in the response variable interpreted by the regression model. "Residual sum of squares (Res SS)" represents the variability in the response not accounted for by the model. The relationship among these three square items can be written as:

$$\sum_{i=1}^{n}\{O_i - \overline{O}\}^2 = \sum_{i=1}^{n}\{\hat{O}_i - \overline{O}\}^2 + \sum_{i=1}^{n}\{O_i - \hat{O}_i\}^2 \text{ or}$$

Tot SS= Reg SS + Res SS

It is further defined that Reg MS = Reg SS/k, Res MS = Res SS/(n-k-1) and Tot MS = Tot SS/(n-1). The $\overline{O}$ is the mean of observed value of the response variable computed by $\overline{O} = \frac{\sum_{i=1}^{n} O_i}{n}$. The k is the number of explanatory variables.

Table 4.1: Trip Generation, Household, and Employment in Each Zone

TAZ	H_i	E_i	O_i / D_i	TAZ	H_i	E_i	O_i / D_i
1	7	191	623	40	37	10	129
2	22	58	162	41	3	0	9
3	42	94	431	42	32	7	83
4	34	38	139	43	4	0	12
5	173	16	497	44	37	42	135
6	105	6	297	45	118	220	518
7	49	429	1362	46	123	46	389
8	83	148	583	47	49	20	228
9	113	23	372	48	111	1	329
10	294	257	1569	49	200	230	854
11	65	32	235	50	273	154	680
12	143	973	1357	51	60	16	200
13	61	0	167	52	52	132	201
14	124	182	860	53	46	83	157
15	86	27	307	54	8	8	38
16	0	28	107	55	12	0	29
17	0	63	276	56	549	174	1805
18	0	2	3	57	43	2	116
19	43	1	152	58	52	39	207
20	1	630	1384	59	16	56	247
21	3	0	7	60	7	1	24
22	314	529	2900	61	82	19	257
23	329	324	2148	62	65	11	162
24	271	71	849	63	57	55	217
25	115	7	407	64	206	215	723
26	319	1385	2161	65	111	57	338
27	4	103	188	66	52	274	378
28	2	12	29	67	40	9	118
29	223	16	777	68	29	0	58
30	38	36	271	69	9	0	25
31	10	22	46	70	23	57	117
32	512	27	1755	71	2	0	4
33	603	954	2623	72	118	5	323
34	24	3	80	73	194	118	752
35	24	103	520	74	142	57	493
36	802	157	2713	75	261	76	876
37	166	109	867	76	158	63	450
38	461	18	1452	77	77	92	320
39	127	31	439	78	130	32	339

First, the statistical test for the model significance is performed at 95 percent confidence level by using "f statistic test." The hypothesis can be expressed as $H_0: \ \alpha_1 = \alpha_2 = 0$. If the value of F = Reg MS/Res MS is greater than $f_{k,n-k-1,1-0.05}$ (or equivalently report p-value is less than 0.05 from the statistical software), the H_0 hypothesis is rejected, which determines that the model is significant in estimating the response using the collective explanatory variables.

After verifying the model significance, goodness of fit is examined, which refers to the fraction of the variability in the response accounted for by the regression model. In this multiple linear regression, the adjusted R square is used in judging the degree to which the model can explain the variability in the response variable, which is defined by 1-Res MS/Tot MS.

The hypothesis for the coefficient of each explanatory variable is further tested using "t test," which examines if one explanatory variable is significant in predicting the response after controlling for other explanatory variables.

Let $L_{HH} = \sum_{i=1}^{n} H_i^2 - (\sum_{i=1}^{n} H_i)^2 / n = (n-1)s_X^2$, and $H_0: \ \alpha_1 = 0$. If the value of $t = \frac{\alpha_1}{\sqrt{\text{Res MS}/L_{HH}}}$ is greater than $|t|_{n-k-1,1-0.05/2}$ (or equivalently the compute reported p-value is less than 0.05), the hypothesis test is rejected. It indicated the coefficient for this explanatory variable is significant at 95 percent confidence, which also means this variable is useful in predicting the response.

The model statistics output for this trip generation model calibration is listed in Table 4.2, followed by the regression equation 4.12.

Table 4.2: Trip Generation Regression Model Statistics

Trip Generation Model Statistics			
Variables	Estimate	t-test	p-value
H (Household)	3.1369	18.62	0.000
E (employment)	1.2663	10.17	0.000
Adjusted R Square	0.91		
(Model) F Stat.	467.97		
(Model) p-value.	0.000		

$$O_i = D_i = 3.1369H_i + 1.2663E_i \quad 4.12$$

The primary goal of this regression analysis is to obtain the statistically significant model for estimating the response variable. As seen from the model statistics, 91 percent of (the values of Adjusted R Square) variation in the trip generation can be generally explained by the combination of household and employment. The F statistic shows that the model is quite significant at 95 percent confidence level with p-value substantially lower than 0.05. Thus, the explanatory variables are collectively significant in forecasting. The t statistics show that the coefficient of each explanatory variable is significant at 95 percent confidence level too, which shows that each explanatory variable is significant in the model. Therefore, this regression model is good for the purposes of explanation and estimation. The observed base-year trip generation for each zone are compared with the estimated trips in Figure 4.3, where horizontal axis denotes observed trips and vertical axis represents estimated trips.

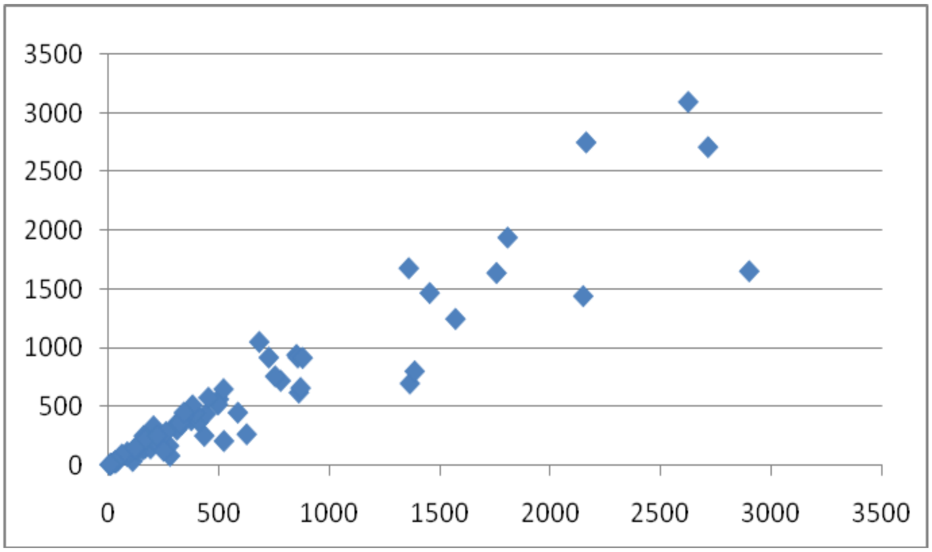

Figure 4.3: Comparison between Observed and Estimated Trip Generation

4.2.2 Trip Distribution Model Calibration

As discussed in section 4.1.2, β that plays an important role in distributing OD trips is empirically determined based on the observed base-year data. The observed OD trip and travel time tables (Table 3.2 and 3.3) are utilized for this model calibration. Trip distribution model calibration refers to the process of determining β value, which ensures that the modeled trip length distribution is as close as possible to the observed trip-length distribution. A trial-and-error (iteration) process is employed in this calibration process.

The calibration process compares the modeled mean travel time with the observed mean travel time in each iteration until these two factors reach a convergence (Caliper Corporation, 2004). Let c_{ij} denotes the observed travel time between OD pair from i to j, ($\forall i \in I$) and ($\forall j \in J$). The observed mean travel time for the study area is defined as $\overline{C_*} = \frac{\sum_{ij} t_{ij} c_{ij}}{T}$. The calibration procedure is constructed as below (Caliper Corporation, 2004).

- Take the inverse of the observed mean travel time as the initial value of β into the travel impedance function; the initial β value can be regarded as β_1.

- Apply the trip distribution model of equation 4.6: $t_{ij} = A_i B_j O_i D_j e^{-\beta u_{ij}}$ associated with the observed O_i and D_j. This will create a new modeled OD trip table.
- Calculate the new modeled mean travel time $\overline{C}_1$ using the same formula as $\overline{C}_*$ based on the new modeled OD trip table. If the convergence ($\frac{\overline{C}_i - \overline{C}_*}{\overline{C}_*} < 1\%$) is reached, the procedure stops; otherwise, update the value of β as follows:
- At (i+1)th iteration, the updated parameter can be calculated according to the following formula $\beta_{i+1} = \frac{(\overline{C}_i - \overline{C}_*)\beta_{i-1} - (\overline{C}_* - \overline{C}_{i-1})\beta_i}{\overline{C}_i - \overline{C}_{i-1}}$. If i=1, $\beta_2 = \frac{\overline{C}_1 \beta_1}{\overline{C}_*}$
- After updating the value of β, return to the next iteration

The calibration procedure stops after eight iterations; the output is the optimal value of β (0.1993), which is larger than zero corresponding to the discussion of β value in model formulation.

4.3 MODEL TEST

The β value and the relationship between O_i / D_i and H_i as well with E_i obtained from the model calibration will be substituted into the combined transportation model. The next step is to sort out the relationship between links and paths for the study area in order to represent the transportation network structure. It is noted that there could be many paths connecting each OD pair in the network, most of which are not reasonable for travelers to choose. Travelers heuristically take only a few of the shortest paths into consideration. For this study, the first three shortest paths between each OD pair are taken into account for formulating the network structure. For each OD pair, the first three shortest paths are chosen based on travel time on the paths. After the relationship between links and paths are obtained, the combined trip distribution and assignment model can be formulated.

$$Min\sum_{a}\int_{0}^{v_a} s_a(w)dw+\frac{1}{0.1993}\sum_{i}\sum_{j}(t_{ij}\ln t_{ij}-t_{ij}) \qquad 4.13a$$

$$S.T.$$

$$\sum_{j} t_{ij} = O_i = 3.1369H_i + 1.2663E_i \qquad 4.13b$$

$$\sum_{i} t_{ij} = D_j = 3.1369H_j + 1.2663E_j \qquad 4.13c$$

$$\sum_{r\in R_{ij}} h_r = t_{ij} \qquad 4.13d$$

$$v_a = \sum_{r\in R} h_r \delta^{ar} \qquad 4.13e$$

$$h_r \geq 0 \qquad 4.13f$$

4.3.1 Algorithm

As previously discussed, this program consists of nonlinear objective functions and linear equality constraints. The strict convexity of the objective function and linear quality constraints ensure that this program has a unique global optimum in terms of link flow and OD trip. This program will be solved by means of an interior point algorithm. A few software packages have been developed to implement this algorithm, such as LOQO, KNITRO, etc. LOQO and KNITRO are utilized to seek the solution of this program.

This algorithm has been developed over the last two decades. It has empirically been shown that this algorithm is efficient and robust in solving large non-linear programming problems (Waltz et al, 2004). The logic behind this algorithm is to change constrained optimization problems into unconstrained problems by placing equality constraints into the objective function with multiplying Lagrange multiplier. Inequality constraints will be placed into the objective function with barrier functions. Barrier functions are designed to prevent the solution from departing the feasible region.

After placing all constraints into the objective function along with Lagrange multipliers and barrier functions, the objective function becomes Lagrangian form for this optimization program. The first order Karush-Kuhn-Tucker (KKT) condition can then be derived for this optimization problem, which results in the standard primal-dual system and consists of a series of equations. Then the algorithm is developed to solve this series of equations to satisfy KKT conditions for optimality. There are many variations in the interior point algorithm associated with computing search step and barrier function. For example, LOQO and KNITRO differ from each other on the method of search step, although both of them use the interior point algorithm. The methodologies in these software packages are briefly introduced.

Vanderbei (1998) incorporated a type of interior point algorithm in the LOQO software package. The logarithm type of barrier function is adopted to remove inequality constraints. After formulating KKT conditions for optimality program, Newton's method associated with feasible search direction is used to solve this program. In the case of the combined transportation model program, all the constraints are linear equality constraints except the non-negativity constraints in terms of path flow. A Lagrangian can be easily formulated with adding the Lagrangian multiplier for the equality constraints and barrier function for the inequality constraints.

Waltz et al (2004) proposed another type of interior point algorithm to solve large-scale nonlinear optimization problems. This algorithm is integrated in the KNITRO software package. It follows the same procedure to formulate the Lagrangian form as LOQO. It also utilizes the same type of barrier functions to eliminate inequality constraints by placing them into the objective function as LOQO. The only difference between these two algorithms lies in the search step in solving the equations of KKT conditions (or primal-dual). KNITRO conducts a line search or a trust region search as its primary step. The line search method is to calculate steps by factoring primal-dual systems of equations; the trust region method is to compute steps by a conjugate gradient iteration.

The Lagrangian form of the combined transportation model programming can be written as below, associated with the interior point algorithm.

$$L = \sum_a \int_0^{v_a} s_a(w)dw + \frac{1}{\beta}\sum_i \sum_j (t_{ij} \ln t_{ij} - t_{ij}) + \sum_i \lambda_i (O_i - \sum_j t_{ij})$$
$$+ \sum_j \sigma_j (D_j - \sum_i t_{ij}) + \sum_i \sum_j u_{ij} (t_{ij} - \sum_{r \in R_{ij}} h_r) + \sum_{r \in R} \theta_r (-h_r + w_r) - \kappa \sum_{r \in R} Ln(w_r)$$

Here $w_r > 0$ is a vector of slack variables for the path set in the system, and κ is the barrier parameter associated with the logarithm type of barrier function.

4.3.2 Model Test

The combined trip distribution-assignment model is tested on a realistic study area of Woodford County, Kentucky. This area has been described in Chapter 3 and consists of 78 zones which produce about 6006 OD pair. There are 723 links and 18,018 paths in this network system.

This optimization program can be solved by both solvers of LOQO and KNITRO, since a unique global optimal solution for this program exists. Both solvers do not have any special requirements about starting point. In KNITRO, any starting point could be selected or KNITRO can compute one (Waltz et al, 2004). LOQO could find a globally optimal solution if the problem is convex; otherwise, it could find a locally optimal solution near to a given starting point (Vanderbei, 1998). However, these two solvers behave differently with respect to the starting point in seeking the optimal solution for the proposed program. Several starting points have been tried on KNITRO and LOQO. It was found that KNITRO always can find the optimal solution no matter what the starting point is. For example, KNITRO was able to converge to the optimal solution even with zero as the starting point. LOQO is more sensitive to the starting point than KNITRO. It often failed to converge to the optimal solution when the starting point was not near to the optimal solution. For example, one starting point was the observed OD trips in Tables 3-2 and the other starting point was 0; LOQO can only find the optimal solution with the starting point of the observed OD trips, and it failed with zero as the starting point. With the same staring point such as the observed OD trips, KNITRO converged to the optimal solution a little slower than LOQO. These two algorithms will continue to be used in Chapter 6 to solve another optimization program.

The optimal solution includes OD trips and link volume. Link volume and Volume/Capacity (V/C) ratio can be used to identify the deficiency along the transportation network, demonstrated in Table 4.3. Since there are too many links over the network, only the first twenty links with the most traffic flow are selected for illustration purposes.

Table 4.3: Selected Link Attributes and Assigned Flow

Link	Flows			V/C
	Background	Assigned	Capacity	
L707	2325.8	4486.8	7000	0.973
L196	2769.5	4461.2	7000	1.033
L679	2903.0	4461.0	7000	1.052
L197	4036.9	3980.4	7000	1.145
L133	11852.9	3681.3	20000	0.777
L356	13024.8	3646.2	20000	0.834
L143	13033.4	3631.7	20000	0.833
L699	174.7	3572.8	7000	0.535
L199	3271.4	3501.3	6000	1.129
L158	11908.7	3481.6	20000	0.770
L515	3254.5	3193.3	6000	1.075
L198	3211.3	3147.1	6000	1.060
L177	2095.8	3025.7	7000	0.732
L209	3141.0	2996.1	6000	1.023
L363	3141.0	2996.1	6000	1.023
L514	3257.9	2933.1	6000	1.032
L178	2078.1	2852.7	7000	0.704
L491	709.9	2846.5	99990	0.036
L70	583.8	2781.1	6000	0.561

Figure 4.4 shows the frequency distribution of the V/C ratio. Results shows that the most traffic is concentrated on the US Highway 60 since this road provides service to a large quantity of through traffic, which is consitent with the observation. Result statistics shows that the study area overall is not congested at all, with an average V/C

ratio of 0.22. Only 3 percent of road segments in the network carry traffic flow over their capacity. Eight percent of the segments have a V/C ratio between 0.6 and 1. Most of the segments carry traffic well under their capacity.

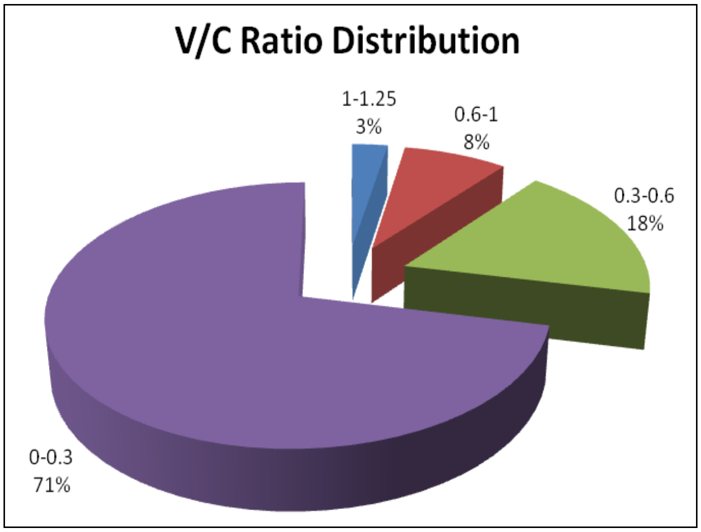

Figure 4.4: V/C Ratio Frequency Distribution

Based on the model outputs, other transportation measures can be developed such as congested travel time between each OD pair and travel time to downtown, which will be used in the land use model development. The next chapter develops the land use model for estimating household and employment distribution.

CHAPTER 5
LAND USE MODEL DEVELOPMENT

The proposed land use model is designed to forecast the number of households/employments that will be located in each planning zone. It is comprised of two sub-models: household and employment distribution models. In this study, land use planning zones have the same spatial configuration as TAZs. This indicates that the land use model and the transportation model are established at the same spatial level. This configuration differs from existing integrated urban models ITLUP and MEPLAN. In those models, a land use planning zone consists of several TAZs. Aggregation of TAZs' information into the land use planning zone resulted in loss of information such as travel time between TAZs. In addition, there are only a few TAZs in this study area; information loss during aggregating TAZs into the land use planning zone is significant for the purpose of analysis.

This chapter first introduces the reasons that existing land use models cannot be applied and the characteristics of the study area. A modeling approach for estimating household/employment distribution at the spatial TAZ level is then discussed. Different categories of factors associated with land use structure and transportation measure are identified from available data sources. Multiple regression equations are developed to find a better combination of these factors for the purposes of explanation and estimation. The model is developed using a relatively straightforward statistical technique, making it easy to understand and develop using the limited resources available in a small urban area.

5.1 MOTIVATION

Regional growth has touched small urban areas (generally defined as communities with a population less than 50,000), resulting in deteriorating traffic conditions. However, MPOs, which coordinate land use planning and transportation planning, are only established in large urban areas with more than 50,000 in population. In recent years, small communities have become more concern about growth. These concerns increased interest in investigating the interrelationship between land use and transportation planning, to encourage smart growth and mitigate traffic congestion. However, for smaller urban areas, existing tools for modeling the interaction between land use and transportation systems originally designed for large urban areas, are not usable.

As discussed in Chapter 2, urban economic models, e.g., MEPLAN, based on input-output frameworks are only suitable for regional or intercity modeling and not for small urban areas. The recently developed urban model UrbanSim, based on discrete choice theory, requires very detailed data at a much disaggregated level such as parcel level. In comparison with UrbanSim, the proposed land use model is easy to use and affordable to implement. For example, UrbanSim has to use GIS to produce a parcel level database, and SAS or SPSS to calibrate the model; the proposed land use model can be easily calibrated in Microsoft Excel. The data inputs for UrbanSim include regional control totals, existing land use, future land use plans, households, employments, environment constraints, development costs, and accessibility. Environment constraints and development costs are not typically available in planning commissions. The proposed land use model does not require these two variables. If UrbanSim was applied to the study area at 200x200 meter parcel level with 10 variables, it requires 123,500 data records (number of variables multiplied by number of parcel cells) for model calibration; the proposed land use model requires only 780 data records (number of variables multiplied by the number of TAZs).

The land use model in the most widely integrated model ITLUP considers the model area as a relatively closed system. It indicates that the majority of jobs in the model area are occupied by workers who live in the model area, and the majority of workers from local households work in the same area. However, in the study area, a significant number of jobs (48% of total jobs) are occupied by workers outside the study area; a significant number of workers from local households (55% of total workers) work outside the study area.

Figure 5.1 shows local residents' spatial commuting pattern based on data from Kentucky State Data Center. It indicates where workers from local households are working, and where workers who are employed in this area live. Only 45 percent of the workers who reside in this study area work in this area. Fifty-five percent work outside the study area: 35 percent work in Fayette County, which is a regional economic center; 11 percent work in Franklin County, where the state capital is located with many state employment opportunities; 9 percent work in other counties or states.

Fifty-two percent of the jobs in this study area are occupied by workers from local households. Forty-eight percent are occupied by workers outside the study area: 18 percent by workers who live in the metropolitan area of Fayette County, 6 percent by workers from Franklin County, and 24 percent by workers from other counties or states.

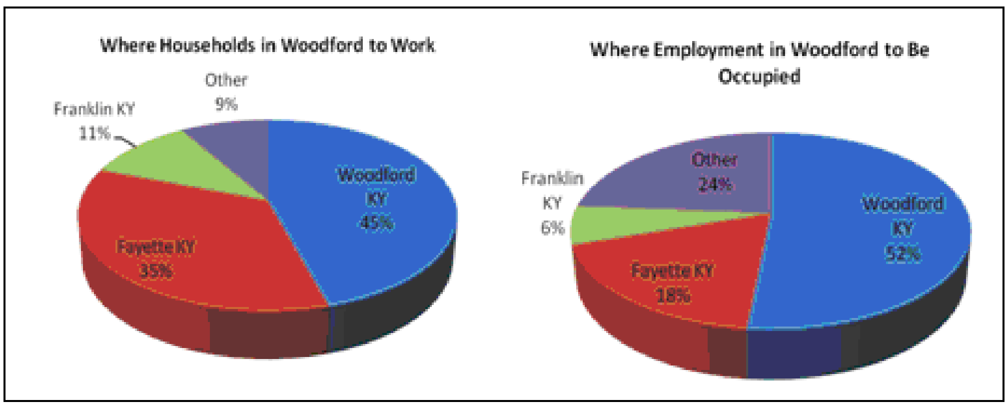

Figure 5.1: Commuting Patterns of the Study Area

It is obvious that the ITLUP is not suitable for this study area because it cannot separate jobs occupied by residents of outside the study area from jobs occupied by Woodford residents. Unrealistic assumptions have to be made when the ITLUP is forced to be applied to this study area. The ITLUP can only focus on the 45 percent of area residents who work inside the study area, and the 52 percent of jobs that are occupied by households from inside the study area. Assumptions made by the ITLUP model were that in each TAZ, 45 percent of area residents work inside the study area, and 52 percent of local jobs are occupied by workers who live inside the study area. A preliminary study was conducted to estimate household distribution using the ITLUP. Results show that the fit between the observed household distribution and the modeled distribution is low, with R square 0.56 even only for those workers employed inside the study area. The household distribution results from the ITLUP model is listed in Appendix 5-A.

It can be seen that existing land use models are not applicable for this small urban area due to their limitation and the area's characteristics. The objective of the proposed land use model is to fit this small urban area's characteristics and take into consideration the limited resources available in small urban areas. The next section describes the process of developing a suitable land use model that can be used in this small urban area as well as in large urban areas.

It is important to note that preliminary analysis shows that household/employment density is a better indicator for household/employment distribution model development than the number of households/employments for. Therefore, the land use model in the following discussion uses the household density model and employment density model.

5.2 VARIABLE SPECIFICATION

Many variables affect household and employment distribution, including demographic and built-environment variables. This study consider two basic categories of variables: land use structure and transportation measures.

5.2.1 Land Use Structure Variables

Land use patterns of Woodford County were analyzed first to examine land-use composition in each zone. In order to eliminate the impact of the zone size, fractions of land use are used instead of land areas of each type of land use. The fraction of a type of land use is defined as the ratio of this land use area over the total area of this TAZ. Eight types of land use are considered in this study: residential, mobile home and multi-family residential, agriculture and preserved agriculture, industrial, commercial, professional office and institutional, vacant, and other land use. In addition, land use mix index (also known as land use balance, or entropy) defined in equation 5.1 is also included. It is a function of land use fractions suggested by Kockelman (1997).

$$MixIndex = -\frac{1}{\ln(n)}\sum_{i}^{n} Fr_i Ln(Fr_i) \qquad 5.1$$

where n is the total number of land use types under consideration, and Fr_i is the fraction of land use type i. The value of mix index ranges from 0 to 1, where 0 infers that there is only one type land use and 1 indicates that all eight types of land use have the same share in a TAZ.

5.2.2 Transportation Measures

Transportation measures describe the transportation system characteristics for each TAZ, as well as its accessibility to different activities. Factors under this category are developed based on the literature of urban economics, urban planning and urban sociology (Waddell et al, 2003a; Waddell et al, 2003b). This literature recognizes many transportation measures, providing inspiration for the development of transportation measures in this dissertation. Transportation measures for this study include distance to major highways, travel time to downtown, travel time to adjacent employment centers, and accessibility measures.

Distance to major highways refers to the length of road segments that connect a TAZ centroid to an on/off ramp or intersection of the nearest major highway. This measure indicates how accessible a TAZ is to US60/US62, Interstate 64, and the Bluegrass Parkway. Travel time to downtown is defined as the congested travel time from a TAZ to the downtown of Versailles, because this downtown has a large number of service and commercial activities. The congested travel time is generated after running the combined trip distribution-assignment model. This measure reflects the degree to which households in a TAZ are able to access service and commercial activities in the downtown area.

There are two major employment centers adjacent to Woodford County, connected through the inter-urban facilities of US 60 and US 62. One is the Lexington metropolitan area; the other is the state capital (Frankfort) with many government jobs. The travel time to adjacent employment center factors include both travel time to Lexington and travel time to Frankfort. These factors describe the accessibility of a TAZ to employment centers.

Accessibility is used to measure the degree to which people in a TAZ reach other activities or a business in a TAZ is reached. Many accessibility measures have been developed for different applications such as access to activity within time threshold, "logsum" accessibility measure, etc. (Miller, 2004). For example, the function of the access to activity within time threshold is to add up all opportunities (i.e., jobs or households) that lie within the travel time threshold (i.e., 15 or 30 minutes). The function of "logsum" measure for a TAZ is to logsum all opportunities as a function of households or employment, and the travel impedance as a function of travel time between this TAZ and other zones.

For the household density model, accessibility measures the degree of ease with which the residents of a TAZ can make trips to other zones in order to accomplish their activities. It is directly proportional to reachable opportunities and inversely proportional to travel time for reaching those opportunities. The combined trip distribution-assignment model in Chapter 4 produced the travel impedance function and the congested travel time between each OD pair, which are used to develop the accessibility function. Based on the concept introduced by Williams (1977), the accessibility function for the household density model is formulated in equation 5.2.

$$ACC_i^h = \sum_j \mathrm{E}_{\mathrm{j}} \times e^{-\beta c_{ij}} \qquad 5.2$$

where

ACC_i^h: The accessibility measure for the household density model

β: The empirically determined coefficient in the combined trip distribution-assignment model as discussed in Chapter 4

For the employment density model, accessibility measures the degree of ease with which a business in a TAZ can attract trips from other zones for gaining benefits (e.g., attracting people to work or shop in this TAZ). This is directly proportional to the number of households at originating zones and inversely proportional to travel time between this TAZ and originating TAZs. Similarly, the accessibility function for the employment density model can be formulated in equation 5.3.

$$ACC_j^e = \sum_i \mathrm{H_i} \times e^{-\beta c_{ij}} \qquad 5.3$$

where

ACC_j^e: The accessibility measure for the employment density model

5.3 CORRELATION ANALYSIS

In this section, the relationship between household/employment density and each factor under the category of land use structure and transportation measure is explored. Land use structure variables include each type of land use fraction ($\mathrm{X_1}$, …, $\mathrm{X_8}$) and land use mix index ($\mathrm{X_9}$). Transportation measures include distance to major highways ($\mathrm{X_{10}}$), travel time to downtown ($\mathrm{X_{11}}$), travel time to Lexington ($\mathrm{X_{12}}$), travel time to Frankfort ($\mathrm{X_{13}}$) and accessibility for the household density model and the employment density model ($\mathrm{X_{14}}$ and $\mathrm{X_{15}}$ respectively). All factors as explanatory variables are summarized in Table 5.1 along with two response variables. Numerical values of these variables are listed in Appendix 5-B. Land use structure data are from the data sources as discussed in Chapter 3; transportation measures are developed based on the output of the combined trip distribution-assignment model.

Table 5.1: Variable Description

Notation	Description
Response Variables	
Y_1	Household density (Households/million ft^2)
Y_2	Employment density (Jobs/million ft^2)
Explanatory Variables	
X_1	Mobile home and multi-family land use fraction
X_2	Residential land use fraction
X_3	Professional office and institutional land use fraction
X_4	Commercial land use fraction
X_5	Industrial land use fraction
X_6	Agriculture and preserved agriculture land use fraction
X_7	Vacant land use fraction
X_8	Other land use fraction
X_9	Land use mix index
X_{10}	Distance to major highways (mile)
X_{11}	Travel time to downtown (min)
X_{12}	Travel time to Lexington (min)
X_{13}	Travel time to Frankfort (min)
X_{14}	Accessibility measure for household density estimation
X_{15}	Accessibility measure for employment density estimation

Correlation analysis is performed to identify the linear association between the response and explanatory variables. Correlation coefficient reflects the direction and strength of the linear relationship between the two variables. The correlation coefficient between Y and X can be computed by the following equation.

$$r_{Y,X} = \frac{s_{Y,X}}{(n-1)s_Y s_X} \tag{5.4}$$

Here $r_{Y,X}$ is the correlation coefficient between these two variables. $s_{Y,X}$ denotes the covariance between these two variables; it can be calculated as $s_{Y,X} = \sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})$ where $\bar{X}$ and $\bar{Y}$ are the means of X and Y respectively.

s_Y , s_X are the sample standard deviations of Y and X respectively: $s_Y = \frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{n-1}$

and $s_X = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$; n is the sample size, which is the total number of TAZs in this study area. The value of the correlation coefficient ranges from -1 to +1; a plus sign represents a positive linear relationship between the two variables; a minus sign indicates that the two variables are negatively correlated. The larger the absolute value, the stronger the linear association that exists between the two variables.

5.3.1 Household Density Correlation Analysis

The scatter plots between household density variables and each explanatory variable are shown in Figure 5.2a-b. They provide visual indication of the relationship (linear, non-linear) between each explanatory variable and the response variable. Correlation analysis is then performed to find major factors that have a significant influence on household density.

Each cell in Figure 5.2a-b shows the relationship between the response variable and the corresponding explanatory variable. The plots apparently demonstrate that there are nonlinear correlations between household density and industry land use fraction, agriculture and preserved agriculture land use fraction, distance to major highway, and travel time to downtown, all of which look like an inversely proportional curvature. Therefore, household density is decreasing with the increase of industrial land use fraction, agriculture and preserved agriculture land use fraction, distance to major highway and travel time to downtown. The relationship between the response and other explanatory variables are difficult to judge based on the plots. Possible transformations such as logarithm, inverse, square, and exponential transformations were tested to find the transformation form that has the strongest correlation with household density. The transformation with the strongest degree of association was chosen for further analysis. Variables of industry land use fraction, agriculture and preserved agriculture land use fraction are not transformed into inverse form since there are zero values for these variables that disable inverse transformation. Ultimate transformation forms for variables of distance to major highway, travel time to downtown, travel time to Lexington, and travel time to Frankfort are inverse form; linear association is the strongest among all transformation forms for other variables.

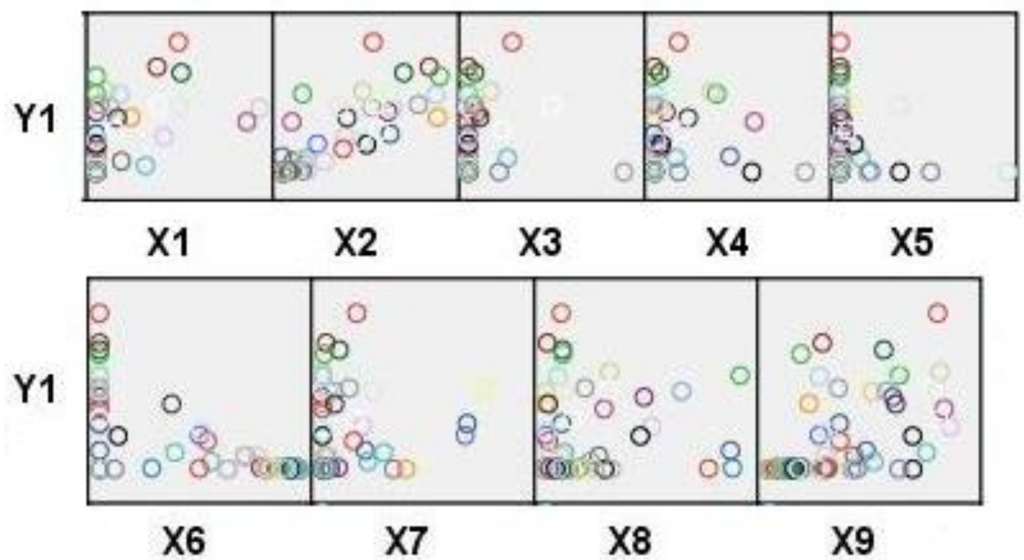

Figure 5.2a: Scatter Plot between Household Density and Land Use Structure Variables

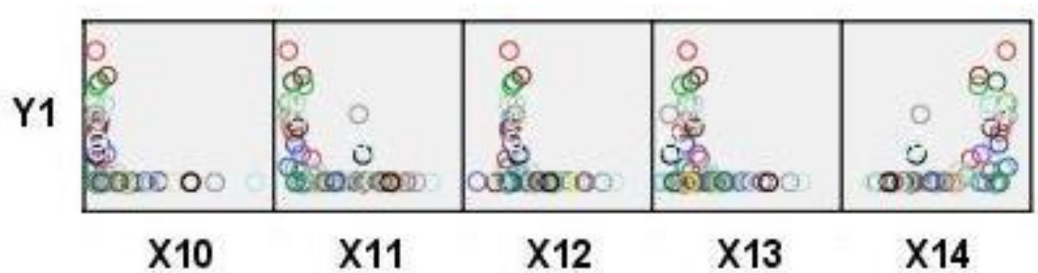

Figure 5.2b: Scatter Plot between Household Density and Transportation Measure

The correlation coefficients between household density and each explanatory variable are presented in Table 5.2. Among all explanatory variables, professional office and institutional land use fraction, commercial land use fraction, industrial land use fraction, vacant land use fraction, other land use fraction, distance to major highway, travel time to Lexington, and travel time to Frankfort are insignificant variables because these variables have correlation coefficient values lower than 0.4, which is used to judge the strength of correlation in this dissertation. Other explanatory variables are significant because they have correlation coefficient values higher than 0.4.

TABLE 5.2 Correlation Coefficients for Household/Employment Density

Correlation Coefficient	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_1	0.59	0.87	0.16	0.16	-0.08	-0.77	0.21
Correlation Coefficient	X_8	X_9	X_{10}^{-1}	X_{11}^{-1}	X_{12}^{-1}	X_{13}^{-1}	X_{14}
Y_1	0.18	0.60	0.34	0.66	0.36	0.34	0.64
Correlation Coefficient	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_2	0.31	0.20	0.10	0.47	0.17	0.48	0.17
Correlation Coefficient	X_8	X_9	X_{10}^{-1}	X_{11}^{-1}	X_{12}^{-1}	X_{13}^{-1}	X_{15}
Y_2	-0.55	0.57	0.70	0.65	0.27	0.23	0.49

Among land use structure variables, the residential land use fraction variable has the most significant relationship with household density, which has a correlation coefficient value of 0.87. It also is the most significant variable among all explanatory variables. This indicates that changing the residential land use fraction will have a significant influence on household density in this study area. The mobile home and multi-family land use fraction variable is also strongly correlated with household density. It can be seen that the agriculture land use fraction has a negative impact on household density; household density decreases with an increase of the agriculture land use fraction. Although land use mix seems to have a positive impact on household density, we need to be cautious with this measure if it appears in the final model. The impact of this measure on household density is complicated under certain circumstances. For example, the correlation coefficient shows that a lower value of land use mix (whose value ranges from 0 to 1) should contribute to a lower value of household density in a TAZ. However, this is not the case when a TAZ has only residential land use. In this circumstance, household density has a higher value although the value of land use mix is the lowest zero. Therefore, land use mix can only be applied to those TAZs that have both residential land use and other types of land use. Among transportation measure variables, travel time to downtown and accessibility show a strong correlation. Correlation coefficients show that households tend to live in more accessible zones; the zones closer to the downtown area have greater household density.

5.3.2 Employment Density Correlation Analysis

The employment density correlation analysis follows the same procedure as household density. The scatter plot between employment density and each category of variables is drawn and shown in Figure 5.3a-b. Analysis was performed to identify significant factors that are strongly correlated with employment density.

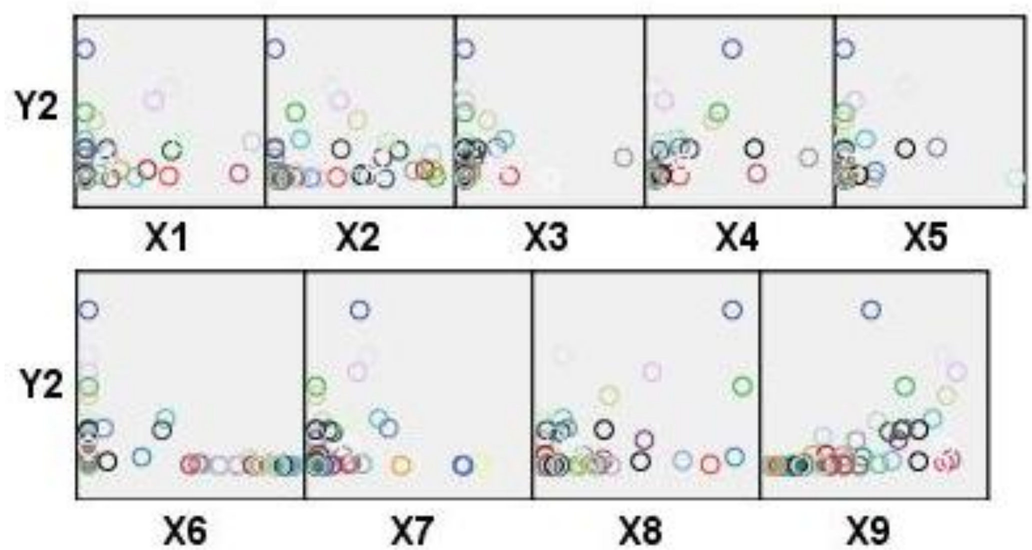

Figure 5.3a: Scatter Plot between Employment Density and Land Use Structure Variables

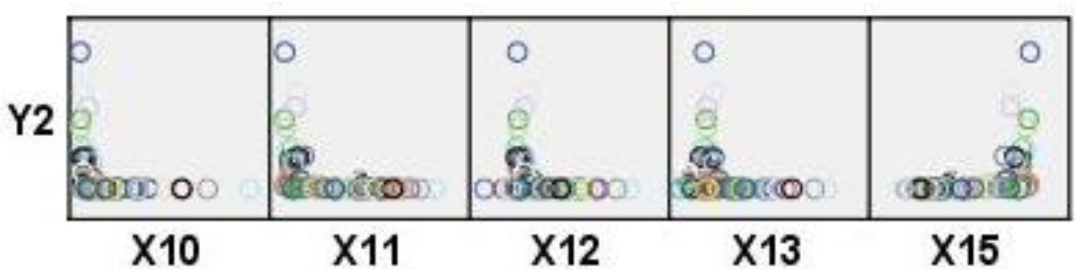

Figure 5.3b: Scatter Plot between Employment Density and Transportation Measures

Each cell in Figure 5.3a-b shows the relationship between employment density and the corresponding explanatory variable. It is observed that there is a nonlinear relationship between employment density and mobile home and multi-family land use fraction, residential land use fraction, agriculture and preserved agriculture land use fraction, distance to major highway, and travel time to downtown, all of which appear to have a inversely proportional curvature. This signifies that employment density decreases with the increase of mobile home and multi-family land use fraction, residential land use fraction, agriculture and preserved agriculture land use fraction, distance to major highway, and travel time to downtown. The relationship between employment density and other explanatory variables is unclear based on their plots. Possible transformations such as logarithm, inversion, power, and exponential are tested to find the desirable transformation form. Variables of mobile home and multi-family land use fraction, residential land use fraction, and agriculture and preserved agriculture land use fraction are not transformed into inverse form since there are many zero values for these variables, making inverse transformation impossible. The ultimate transformation forms for variables of distance to major highway, travel time to downtown, travel time to Lexington, and travel time to Frankfort are inverse form, which are utilized in further analysis.

The correlation coefficients between employment density and each explanatory variable are displayed in Table 5.2. Half of the explanatory variables are insignificantly correlated with the employment density variables since correlation coefficient values are lower than 0.4. These variables include mobile home and multi-family land use fraction, residential land use fraction, professional office and institutional land use fraction, industrial land use fraction, vacant land use fraction, travel time to Lexington, and travel time to Frankfort.

Among land use structure variables, the commercial land use fraction is strongly correlated with employment density. The correlation coefficient between the agriculture land use fraction and employment density shows that it exerts a negative impact on employment density. Land use mix index also has a significant influence on employment density, which shows that employment density rises with the increase of land use mix. Attentions should be paid if it appears in the final model, because the impact of this measure on employment density is complicated under certain circumstances. For example, the correlation coefficient shows that a lower value of land use mix (whose value ranges from 0 to 1) will result in a lower value of employment density. However, this is not the case when a TAZ has only commercial/service/industry land use. In this circumstance, employment density has a higher value although the value of land use mix is the lowest zero. Therefore, land use mix can only be applied to those TAZs that have both commercial/service/industry land use and other types of land use. Among transportation measure variables, distance to major highway has the strongest correlation with employment density, with a correlation coefficient value of 0.70. It also is the most significant factor in all explanatory variables. Employment density tends to decline with an increase in distance to major highway, which suggests that employment has an inclination to locate along major highways. Travel time to downtown and accessibility are strongly correlated with employment density, too. Employment density seems to decrease with the increase of travel time to downtown and increase with the increase of accessibility. It indicates that jobs tend to be located around the downtown area and in those TAZs with higher accessibility.

5.4 REGRESSION MODELING ANALYSIS

The explanatory variables that are strongly correlated with the response variables will be considered in model development. These explanatory variables are candidate variables for multiple linear regression model development. A model with all these explanatory variables included may not be the best based on statistical tests. Statistical techniques are used to find an optimal subset of candidate variables. These techniques include forward selection, backward elimination, and stepwise selection (Rosner, 2005).

5.4.1 Statistical Techniques

Forward selection procedure begins with only an intercept in the model. First, a single explanatory variable model is identified from all possible one-variable models which yield the smallest Res SS as defined in Chapter 4. In this study, the significance level is 0.05, so p-value for this variable must be less than 0.05. If p-value for the first selected value is less than 0.05, this variable is included in the model; otherwise, none of the candidate variables should be included in the model. The next step is to fit all possible two-variable models with the first variable in the model. The second variable is identified, which yields the largest further reduction in Res SS. The p-value for the second variable is also required to be less than the significance level. This procedure continues until a large p-value over 0.05 is obtained or, much less commonly, until all explanatory variables are included in the model.

Backward elimination procedure starts by including all explanatory variables in the model. A single variable is then identified whose removal will cause the smallest increase of Res SS. This variable will be removed if its p-value is larger than 0.05; otherwise, it suggests that all candidate variables should remain in the model. If the first selected variable is removed, the procedure goes on to identify the second explanatory variable whose removal will lead to the smallest increase in Res SS. This process continues until the p-value of each variable in the model is smaller than 0.05, which indicates that none of the candidate explanatory variables should be removed from the model.

Stepwise selection procedure is the hybrid of forward selection and backward elimination. It starts out the same way as forward selection. Each time a new explanatory variable is included in the model based on the forward selection method, backward elimination will be conducted to test if any of the previously added explanatory variables can be removed from the model.

During regression model development, collinearity may appear among explanatory variables. Strong correlation between explanatory variables produces the collinearity. For example, there is a strong correlation between travel time to downtown and accessibility during household density model development; their correlation coefficient is 0.79. Collinearity will have a negative impact in estimating the coefficient of explanatory variables. An appropriate combination of explanatory variables will be chosen to rule out collinearity. Collinearity will be diagnosed using statistics VIF (Variance of Inflation Factor) for each explanatory variable. For example, VIF for explanatory variable X_1 is defined as $1/(1-R_1^2)$. R_1^2 is calculated from the auxiliary regression model, where X_1 will be treated as the response variable and other explanatory variables still remain as explanatory variables. There is no specific threshold value for VIF to determine collinearity. A rule of thumb is to pay attention to collinearity when VIF for an explanatory variable is greater than 10 (Rosner, 2005). For example, VIF of X_1 is larger than 10, which indicates that this explanatory variable can be removed from the model since it is approximately a linear combination of other explanatory variables.

Forward, backward and stepwise selection procedures are implemented in the SPSS software package. These three procedures generate three candidate regression models. The adjusted R square and VIF will be used to judge these competing candidate models.

5.4.2 Household Density Model Development

Household density model development is to find the appropriate combination of explanatory variables that can be used to estimate household density for each TAZ. During the model development, the aggregation of similar explanatory variables is conducted to seek better estimation. For example, residential land use fraction is added into mobile home and multi-family land use fraction to create total residential land use fraction since they belong to the same category of land use. It shows that total land use fraction can provide a better estimation for household density with bigger adjusted R square.

In order for the regression model to be more physically meaningful, the constant is excluded from the model. Unreasonable estimation can take place if the constant stays in the model. For example, some households are estimated for a TAZ because of the constant even if this TAZ has no residential land use. After evaluating three candidate models from forward, backward, and stepwise selection procedures, the following linear equation out of the stepwise selection procedure is chosen because it has the best fit for the data. The statistical outputs of the household density model are listed in Table 5.3.

$$\text{Household Density} = 64.762 \times (\text{total residential land use fraction}) + 0.224 \times (\text{accessibility})$$

Table 5.3: Household Density Model Statistics

Variables	Estimate	t-test	Sig.(p-Value)
Total residential land use fraction	64.762	13.267	0
Accessibility	0.224	3.052	0.003
Adjusted R Square	0.87		
F Stat.	264.144		
Sig.	0		

Results show that 87 percent of variation (Adjusted R square 0.87) in household density can be explained by the combination of total residential land use fraction and accessibility measure. All statistics for this household density model are significant at 95 percent confidence level. The f statistic for the regression model significance shows that the explanatory variables are collectively useful in predicting household density. The t statistic for the coefficient of each explanatory variable shows that each variable is useful in forecasting the response after controlling other explanatory variables.

As expected, total residential land use fraction has a strong and positive impact on household density; household density tends to increase with its increase. Accessibility also plays an important role in estimating household density; households tend to live in zones with higher accessibility. We need to be cautious when applying this model since each of these two explanatory variables has its own physical meaning in the urban planning environment. For example, accessibility measure always has a value larger than zero, which can create a situation where the estimated value of household density for a zone is always larger than zero even when there is no residential land use at all in this zone. This is not consistent with reality in urban planning. Therefore, this model is only applied into these zones whose total residential land use fraction is larger than zero. For each of those zones with no residential land use, household density is set to zero. In summary, this model is able to assess the role of transportation decision and land use policy in household distribution since it includes both transportation and land use explanatory variables. The comparison between the observed household density and the estimated household density is shown in Figure 5.4.

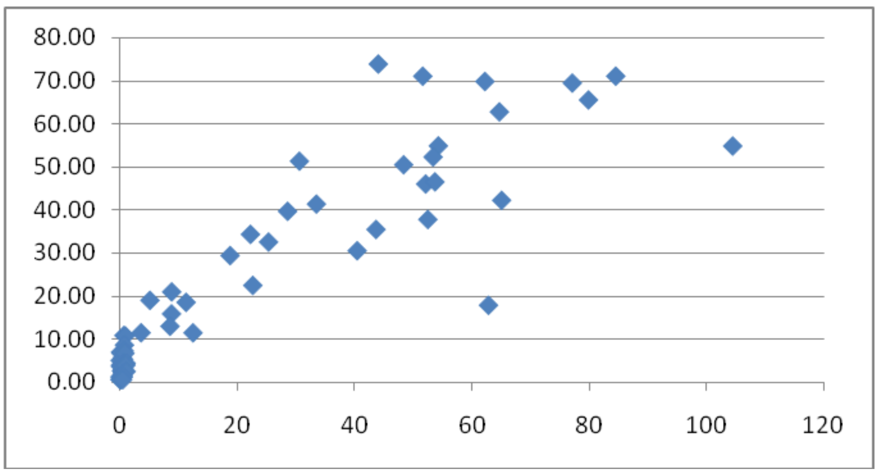

Figure 5.4: Comparison between Observed and Estimated Household Density

5.4.3 Employment Density Model Development

The employment density model is developed by following a similar procedure similar to the household density model. The aggregation of similar explanatory variables is examined during model development. For example, professional office and institutional land use fraction, commercial land use fraction, and industrial land use fraction are aggregated together to form the employment land use fraction variable.

Similarly, the constant is excluded from the model. After evaluating the candidate models from forward, backward, and stepwise selection procedures, the following linear equation out of stepwise is chosen to represent the employment density model. The statistical outputs of the employment density model are listed in Table 5.4.

$$\begin{aligned}\text{Employment Density} = {} & 9.411/(\text{distance to major highway}) \\ & + 58.724 \times (\text{land use mix index})\end{aligned}$$

Table 5.4: Employment Density Model Statistics

Variables	Estimate	t-test	Sig.
Distance to major highway	9.411	6.276	0
Land use mix	58.742	3.185	0.002
Adjusted R Square	0.606		
F Stat.	63.098		
Sig.	0		

The model output shows that 60.6 percent of variation in employment density can be explained by this regression model. F statistic for the regression model is significant at 95 percent confidence level; and t statistics for the coefficients of both explanatory variables are significant at 95 percent confidence level. The model shows that the zones closer to major highways tend to have denser employment than those farther from major highways.

It is important to note that employment distribution typically has a historical trend, or it is more exogenous than endogenous in the context of urban planning. For example, the EMPAL model discussed in Chapter 2 allocates employments into each planning zone based on previous employment distribution. This historical relationship between current employment distribution (2000) and previous employment distribution (1995) is further explored. It is found that current employment distribution is closely related to previous employment distribution (1995). The employment distribution model with previous employment distribution as the explanatory shows an extraordinary high adjusted R square of 0.994. Therefore, the following linear equation is selected to be the final employment density model in this study. The statistical outputs of the employment density model are listed in Table 5.5. All the statistics for both the model and the explanatory variables are significant at 95 percent confidence level.

$$\text{Employment Density} = 1.0926(\text{Previous employment density})$$

Table 5.5: Employment Density History Model Statistics

Variables	Estimate	t-test	Sig.
Previous employment density	1.0926	114.5	0
R Square	0.994		
F Stat.	13122		
Sig.	0		

The comparison between the observed and the estimated employment density is shown in Figure 5.5.

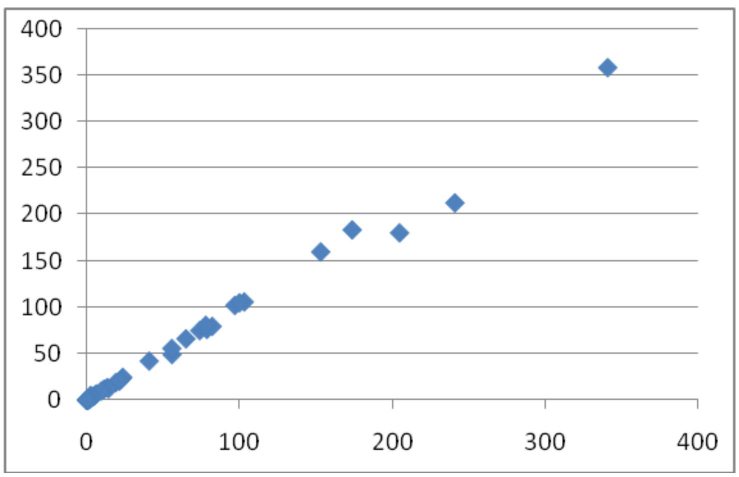

Figure 5.5: Comparison between Observed and Estimated Employment Density

The household and employment density models are able to estimate the number of households and employments in each zone (TAZ) by multiplying zone area, which is the critical input to the proposed combined trip distribution and assignment model as discussed in Chapter 4. On the other hand, the output of the combined trip distribution and assignment model produce congested travel time on the network, which is an important input to household density model. The interaction between these two models will be discussed in detail in the next chapter.

CHAPTER 6
THE INTERACTION

6.1 OVERVIEW

After the combined transportation model and the land use model have been developed and calibrated as discussed in the last two chapters, the third component of the proposed integrated model framework (interaction between the land use model and the transportation model) is examined in this chapter. As discussed in the last two chapters, the land use model output serves as the input to the transportation model, and vice versa. This generates the interaction between these two models. In this study, the interaction is designed to achieve consistency between the land use model outputs and the transportation model outputs. Consistency is achieved when the transportation model outputs put into the land use model are able to produce the same land use model outputs as those initially put into the transportation model, and vice versa. In the existing ITLUP framework, a feedback loop between the DRAM/EMAPL model and the traditional four-step transportation model is formulated to reach this consistency as discussed in Chapter 2.

In this dissertation, two different procedures are developed to achieve consistency between the land use model outputs and the transportation model outputs. One is feedback model configuration, which solves the land use model and the transportation model iteratively. The other is simultaneous model configuration, which formulates the land use model and the transportation models into one optimization problem. In both configurations, use equilibrium conditions are satisfied.

6.2 FEEDBACK MODEL CONFIGURATION

Most existing operational integrated urban models take "lagged transportation model output in terms of travel time" as the input for land use models to estimate household/employment distribution. This does not take into consideration the consistency between land use model output and transportation model output. Only ITLUP uses a feedback loop to iteratively solve its land use model and transportation model to reach consistency. In ITLUP, the transportation model is a traditional four-step travel demand model, and land use models are DRAM/EMPAL.

Feedback model configuration in this dissertation can be briefly described as follows: After solving the transportation model, the transportation model output such as travel time is input to the household distribution model. The household distribution model outputs subsequently become inputs to the transportation model to re-estimate the travel time, which will be put back into the land use model again to re-estimate the household distribution. This feedback loop continues until pre-defined convergence is reached.

6.2.1 Formulation

The developments of the combined trip distribution-assignment model and the regression-based land use model have been discussed in Chapter 4 and 5 respectively. This section describes how to formulate the interaction between them.

The combined trip distribution-assignment transportation model has been developed as illustrated in Equation: 4.13a-f. The model outputs can be used to generate a variety of transportation performance measures for decision-makers to evaluate the transportation system. In addition, one transportation measure is congested travel time between TAZs (c_{ij}), which is a key input to the household distribution model in this study.

In the land use model, the final employment distribution model as described in Table 5.5 does not include any transportation measures or land use structure variables, which will be regarded as exogenously determined in this study. The household distribution model has been developed as a function of total residential land use fraction and accessibility measure as described in Table 5.3. The number of households in each TAZ can be estimated based on this household distribution model with the following equation:

$$H_i = Area_i \times [64.762 \times (X_{1i} + X_{2i}) + 0.224 \times \sum_j \mathrm{E_j} \times e^{-\beta c_{ij}}] \qquad 6.1$$

where $Area_i$ is the area of TAZ i for $i \in I$; it is usually a constant during the urban planning process. The number of households in each TAZ (H_i) is a key input to the transportation model.

A feedback loop is formed between the land use model and the transportation model since one model's output serves as input to the other model. The feedback loop model configuration is illustrated in Figure 6.1. An iterative approach is established between the land use and transportation models in feedback model configuration. The approach involves solving the land use model and the transportation model iteratively based on each other's outputs in the previous round. The household distribution model (equation 6.1) is solved to obtain the number of households in each TAZ (H_i); then, H_i is substituted into the transportation model (equations 4.13a-f). Congested travel time c_{ij} is obtained after solving the transportation model. c_{ij} is then put into the household distribution model to re-estimate H_i; re-estimated H_i is put back into the transportation model again to re-estimate c_{ij}. This feedback loop continues until predefined convergence criteria are reached.

In this study, the predefined convergence criteria are specified as: in two consecutive iterations, (1) less than 5 percent of OD pairs have OD trip variation more than 5 percent; (2) less than 5 percent of links have link flow variation more than 5 percent; and (3) less than 1 percent of zones have household distribution variation more than 5 percent. For example, the variation with respect to OD trip variable at k-th iteration is calculated as the equation: $\left|\frac{t_{ij}^k - t_{ij}^{k-1}}{t_{ij}^{k-1}}\right| \times 100$. All three conditions have to be satisfied before the procedure stops. The consistency between the land use model outputs and the transportation model outputs is regarded as being achieved once the predefined convergence criteria are met.

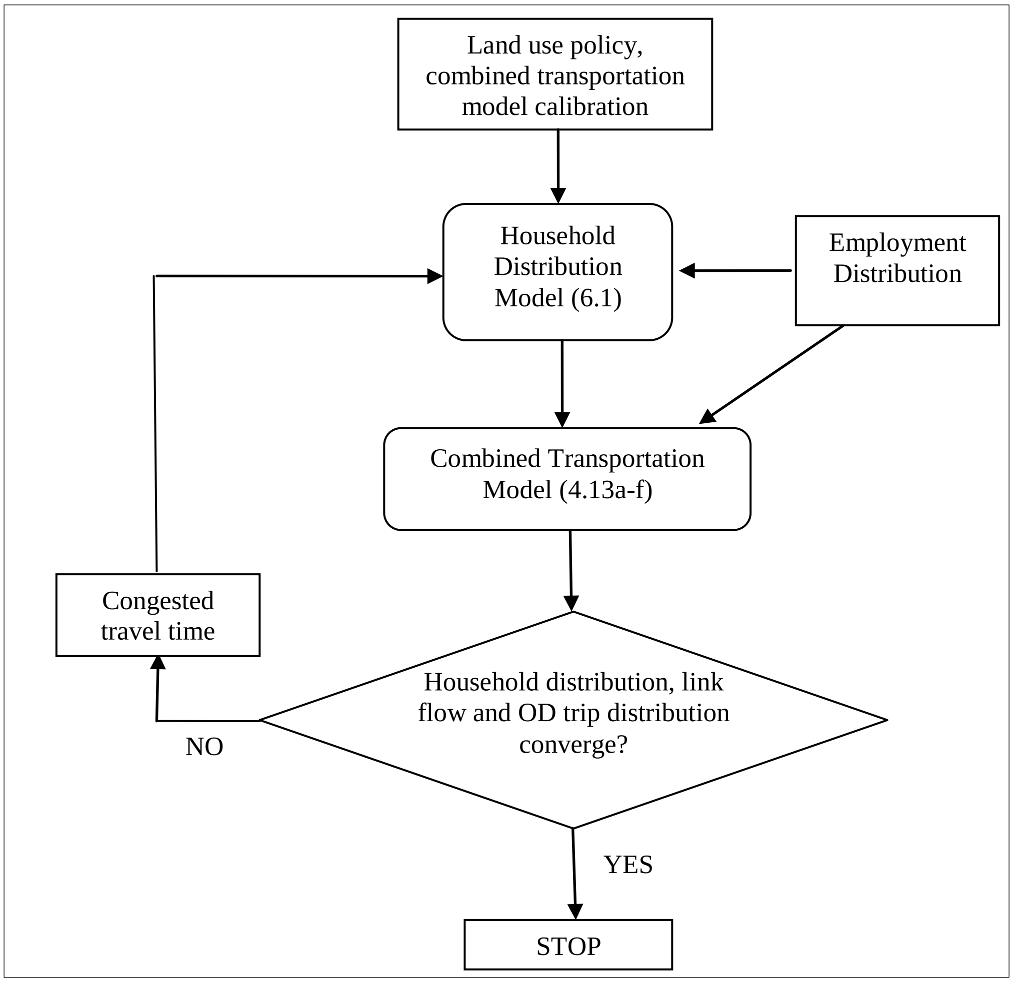

Figure 6.1: Feedback Model Configuration

6.2.2 Numerical Example

Household distribution model equations (6.1) can be solved when input variables are ready. The combined trip distribution-assignment model (4.13a-f) can be solved using either KNITRO or LOQO as discussed in Chapter 4. The specific steps involved in solving the feedback model are outlined below.

- Determine E_i by solving the employment distribution model
- Set initial values c_{ij}^{k-1} which can be obtained from the base year transportation model; and set k=1 where k means k-th iteration
- Solve the household distribution model to obtain H_i^k
- Solve the combined trip distribution and assignment model to attain v_a^k, t_{ij}^k, c_{ij}^k
- Check the convergence criteria with respect to H_i^k , v_a^k , t_{ij}^k ; if they are satisfied, stop the iteration; otherwise, substitute c_{ij}^k into step 2 to replace c_{ij}^{k-1}, and repeat the procedure

The household distribution model and the combined trip distribution-assignment model are resolved iteratively. KNITRO or LOQO are used to solve the combined trip distribution-assignment model. Both solvers are able to successfully find the optimal solutions for this transportation model. KNITRO always can find the optimal solution for the combined transportation model no matter where the starting point is; however, LOQO can find the optimal solution only if the starting point is close to the optimal solution, such as using the base year data as the starting point.

The base year data associated with the socio-economic and transportation network is plugged into this integrated model to test the model performance. Although the convergence between the household distribution model and the combined transportation model is not guaranteed in theory, this feedback model converges quickly for the study area. The convergence is attained after three iterations between the land use model and transportation model. During these three iterations, the convergence is reached according to the pre-defined convergence criteria. For the links with big variations on link flows between the first and second iterations, the variations were becoming smaller in the next iteration. For the links with smaller variation than the convergence criteria between the first and second iterations, the variation changed very little in the next iteration and still remained under the convergence criteria. There is no major fluctuation on the variations as to link flows, OD trips, and household distribution between these three iterations.

After three iterations, pre-defined convergence criteria have been entirely satisfied for the base year scenario. Results with respect to link flow, OD trip, and household distribution in the last two iterations are reported. Since there are too many OD pairs and links, only the first twenty links with the most traffic flow and the first twenty OD pairs with the highest OD trips are chosen to be displayed in Table 6.1. The results show that no link has link flow variation of more than 5 percent in the last two iterations, no OD pair has trip distribution variation greater than 5 percent, and only 26 OD pairs out of 6006 pairs (0.4 percent) have trip distribution variation of more than 1 percent. With respect to the household distribution variable, none of the TAZs has the number of household variation more than 5 percent. It can be seen that consistency between the land use model outputs and the transportation model outputs is achieved after the pre-defined convergence criteria are met. However, consistency is not achieved between the first and the second iterations because the predefined convergence criteria are not reached. For example, between the first and the second iterations, 6.9 percent of links have flow variation of more than 5 percent, which does not satisfy the pre-defined convergence criteria; 10.2 percent of OD pairs have trip distribution variation of more than 5 percent, which does not meet the convergence criteria; 6.4 percent of TAZs have the number of household variation of more than 5 percent, which falls out of the convergence criteria. In addition, this consistency indicates that a dynamic equilibrium between land use and transportation systems is reached. Households are satisfied with their locations and travel time to work and other activities after iteratively household distribution. Transportation systems can accommodate travel demand and serve travel needs at an acceptable level.

The model outputs such as link volume can be used to help decision-makers evaluate overall performance of the existing road network along with link capacity. V/C ratio is a proper index to measure the degree to which network capacity is able to accommodate travel demand. Results show that the test area has a quite low level of congestion, with average V/C ratio of 0.32.

Table 6.1: Selected Feedback Model Output

Link	Link Flow (third iteration)	Link Flow (second iteration)	Change between two iterations(percent)	Origination	Destination	OD Trips	OD Trips	Change between two iterations (percent)	Zone	Household (third iteration)	Household (second iteration)	Change between two iterations (percent)
L196	6856	6855	0.0	Z74	Z75	1431	1432	0.1	56	1635	1635	0.0
L679	6823	6828	0.1	Z75	Z74	1431	1432	0.1	73	1335	1328	0.6
L356	6559	6557	0.0	Z72	Z73	850	845	0.6	72	1257	1255	0.2
L143	6354	6352	0.0	Z73	Z72	850	845	0.6	36	1180	1179	0.0
L133	6252	6240	0.2	Z36	Z56	587	587	0.0	75	1163	1163	0.0
L707	6194	6193	0.0	Z56	Z36	587	587	0.0	74	1045	1045	0.0
L110	6044	6031	0.2	Z65	Z64	586	586	0.0	38	1027	1026	0.0
L111	6044	6031	0.2	Z64	Z65	586	586	0.0	33	839	836	0.3
L158	5900	5890	0.2	Z76	Z78	566	566	0.1	64	809	809	0.0
L633	5806	5792	0.2	Z78	Z76	566	566	0.1	62	733	733	0.0
L634	5806	5792	0.2	Z64	Z62	502	503	0.2	58	579	576	0.4
L105	5791	5778	0.2	Z62	Z64	502	503	0.2	37	578	578	0.0
L106	5791	5778	0.2	Z75	Z76	365	365	0.1	50	575	573	0.4
L107	5791	5778	0.2	Z76	Z75	365	365	0.0	39	572	571	0.2
L108	5791	5778	0.2	Z56	Z72	356	356	0.0	51	472	469	0.6
L109	5791	5778	0.2	Z72	Z56	356	356	0.0	76	459	458	0.2
L632	5477	5462	0.3	Z73	Z56	339	347	2.4	61	453	452	0.1
L409	5471	5456	0.3	Z56	Z73	339	347	2.4	32	435	435	0.0
L403	5350	5348	0.0	Z37	Z36	330	330	0.0	30	384	384	0.1
L636	5350	5348	0.0	Z36	Z37	330	330	0.0	29	377	377	0.1

Average V/C ratio is biased downward because it is observed that a large proportion of links are lightly traveled. Therefore, frequency distribution of V/C ratio is presented in Figure 6.2 to reflect the level of network congestion. V/C ratio statistics show that there are no extremely congested links on the network. Only 5 percent of links have V/C ratios larger than 1, which indicates that 95 percent of links carry traffic under their capacities. Fifty-nine percent of links are lightly loaded by showing a V/C ratio of less than 0.3; 21 percent of links have a V/C ratio between 0.3 and 0.6; 15 percent of links have a V/C ratio between 0.6 and 1.

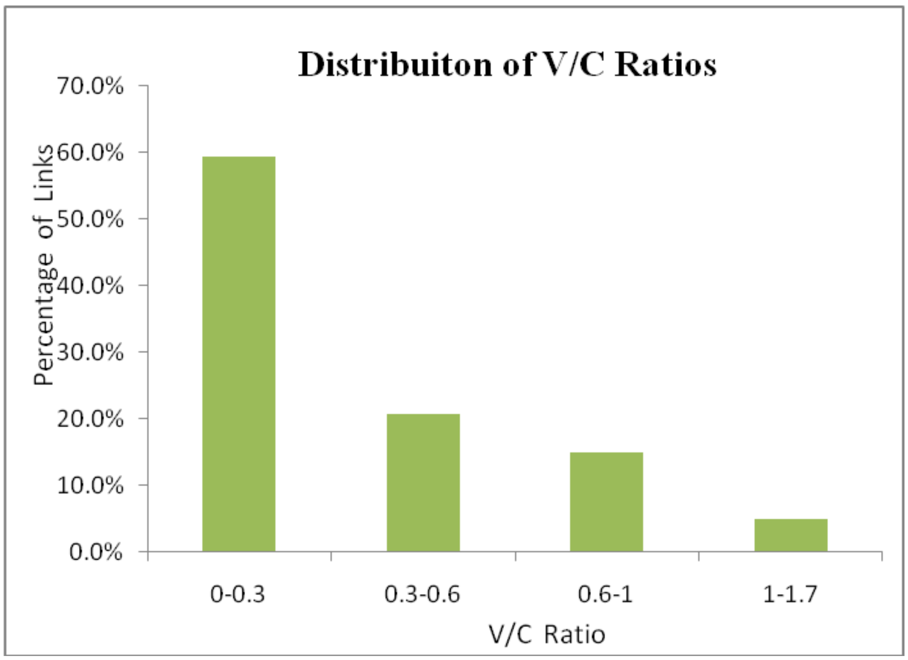

Figure 6.2: Frequency Distribution of V/C Ratio from Feedback Model

6.2.3 Future Model Application

This section is to test the feedback integrated model performance on future scenarios. One potential residential land use development is evaluated using this feedback model framework. One major development decision is being made to TAZ 36: a subdivision development will change the residential land use fraction from 36.4 percent to 46.4 percent. This TAZ is near the board of Versailles along major road US60. The location of TAZ 36 is displayed in Figure 6.3.

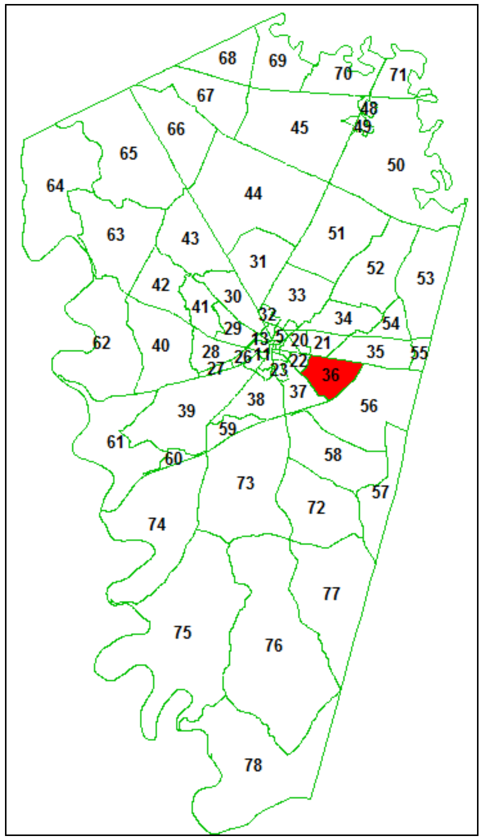

Figure 6.3: Development Area

Residential land use fraction (46.4 percent) is substituted into the household distribution model along with the base year travel-time matrix to preliminarily estimate future household distribution. The estimated future household distribution is subsequently put into the transportation model to forecast the future travel-time matrix. The estimated future travel-time matrix is again substituted into the household distribution model to re-estimate future household distribution. This feedback loop between the land use model and the transportation model continues until the specified convergence criteria are reached.

The convergence is reached after three iterations. The results show that in the future, the number of households in TAZ 36 will increase, and the number of households in neighborhood TAZ 56 will decrease. The number of households in TAZ 36 will increase from 1,180 to 1,440 due to the augmentation of residential land use area. Additional trips generated by the newly added 260 households will use surrounding road networks, which leads to the decrease of accessibility in neighborhood TAZ 56. Therefore, the number of households in neighborhood TAZ 56 will be reduced from 1,654 to 1,653 as a result of lower accessibility in comparison to the base year. The reason there is no significant household change in TAZ56 is because of a minor change in accessibility measure in this TAZ. This minor change in accessibility is associated with an insignificant difference between future traffic flow and base year traffic flow over the network. Overall, the new 260 households do not have a significant impact on the congestion level of the network because it is not congested either in the base year or in the future. For example, the road segment of US 60 along the development area has a daily traffic volume of 16,779, and the V/C ratio is 0.76 with its daily capacity of 22,000 in the base year. With the newly added households in TAZ 36, daily traffic flow on this segment will be 16,832 and the V/C ratio will be 0.77.

The traffic flow on each link along with the V/C ratio for the future scenario is estimated. The comparison on V/C ratio frequency distribution between the base year and the future is illustrated in Figure 6.4. It can be seen that the new development in TAZ 36 does not have a major impact on traffic flow distribution compared with base year. The comparison shows that the future year average V/C ratio increases slightly to 0.33 from 0.32 in the base year. Therefore, there is no significant difference in congestion level between the base year and the future. The comparison also shows that the percent of congested links increase from 5 percent in the base year to 5.3 percent in the future. This indicates that the newly added 260 households in TAZ 36 generate a bit more congested links in the future because these new households travel to other TAZs to achieve their needs. Additionally, the comparison shows that the percent of the links whose V/C ratio is between 0 and 0.3 increases from the base year 59.3 to the future 60. This implies that the OD trips are redistributed because of the new development in TAZ 36, which produces a little more links with less traffic compared to the base year.

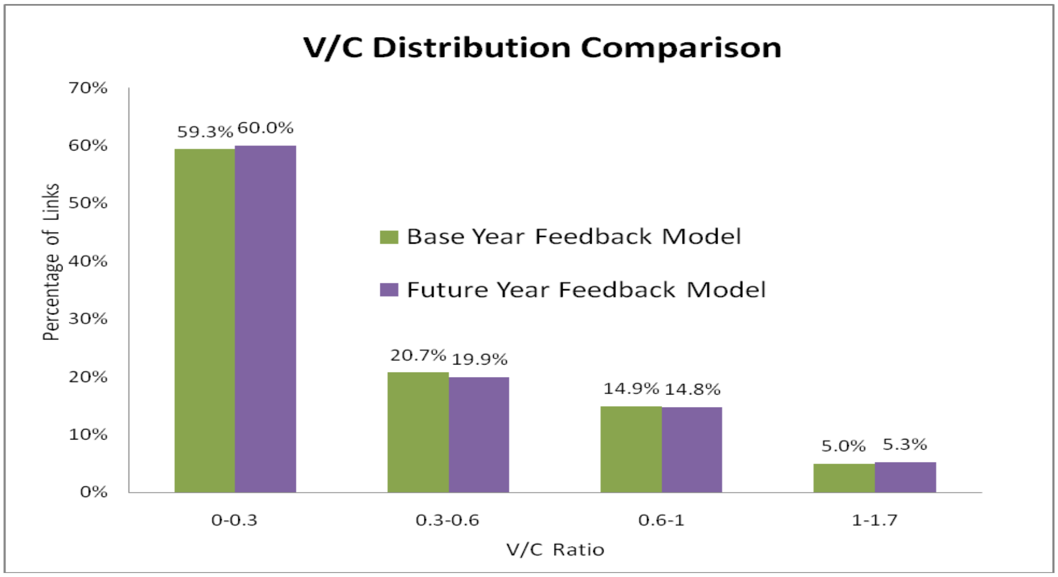

Figure 6.4: Comparisons of V/C Ratio Frequency Distribution between the Base Year and the Future

Similarly, this feedback integrated model can be used to evaluate the impact of alternative transportation projects on this study area in the future. Although the convergence can be heuristically reached for this study area, there is no guarantee that this feedback model is always converged in theory for other study areas or applications. It is encouraged to attempt to eliminate the feedback loop by reformulating the model. The next section proposes an approach to simultaneously solve the land use model and the transportation model instead of using an iterative procedure, which creates another methodology of examining the interaction between the land use model and the transportation model.

6.3 SIMULTANEOUS MODEL CONFIGURATION

Integration of the regression-based land use model and the combined trip distribution-assignment model can be investigated under a simultaneous model structure. Simultaneous model configuration refers to solving the land use model and the transportation model at the same time rather than using an iterative procedure as shown in the feedback model configuration.

6.3.1 Simultaneous Model Framework

As discussed in the feedback model configuration, H_i serves as an input to the combined transportation model. If household distribution model equation 6.1 is put into trip generation constraints 4.13 b-c, equations 4.13b-c will be transformed into the following forms:

$$\sum_j t_{ij} = O_i = 3.1369 \times \{Area_i \times [64.762 \times (X_{1i} + X_{2i}) + 0.224 \times \sum_j \mathrm{E}_j \times e^{-\beta c_{ij}}]\} + 1.2663 E_i$$

6.2a

$$\sum_i t_{ij} = D_j = 3.1369 \times \{Area_j \times [64.762 \times (X_{1j} + X_{2j}) + 0.224 \times \sum_i \mathrm{E}_i \times e^{-\beta c_{ij}}\} + 1.2663 E_j$$

6.2b

O_i and D_j turn into functions of total residential land use fraction, TAZ area, employment distribution, and congested travel time (c_{ij}) as illustrated in equations 6.2a and 6.2b. All variables in O_i and D_j equations 6.2a-b are given except congested travel time (c_{ij}). A challenge for formulating the simultaneous integrated model lies in the formulation of c_{ij}, the minimum congested travel time between an OD pair form i to j. It is the result of solving the user equilibrium transportation model such as the combined transportation model 4.13a-f. Under user equilibrium conditions, all used paths have equal and minimum travel time; unused paths do not carry any trip and have a higher travel time; no road user can improve his/her travel time by switching paths. In the simultaneous integrated model framework, there is no use equilibrium transportation model to offer c_{ij}. An alternative formulation needs to be found for c_{ij}. According to user equilibrium conditions, c_{ij} is the travel time on used paths; therefore, c_{ij} can be formulated if the used path set can be found. Chen and Bernstein (2004) proposed a methodology to find the used path set for toll road modeling. This method makes use of a maximum entropy model proposed by Larsson, et. al (2001), which is used to find the most likely path flow pattern under the constraints of a user equilibrium link flow pattern. In this entropy model formulation, the equilibrium link flow pattern is regarded as the macro state. The path choices of individual travelers are defined as a set of micro states. A variety of micro states will give rise to the same macro state. All micro states are equally probable to take place due to the assumption that an individual traveler's behavior is the same in choosing the shortest paths. Based on the well-known entropy concept, the path flow pattern that engenders the greatest number of micro states under the constraints of macro state is the most likely path flow pattern. The maximum entropy model can be formulated in equations 6.3a-d (Larsson et al, 2001).

$$Max - \sum_{r \in R} h_r Lnh_r \qquad 6.3a$$

Subject to

$$\sum_{r \in R_{ij}} h_r = t_{ij} \qquad 6.3b$$

$$h_r > 0 \qquad 6.3c$$

$$\sum_{r \in R} h_r \delta^{ar} = v_a \qquad 6.3d$$

This optimization program 6.3a-d has strictly convex objective function, as well as linear constraints. In this program, equilibrium link flow pattern (v_a) and OD trips (t_{ij}) are given variables generated from the outputs of the combined trip distribution-assignment transportation model 4.13a-f. This program has a unique global optimal solution with respect to path flow. The optimal solution is the most likely path flow pattern under user equilibrium link flow pattern.

It is important to mention that the most likely path flows out of the entropy model (6.4a-d) are always greater than zero ($h_r > 0$) because of the logarithm type of objective function (6.4a). Under the optimal solution, all path flows are greater than zero, and none of them are equal to zero. However, from the perspective of transportation engineers/planners, any path whose path flow is less than 1 is considered the unused path, since trips are always larger than 1 in reality. The paths with flows more than or equal to

1 will be regarded as the used paths. Therefore the "reasonable" used path set can be extracted from the most likely path flow pattern in conjunction with this reasonable threshold/tolerance (1).

Let UR_{ij} denotes the used path set between origin i and destination j, for $i \in I, j \in J$; and UR represent the used path set in the whole study area, which is the union of UR_{ij}. After obtaining the most likely path flow pattern by solving the entropy maximization problem (6.3a-d), UR_{ij} and UR are derived from the most likely flow pattern. Thus, c_{ij} can be formulated in equation 6.4.

$$c_{ij} = \sum_a s_a(v_a) \times \delta_{ij}^{ar} \quad \text{For } r \in UR_{ij} \qquad 6.4$$

Equation 6.4 indicates that all used paths between OD pairs (from i to j, for $i \in I, j \in J$) have equal and minimal travel time (c_{ij}).

In an attempt to formulate the simultaneous mode framework, it is worth recalling the transportation model formulation (4.13a-f). In the objective function of the transportation model: $Min \sum_a \int_0^{v_a} s_a(w)dw + \frac{1}{0.1993} \sum_i \sum_j (t_{ij} \ln t_{ij} - t_{ij})$, the term $\sum_a \int_0^{v_a} s_a(w)dw$ does not have any economic or behavioral meaning; it is strictly constructed as a mathematical formulation to obtain user equilibrium conditions (all the used paths between OD pairs have equal and minimal travel time). In the simultaneous model formulation, the unused paths have been removed by the most likely path flow pattern and reasonable threshold; the formulation of c_{ij} (equation 6.4) has guaranteed user equilibrium conditions. Therefore $\sum_a \int_0^{v_a} s_a(w)dw$ is removed from the objective function. After identifying the used path set (UR_{ij} and UR) and formulating c_{ij}, the simultaneous model can be formulated as the following optimization program.

$$Min \frac{1}{\beta}\sum_{i}\sum_{j}(t_{ij}\ln t_{ij} - t_{ij}) \tag{6.5a}$$

Subject to

$$\sum_{j} t_{ij} = O_i = 3.1369 \times \{Area_i \times [64.762 \times (X_{1i} + X_{2i}) + 0.092\sum_{j} \mathrm{E_j} \times e^{-\beta c_{ij}}]\} + 1.2663E_i \tag{6.5b}$$

$$\sum_{i} t_{ij} = D_j = 3.1369 \times \{Area_j \times [64.762 \times (X_{1j} + X_{2j}) + 0.092\sum_{i} \mathrm{E_i} \times e^{-\beta c_{ij}}]\} + 1.2663E_j \tag{6.5c}$$

$$\sum_{r \in UR_{ij}} h_r = t_{ij} \tag{6.5d}$$

$$v_a = \sum_{r \in UR} h_r \delta^{ar} \tag{6.5e}$$

$$c_{ij} = \sum_{a} s_a(v_a) \times \delta_{ij}^{ar} \quad \text{For } r \in UR_{ij} \tag{6.5f}$$

$$h_r \geq 0 \tag{6.5g}$$

The item ($\frac{1}{\beta}\sum_{i}\sum_{j}(t_{ij}\ln t_{ij} - t_{ij})$) in the objective function is to produce trip distribution under the entropy concept as discussed in Chapter 4. The solutions of this simultaneous model (equations 6.5a-g) have satisfied the same user equilibrium condition and entropy concept as the solutions of the combined trip distribution-assignment model. In this simultaneous model framework, variables of $Area$, X_1, X_2, are given; t_{ij}, c_{ij}, v_a, are unknown variables whose solutions are the outputs of this simultaneous model. The simultaneous model structure can be illustrated in Figure 6.5.

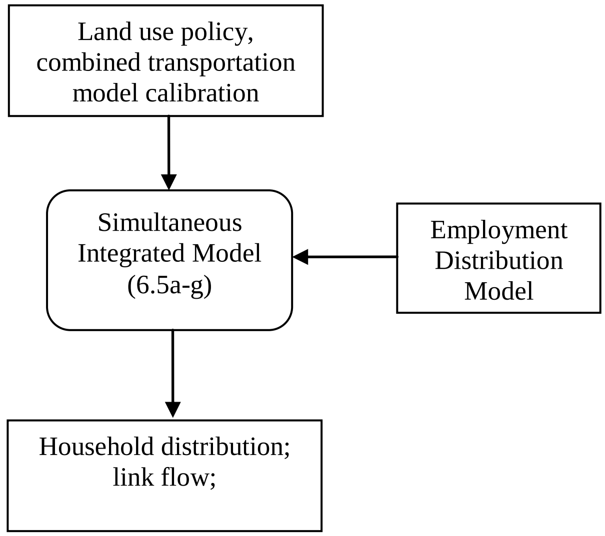

Figure 6.5: Simultaneous Model Configuration

6.3.2 Simultaneous Model Solution

This optimization program of the simultaneous model is a complicated nonlinear problem with a nonlinear objective function, linear equality constraints, and nonlinear equality constraints. The strict convexity of the objective function (equation 6.5a) has been proven in Chapter 4. Equations (6.5b, c, f) including variable (c_{ij}) are nonlinear equality equations. The existence of nonlinear equality constraints creates complexity for this program because the feasible region defined by these nonlinear equality constraints is non-convex. Therefore, the global optimal solution of this optimization program is not guaranteed. However, a reasonable local optimal solution can satisfy planning needs from the perspective of a transportation planner/engineer if it can be attained. The procedure to solve this simultaneous model is discussed using base year data.

The program starts with solving the combined trip distribution-assignment transportation model (equations 4.13a-f) using base year data, whose outputs are link flow (v_a) and OD trips (t_{ij}). v_a and t_{ij} are then substituted into the entropy model (equations 6.3a-d) to generate the most likely path flow (h_r). Based on the most likely path flow (h_r) and the reasonable threshold (1 trip), the used path set can be identified as

UR^0. UR^0 is placed into the simultaneous model (6.5a-g) to forecast link flow, OD trip, and household distribution. Since the household distribution model outputs and the transportation model outputs are simultaneously obtained, consistency between the household distribution and the transportation model outputs is reached.

However, special attention needs to be paid to the used path set. It is possible for the set of used path to vary after solving the simultaneous model. Once the simultaneous model is solved, the updated used path set (defined as UR^1) can be derived from the outputs of the simultaneous model by following the same procedure as finding UR^0. If the initial used path set UR^0 is different from UR^1, the updated used path set UR^1 will be substituted into the simultaneous model (6.5a-g) again to estimate link flow, OD trip, and household distribution. The iterations with regard to the used path set will continue until the used path set is the same during the last two consecutive iterations. UR^i is defined as the used path set at i-th iteration. The procedure for solving the simultaneous model is summarized as follows:

- Solve the combined trip distribution-assignment transportation model (equations 4.13a-f) to obtain link flows and the OD trips
- Solve the most likely path flow model (6.3a-d); use the reasonable threshold (1 trip) to identify the initial used path set UR^0
- Solve the simultaneous model (equations 6.5a-g) using the path set UR^{i-1} to obtain link flow, OD trip, and household distribution; set i=1 where i means i-th iteration
- Solve the most likely path flow model (6.3a-d) using the outputs of the simultaneous model at iteration i; identify the updated used path set UR^i based on the reasonable tolerance
- Check the consistency between UR^i and UR^{i-1}; if they are entirely identical, stop the iteration, and report link flow, OD trip, and household distribution; otherwise, substitute UR^i into step 3 replacing UR^{i-1} to repeat the procedure.

The simultaneous model framework is tested on the same network and zone structure of Woodford County as the feedback model framework. The interior point algorithm associated with solvers of KNITRO and LOQO is used to solve this simultaneous model. During the procedure of searching local optimal solutions for this simultaneous model (equations 6.5a-g), LOQO solver fails to converge no matter what starting point is used. KNITRO solver succeeds in searching local optimal solutions for this simultaneous model. However, KNITRO can only converge to the local optimal solutions using the reasonable starting point. The reasonable starting point refers to using the base year data such as link flow and OD trip to solve the base year simultaneous mode, and using the future year data (pre-estimated link flow and OD trip) for solving the future simultaneous model. The future year data as starting point will be discussed in the section on Future Model Application.

The simultaneous model framework is applied to both base year and future year scenarios. Although the same used path set between two consecutive iterations is not mathematically guaranteed, the same used path set is generated after two iterations for both base year and future model application.

A reasonable local optimal solution for the base year is attained after two iterations with regard to the used path set. In the first iteration, there is no significant change between the initial used path set UR^0 and the first iteration used path set UR^1. For example, the paths with number of trips greater than 5 is the same between UR^0 and UR^1; the variation only takes place on those paths with number of trips less than 5; 2 percent of paths do not appear in UR^1 because the number of trips on them are decreased from a little greater than 1 to less than 1. During the second iteration, the same used path set is reached between UR^1 and UR^2.

After the used path set is identical between the last two iterations, the outputs of the simultaneous model are reported as the final output. The outputs associated with link flow, OD trip and household distribution are illustrated in Table 6.2. The same links, OD pair, and household distribution as the feedback model framework are selected to be displayed. A comparison of the results between the feedback model and the simultaneous model will be discussed in section 6.4.

Table 6.2: Selected Simultaneous Model Output

Link	Link flow	Origination	Destination	OD Trips	Zone	Households
L196	8895	Z74	Z75	1407	Z56	1628
L679	7581	Z75	Z74	1407	Z73	1132
L356	6566	Z72	Z73	259	Z72	1206
L143	7156	Z73	Z72	259	Z36	1174
L133	9218	Z36	Z56	533	Z75	1157
L707	7413	Z56	Z36	533	Z74	1041
L110	7570	Z65	Z64	558	Z38	1018
L111	7570	Z64	Z65	558	Z33	761
L158	8532	Z76	Z78	445	Z64	806
L633	9289	Z78	Z76	445	Z62	731
L634	9289	Z64	Z62	487	Z58	478
L105	7335	Z62	Z64	487	Z37	574
L106	7335	Z75	Z76	320	Z50	514
L107	7335	Z76	Z75	320	Z39	560
L108	7335	Z56	Z72	266	Z51	396
L109	7335	Z72	Z56	266	Z76	436
L632	8442	Z73	Z56	272	Z61	449
L409	6646	Z56	Z73	272	Z32	433
L403	5328	Z37	Z36	262	Z30	366
L636	5328	Z36	Z37	262	Z29	372

Similarly, the simultaneous model outputs can play the same role as the feedback model outputs in evaluating overall performance of the existing road network. Also, the V/C ratio can be used to assess the level of congestion in the road network. The V/C ratio frequency distribution is illustrated in Figure 6.6, which shows that the study area has a low level of congestion with an average V/C ratio of 0.35.

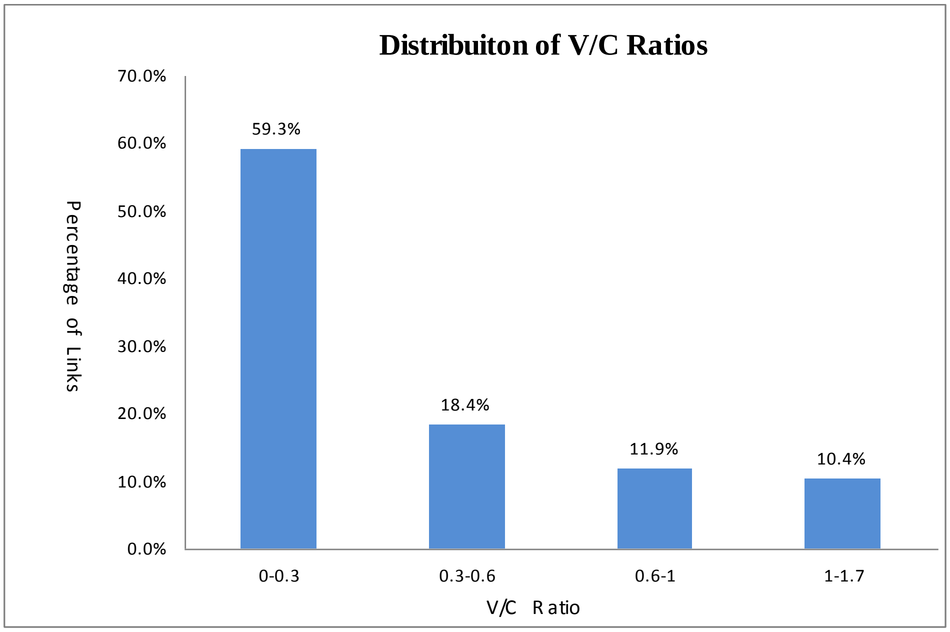

Figure 6.6: V/C Ratio Frequency Distribution from Simultaneous Model Output

Figure 6.6 shows that 10.4 percent of links have a V/C ratio of between 1 and 1.7, which indicates no extremely congested links on the existing road network. 89.6 percent of links carry traffic under their capacities; 11.9 percent of links have a V/C ratio of between 0.6 and 1; 18.4 percent of links have a V/C ratio of between 0.3 and 0.6; and 59.3 percent of links have a V/C ratio of under 0.3.

6.3.3 Future Model Application

Similarly, the simultaneous model can be applied to future year scenarios. The same potential residential land development as the feedback model is used for testing the simultaneous model framework, shown in Figure 6.3.

Finding the used path set is the critical step in formulating and solving the simultaneous model. The base year used path set is most likely different from the future year used path set since more trips are generated in the future due to the new land development. More trips are able to result in more used paths, some of which may be unused in the base year. Preliminary future year model testing has proven that a future year model is not able to converge to a local optimal solution when using the base year used path set. Therefore, the future year used path set needs to be determined before running the future year simultaneous model. The future combined trip distribution-assignment transportation model is solved to assist in finding the future year used path set. When solving the future transportation model, the future household distribution as a key input to the transportation model is estimated using the base year congested travel time matrix. Based on link flow and OD trips from the future transportation model output, equations 6.4a-d are performed to obtain the most likely path flow pattern. Reasonable tolerance (1 trip) is then used to identify the initial used path set. These link flow and OD trips also will be used as the starting point to solve the future simultaneous model.

The future year model is successfully solved after two iterations by showing the identical used path set between the first and second iterations, and the convergence to a local optimal point. The generated output such as household distribution is shown in Figure 6.7. Results show that the number of households in TAZ 36 increases from 1,173 in the base year (base year simultaneous model output) to 1,434 in the future due to the increasing residential land use. The newly generated households will produce additional trips (819) in comparison with the base year. These additional 819 trips lead to a few more trips on neighborhood road networks, which decrease the accessibility of neighborhood TAZs. The decreasing accessibility consequently reduces the number of households in these TAZs. For example, the number of households in neighborhood TAZ 56 will be reduced from1,628 to 1,627. This very small decrease is a result of the small change in its accessibility measure caused by the insignificant change in traffic flow between the base year and the future. In this case, future household distribution does not have a significant traffic impact on the network in comparison with the base year. For example, the road segment of US 60 along the development area (link no.145) has a daily traffic volume of 17,297 (the base year simultaneous model output), with a daily capacity of 22,000, and a V/C ratio of 0.786 in the base year; it has a daily traffic of 17,357, and a V/C ratio of 0.789 in the future.

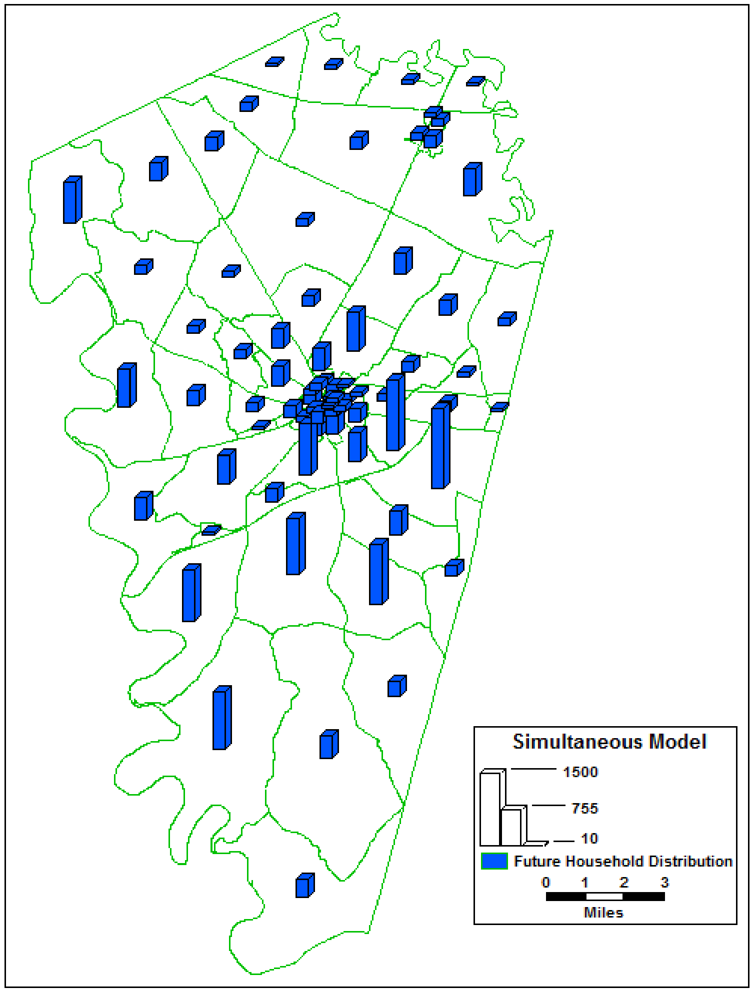

Figure 6.7: Future Household Distribution From Simultaneous Model

As can be seen, both the feedback model framework and the simultaneous model framework can be used to estimate the future traffic condition and household distribution. The next section will compare these two model frameworks.

6.4 MODEL COMPARISON

The feedback model framework and the simultaneous model framework have different model structures, which result in different model outputs. This section compares these two models associated with base year model outputs and model structure.

6.4.1 Model Structure Comparison

The feedback model and the simultaneous model have some similarities. First, both models can produce link flow, OD trip, and household distribution output. Second, both model outputs are able to satisfy user equilibrium conditions which state that all used paths between each OD pair have equal travel time and no road user can improve his/her travel time by switching paths (Wardrop, 1952). Lastly, both models are capable of generating consistency between the land use model output and the transportation model output.

However, the structure of the feedback model is different from the simultaneous model. In the feedback model structure, the transportation model and the land use model are separately developed and solved; the feedback solution procedure is built between them to resolve these two models. In the simultaneous model framework, the land use model is embedded into the transportation model constraints by adding the new variable c_{ij}, which converts the transportation model into the simultaneous model. In order to formulate c_{ij}, the maximum entropy model (6.3a-d) is first solved to help in finding the used path set.

In the feedback model, the used path set does not take part in the formulation at all; instead, the first three shortest paths are input to the transportation model. But in the simultaneous model, the used path set is critical input; it is pre-defined based on the most likely path flow pattern and then participates in the model formulation. The used path set has already been determined before solving the simultaneous model. In the feedback model, the transportation model outputs show that hundreds of paths have path volume below 1. In the simultaneous model, the pre-defined used path set only consists of the paths whose volume is not less than 1.

In the feedback model framework, an iteration procedure is established between the transportation model and the land use model. The iterations continue until the pre-defined convergence criteria are reached. Although there is no iteration procedure between the transportation model and the land use model in the simultaneous model framework, there does exists an iteration procedure between the simultaneous model and the maximum entropy model (6.3a-d); the iterations continue until the used path set is identical in the last two iterations.

In the feedback model framework, land use model equations can be easily solved; the transportation model is formulated as an optimization program with one and only one global optimal solution. Both KNITRO and LOQO can successfully converge to the optimal point; KNITRO is able to converge to the optimal point no matter what the starting point is. In the simultaneous model framework, there is no guarantee that the global optimal solution exists and can be found since it is a complicated nonlinear optimization program. Only KNITRO solver is capable of converging to the local optimal point using the appropriate starting point.

6.4.2 Model Output Comparison

It is expected that the simultaneous model output will be different from the feedback model output to some degree because of the difference in model structure. In the feedback model structure, the used path set is not involved in the model formulation at all; the first three shortest paths are input to the transportation model, and there are hundreds of paths in the model output with a path volume below 1. However, in the simultaneous model structure, the used path set is pre-defined based on the most likely path flow model; the used path set only includes the paths whose volume is not less than 1. Therefore, OD trips use fewer paths in the simultaneous model than the feedback model. Thus, the simultaneous model distributes OD trips over fewer links than the feedback model overall. The distribution of OD trips over fewer links causes the simultaneous model to produce more congested links and longer travel time between OD pairs than the feedback model. The longer travel time generates less accessibility for each TAZ in general. Because of the smaller accessibility measure, the simultaneous model outputs show relatively fewer households in each TAZ than the feedback model.

The used path set comparison between the feedback model outputs and the simultaneous model outputs shows that there are certain differences between them. In the simultaneous model outputs, out of a total of 6,006 OD pairs, 318 OD pairs have two used paths between them; all other OD pairs have only one used paths. In the feedback model outputs, out of a total of 6,006 OD pairs, 489 OD pairs have two used paths, and all other OD pairs have only one used path. In the outputs of both models, no OD pairs has more than two used paths. The 318 OD pairs in the simultaneous model outputs also have two used paths in the feedback model outputs. The feedback model outputs have an additional 171 OD pairs with two used paths.

The outputs of these two models are compared with regard to link flow, OD trip, and household distribution using base year data. The variation is calculated to measure the difference between these two model outputs. For example, the variation as to OD trips is defined as: $\left|\frac{t_{ij}^{f} - t_{ij}^{s}}{t_{ij}^{s}}\right| \times 100$, where the superscripts s and f represent the simultaneous model and the feedback model respectively. The comparisons of link flow, OD trip, and household distribution as well as percent of variation are listed in Table 6.3. Frequency distribution of V/C ratio is shown in Figure 6.8. The comparison associated with household distribution is illustrated in Figure 6.9. As to the link flow comparison between the two model outputs, certain difference are expected because of different model structure and model inputs. The comparisons show that 399 links out of a total of 723 links have link flow variation below 5 percent; 189 links have link flow differences between 5 percent and 20 percent; 135 links have link flow differences between 20 percent and 50 percent; no links have link flow differences of more than 50 percent.

As shown in Figure 6.8, 10.4 percent of links from the simultaneous model have a V/C ratio greater than one, which is 5.4 percent higher than the feedback model because of more congested links. Consequently, the simultaneous model generates a lower percentage of links with smaller V/C ratios than the feedback model. For example, 18.4 percent of the links in the simultaneous model have a V/C ratio of between 0.3 and 0.6, which is 2.3 percent less than in the feedback model. In the simultaneous model, 11.9 percent of links have a V/C ratio of ranging from 0.6 to 1, which is 3 percent less than in the feedback model. Both models have the same percent of links with a V/C ratio below 0.3. The average V/C ratio in the simultaneous model is 0.35, while the average V/C ratio in the feedback model is 0.32.

Household distribution comparison shows that the estimated total households from the simultaneous model are slightly less than the feedback model. This causes the total number of OD trips from the simultaneous model to be lower than in the feedback model. Although the simultaneous model produces fewer total trips than the feedback model, V/C ratio distribution shows that the simultaneous model has higher percentages of links with a V/C ratio of over 1. The fewer total trips from the simultaneous model are distributed in a smaller portion of road segments compared to the feedback model. This is consistent with previous analysis showing that OD trips are distributed over fewer links in the simultaneous model than in the feedback model due to the different used path set.

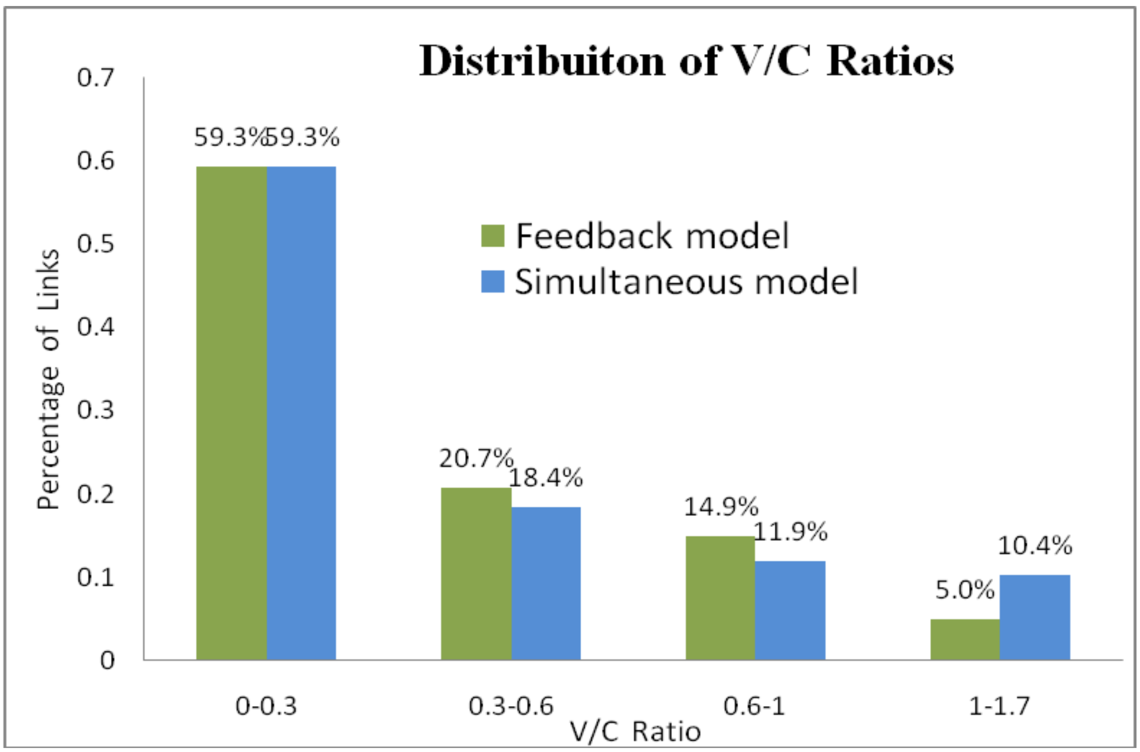

Figure 6.8: Comparisons of V/C Ratio Frequency Distribution between Feedback and Simultaneous Model

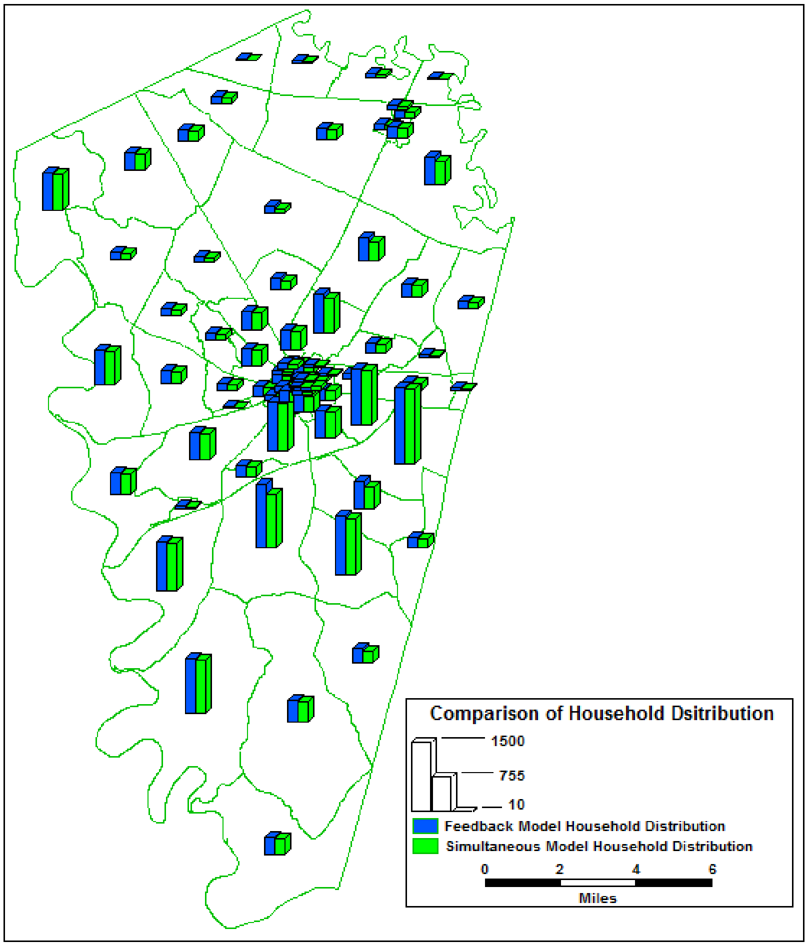

Figure 6.9: Comparisons of Household Distribution between Feedback and Simultaneous Model

Table 6.3: Comparisons between Selected Feedback and Simultaneous Model Output

Link	Link flow (feedback model)	Link flow (simultaneous model)	Variation between these two models (percent)	Origination	Destination	OD Trips (Feedback)	OD Trips (Simultaneous)	Variation between these two models (percent))	Zone	Household (feedback model)	Household (simultaneous model)	Variation between these two models (percent))
L196	6856	8895	22.9	Z74	Z75	1431	1407	1.7	Z56	1635	1628	0.4
L679	6823	7581	10.0	Z75	Z74	1431	1407	1.7	Z73	1335	1132	17.9
L356	6559	6566	0.1	Z72	Z73	850	259	228.2	Z72	1257	1206	4.2
L143	6354	7156	11.2	Z73	Z72	850	259	228.2	Z36	1180	1174	0.5
L133	6252	9218	32.2	Z36	Z56	587	533	10.1	Z75	1163	1157	0.5
L707	6194	7413	16.4	Z56	Z36	587	533	10.1	Z74	1045	1041	0.4
L110	6044	7570	20.2	Z65	Z64	586	558	5.0	Z38	1027	1018	0.9
L111	6044	7570	20.2	Z64	Z65	586	558	5.0	Z33	839	761	10.2
L158	5900	8532	30.8	Z76	Z78	566	445	27.2	Z64	809	806	0.4
L633	5806	9289	37.5	Z78	Z76	566	445	27.2	Z62	733	731	0.3
L634	5806	9289	37.5	Z64	Z62	502	487	3.1	Z58	579	478	21.1
L105	5791	7335	21.0	Z62	Z64	502	487	3.1	Z37	578	574	0.7
L106	5791	7335	21.0	Z75	Z76	365	320	14.1	Z50	575	514	11.9
L107	5791	7335	21.0	Z76	Z75	365	320	14.1	Z39	572	560	2.1
L108	5791	7335	21.0	Z56	Z72	356	266	33.8	Z51	472	396	19.2
L109	5791	7335	21.0	Z72	Z56	356	266	33.8	Z76	459	436	5.3
L632	5477	8442	35.1	Z73	Z56	339	272	24.6	Z61	453	449	0.9
L409	5471	6646	17.7	Z56	Z73	339	272	24.6	Z32	435	433	0.5
L403	5350	5328	0.4	Z37	Z36	330	262	26.0	Z30	384	366	4.9
L636	5350	5328	0.4	Z36	Z37	330	262	26.0	Z29	377	372	1.3

6.5 INTEGRATED MODEL CAPABILITY

The proposed integrated model framework is composed of three components: the transportation model, the land use model, and the interaction between these two models, which have been discussed in Chapter 4, 5, and 6 respectively. The interactions between the land use model and the transportation model are investigated by two different methodologies: feedback model framework and simultaneous model framework. Both of these frameworks can be used to produce consistency between the land use model outputs and the transportation model outputs.

Based upon the estimated statistical parameters, the proposed integrated model framework can provide feasible solutions associated with traffic flows and household distribution. For example, the transportation model calibration shows that 91 percent of variation in originating/destined trips can be explained by the trip generation model as discussed in Chapter 4. The originating/destined trips are obtained from the OD trip table in the outputs of the WTDM, which has the satisfactory error bound between the observed traffic volume and the modeled traffic volume on 120 traffic count stations as discussed in Chapter 3. Also, the land used model calibration in Chapter 5 shows that 87 percent of variation in household distribution can be explained by the land use model, which demonstrates an acceptable error between the observed household density and the modeled household density.

The proposed land use model includes land use structure variables such as mobile home and multi-family land use fraction and residential land use fraction. It has the capability of simulating the impact of changing land use structure on household distribution. Also, the land use model includes transportation measures in terms of travel time; it allows the model to evaluate how household distribution will respond to network changes in the transportation system. The transportation model includes the variables of household and employment distribution, network structure, and network attributes such as speed and capacity. It can evaluate the impact of household and employment distribution on transportation system performance, and the impact of network changes such as road improvements or new road development on transportation system performance. The consistent solution between the land use model output and the transportation model output is obtained using either the feedback model framework or the simultaneous model framework. It can demonstrate that household/employment distribution is in accord with transportation system performance.

The proposed model framework not only provides procedures for evaluating land use and transportation policies, but also offers a clear implication of dynamic equilibrium between land use and transportation systems. For example, correlation between household densities and accessibility provided by transportation systems can be used to estimate household distribution associated with certain transportation investments. The study focuses on the estimation of household distribution in a macro state. For example, the model outputs are the number of households in a TAZ instead of which individual households are relocated for what reason or which new household moves in. Also, current practices often estimate new transportation facilities based on system performance measures such as V/C ratio from travel demand model outputs, but do not take into consideration induced travel demand due to household redistribution. The proposed model framework can be used to project both new transportation facilities and consequently induced demand. In general, the model outputs imply a dynamic equilibrium between household distribution and relevant transportation system performance. The proposed model framework can be transferred to other areas for their own applications. For example, the procedures for developing the land use model can be used in other areas for forecasting demographic distribution.

CHAPTER 7
CONCLUSIONS AND FUTURE RESEARCH

7.1 CONCLUSIONS

The growth of American's urban area in the past few decades has been accompanied by transportation problems such as congestion and pollution. These problems are not only caused by transportation-system design but also are related to land-use planning. There has been growing recognition that the interactive relationship between land use and transportation needs to be understood and analyzed in a consistent and systematic way. Integrated urban models have recently been introduced to examine the interactive relationship between land use and transportation. The general consensus in this field is that each model has its own limitations because of its specific application purposes. This dissertation develops a new type of integrated land use and transportation model framework: integration of a regression-based land use model and a combined trip distribution-assignment transportation model. This new model can be applied to both metropolitan areas and small urban areas.

The proposed new integrated model framework consists of three components: the land use model, the transportation model, and the interaction between these two models. The combined trip distribution-assignment model serving as the transportation model has rarely been examined in existing integrated urban models. This dissertation explores the formulation, calibration and application of the combined trip distribution-assignment transportation model in the context of an integrated model framework.

The land use model is then developed using correlation and regression analysis based on land-use structure factors and transportation measures. This method has not been used in the current literature of integrated urban models. The regression equation of the land use model is effective in capturing the features of household distribution. For example, 87 percent of variation in household density can be explained by the combination of total residential land use fraction and accessibility measure.

The interaction between the land use model and the transportation model is investigated by two model frameworks: feedback model framework and simultaneous model framework. Both of these are effective in estimating link flow, OD trip, and household distribution in a consistent way. In the feedback model framework, a feedback loop is built between the land use model and the transportation model. The procedure can converge to pre-defined criteria after three iterations when this framework is applied to both base year and future scenarios. In the simultaneous model framework, the combined trip distribution-assignment model is converted into the simultaneous model by formulating the land use model into the transportation model constraints. This is achieved by introducing the used path set. Iterations with regard to the used path set are examined in the simultaneous model. Model testing for the base year and the future shows that the same used path is obtained after two iterations.

It is worth mentioning that the calibrated parameters for the proposed transportation model and land use model are only suitable for this study city. It cannot be transferred to other areas. For a specific application, it is recommended that the same methodology in the model framework can be utilized to develop a regression-based land use model, a combined trip distribution-assignment model, and the integration of these two models.

This dissertation presents the first instance of integration of a regression-based land use model and a combined trip distribution-assignment transportation model. It can be efficiently solved using modern computation solvers and can be implemented by both metropolitan areas and small urban areas with limited resources.

7.2 RECOMMENDATIONS FOR FUTURE RESEARCH

The focus of this study is to develop a framework for integration of the regression-based land use model and the combined trip distribution-assignment transportation model. Continued research to improve the transportation model and the land use model will be of considerable benefit to the reliability and completeness of this model framework.

The proposed combined trip distribution-assignment model concentrated on estimating total trips generated by each TAZ, rather than estimating trips by purpose. Incorporating trip purpose into trip generation is recommended, which would give a better understanding of travel behavior and may improve the model's accuracy. The transportation model is developed on a daily pattern; it is reasonable to assume that the number of trips originating from a zone is equal to the number of trips destined to this zone. However, they are not equal to each other on an hourly basis. It would be interesting to develop a peak-hour transportation model for integration with land use models. In the calibration of the transportation model, base year OD trips serve as key inputs derived from the traditional four-step travel demand model instead of by observation. It is recommended to use actual observed OD trips when data is available.

Several ideas that arose during the research as well as during the land use model analysis process could be worth investigating more thoroughly. It is commonly recognized that households with different levels of income show distinct preference in choosing their residence locations. For example, households with high income prefer living in suburban areas in a bigger house with a longer commuting distance; low-income households prefer living close to downtown in a smaller house with a shorter commuting distance. So it is recommended that land use models could be developed for each category of household by income, and for each category of employment by industry type if additional data becomes available, which may more accurately capture household/employment location choice preferences. The proposed land use model development analyzed a list of land use and transportation factors based on data availability. It would be interesting to study more factors if data is available. For example, market force factors including house price, house size, land value, terrain, etc.

are worth being explored since they do play an important role in both household and employment location choice.

APPENDIX

Appendix 2-A1: ITLUP Land Use Model Equations

DRAM Equations:

$$\hat{N}_{i,t} = \sum_j \hat{E}_{j,t} r_t^n \left[\frac{W_{i,t}^h f^h(\hat{c}_{ij,t})}{\sum_k W_{k,t}^h f^h(\hat{c}_{kj,t})} \right]$$

Where:

$\hat{N}_{i,t}$: The estimated number of households in zone i at time t

$\hat{E}_{j,t}$: The estimated number of employments in zone j at time t

$\hat{c}_{ij,t}$: The congested travel time between zone i and j at time t

$f^h(\hat{c}_{ij,t})$: Travel impedance function or accessibility function

r_t^n: Regional activity ratio of households to employments at time t

$W_{i,t}^h$: The attractiveness function of zone i at time t

There are different forms of travel impedance function such as exponential, inverse power and Gamma function. Which category of function is used mostly relies on which function can effectively fit the trip distribution data. The modified form of Gamma function (a.k.a. Tanner function) is recommended by Putman as illustrated below. According to the curve of Tanner function, most workers do not reside at locations too far away from or too close to their workplaces. There exists a desired trip distance from residence to workplace.

$$f^h(\hat{c}_{ij,t}) = \hat{c}_{ij,t}^{c_h} e^{d_h \hat{c}_{ij,t}}$$

where

c_h, d_h - Empirically derived parameters

The attractiveness function for household distribution is expressed by the following equation.

$$W_{i,t} = \left(L_{i,t}\right)^{\theta} \prod_k \{1 + \frac{N_{i,t}^k}{\sum_k N_{i,t}^k}\}^{\gamma_k}$$

where

$L_{i,t}$: The total land area of zone i at time t

$N_{i,t}^k$: The number of households in zone i in the k income level at time t

$\theta, \gamma_1 \cdots \gamma_k$: Empirically estimated parameters

EMPAL Equations:

$$E_{j,t} = \left(\frac{e^{\delta}}{1+e^{\delta}}\right) r_t^h \sum H_{i,t-1} \left[\frac{W_{j,t-1} e^{\beta_p c_{ji,t}^p + \beta_{op} c_{ji,t}^{op}}}{\sum_k W_{k,t-1} e^{\beta_p c_{ji,t}^p + \beta_{op} c_{ji,t}^{op}}}\right] + \left(\frac{1}{1+e^{\delta}}\right) r_t^e E_{j,t-1}$$

where

r_t^h : The regional ratio of employments at time t to households at time $t-1$

$H_{i,t-1}$: The number of household in zone i at time $t-1$

$W_{j,t-1}$: The attractiveness index of zone j at time $t-1$

r_t^e : The regional ratio of employments at time t to employments at time $t-1$

σ : Empirically derived parameter

The attractiveness function employment distribution can be conveyed as the equation below:

$$W_{j,t-1} = \left(E_{j,t-1}\right)^{\delta_1} \times \left(L_j\right)^{\delta_2}$$

where

L_j : The total land area of zone j

σ_1, σ_2 : Empirically derived parameters

Appendix 2-A2: Quasi-Gravity Model Equations

$$Min \sum_a \int_0^{v_a} s_a(t)dt$$

subject to:

$$\sum_i t_{ij} = E_j$$

$$\frac{1}{T} \sum_i \sum_j t_{ij} h_i = \bar{h}$$

$$-\sum_i \sum_j \frac{t_{ij}}{T} Ln \frac{t_{ij}}{T} = S$$

$$\sum_{r \in R_{ij}} h_r = t_{ij}$$

$$v_a = \sum_{r \in R} h_r \delta^{ar}$$

$$h_r \geq 0$$

where:

t_{ij} : The number of trips from origin zone i to destination zone j

h_r : The number of trips on path r between zone i and j; R_{ij} is the set of path from zone i to j, I is the set of origination zones and J is the set of destination zones

$s_a(v_a)$: The travel cost on link a, which is an increasing function of link flow v_a for $a \in A$; A is the set of all links in the transportation network

δ^{ar} : The incidence coefficient that describes the relationship between path and link, $\delta^{ar} = 1$ if the link a is on path r; $\delta^{ar} = 0$ otherwise

E_j : The number of employment or jobs in zone j

h_i : The residential benefit of choosing zone i to live in since all zones are not equally attractive

$\bar{h}$: The mean benefit of living in the study region

Appendix 2-B: MEPLAN Model Equations

$$T_{cj}^{n} = D_{cj}^{n} + Q_{cj}^{n} \quad \text{with}$$

$$D_{cj}^{n} = \sum_{m} a_{j}^{mn} T_{gj}^{m}$$

where

j : The index for land use zones

m : The index for economic sector

n : The index for economic sector

T_{cj}^{n} : The total volume of factor n consumed in zone j

D_{cj}^{n} : The endogenous component of total volume of factor n consumed in zone j

Q_{cj}^{n} : The exogenous component of total volume of factor n consumed in zone j

a_j^{mn} : The volume of factor n consumed in the production of a unit of factor m in zone j

T_{gj}^{m} : The total volume of factor m produced in zone j

After the volume of sector n consumed in zone j is derived out, random utility choice modeling in a spatial context is developed to seek the volume of sector n that will be produced in zone i . The following formula is used to allocate this production.

$$t_{ij}^{n} = T_{cj}^{n} \frac{\exp[\lambda^{n}(T_{bi}^{n} + d_{ij}^{n} + s_{i}^{n} + Q_{\varepsilon i}^{n} + D_{\varepsilon i}^{n})]}{\sum_{i} \exp[\lambda^{n}(T_{bi}^{n} + d_{ij}^{n} + s_{i}^{n} + Q_{\varepsilon i}^{n} + D_{\varepsilon i}^{n})]}$$

where

i : The index for land use zones

t_{ij}^{n} : The volume of economic sector n produced in zone i and consumed in zone j

λ^{n} : The dispersion parameter associated with economic sector n

T_{bi}^{n} : The cost of producing one unit of sector n in zone i

d_{ij}^{n} : The disutility associated with transporting one unit of sector n from zone i to zone j

s_{i}^{n} : A size term that accounts for a *priori* likelihood that one unit of sector n i produced in zone i

$Q_{\varepsilon i}^{n}$: The exogenous component of zone-specific disutility associated with producing sector n in zone i

$D_{\varepsilon i}^{n}$: The endogenous component of zone-specific disutility associated with producing sector n in zone i

Appendix 2-C: Location Choice Model Equation in five-Stage Urban Model

The probability ($P_{i/h}$) that the customer h will choose lot i can be depicted as be:

$$P_{i/h} = \frac{\exp[\mu(WP_{hi} - P_i)]}{\sum_{j\in S} \exp[\mu(WP_{hj} - P_j)]}$$

where h represents customers; i is the index for land lot (planning zones), S is the set of land lots; WP is the willingness to pay function; P is the price function. μ is empirically derived parameter.

Appendix 3-A: Traditional Four-Step Travel Demand Model Development

There are three major steps in this traditional travel demand model development for the study area: trip generation, trip distribution, and traffic assignment. The step of mode split is skipped due to data availability.

Trip Generation

Internal Trip Generation

The model uses a cross-classification trip method for trip generation. The trip production rate varies by household size and income, and the trip attraction rate is mainly associated with employment classification and households. Household data is obtained from the 2000 U.S. Census survey; employment data is provided by Dun & Bradstreet (D&B). In this Woodford County model, households are categorized into low (<25th percentile), medium (25th-75th percentile) and high (>75th percentile) income group. The household income data is only available at the bigger spatial level of census block group. One census block group is composed of several TAZs. It is assumed that household income distribution is even cross the TAZs in the same census block group. This indicates that household income data is not adequately accurate in developing this travel demand model. The employment data were obtained originally by category of standard industrial classification (SIC). The classification system is illustrated as below.

01-09	Agriculture, Forestry, and Fishing
10-14	Mining
15-17	Construction
20-39	Manufacturing
40-49	Transportation, Communications, and Utilities
50-51	Wholesale Trade
52-59	Retail Trade
60-67	Finance, Insurance, and Real Estate
70-89	Services
91-97	Public Administration
99	Non-classifiable Establishments

For modeling purposes, employment data is divided into three different groups because trip attraction rate is highly related with industry type. According to NCHRP Report 365, the employment data is divided into the categories of Basic, Service, and Retail. Each category is composed of corresponding industry types.

Basic:	Major groups 1 through 51 and 91 through 99
Service:	Major groups 60 through 90
Retail:	Major groups 52 through 59.

Some verification effort has been made to enhance the accuracy of the data as close to reality as possible. Control totals for Woodford County employment were obtained from the Kentucky Cabinet for Economic Development (CED). The total number of employees as listed by D&B is mostly consistent with the CED record. The locations of major employment where the number of employees is more than 200 as indicated by CED match with those listed by D&B at the TAZ level. The inputs for trip generation can include the number of household by income and size for each TAZ, the amount of employment by basic, service and retail type, and the production and attraction rates.

Trip Production

The trip production is estimated for each different trip purpose: Home Based Work (HBW), Home Based Other (HBO), and Non-Home-Based (NHB). The trip rate recommended by NCHRP Report 365 is used, as shown in Table A1.

Using the trip rates by income and household size, the total trips by trip purpose can be estimated. For example, TAZ no.24 has 68 low-income households, 115 middle-income households and 73 high-income households. The trip production for this TAZ is then estimated as:

HBW trip productions = 68 x 1.04 + 115 x 1.53 + 73 x 1.84 = 381 trips

HBO trip productions = 68 x 3.90 + 115 x 4.09 + 73 x 5.06 = 1105 trips

NHB trip productions = 68 x 1.56 + 115 x 1.68 + 73 x 2.30 = 467 trips

Therefore, the total trips produced from this TAZ are 1953 per day.

Table A1: Trip Production Rate

Household Income Group	HHSIZE	Rate_HBW	Rate_HBO	Rate_NHB
Low	1	0.58	2.16	0.86
Low	2	1.04	3.90	1.56
Low	3	1.46	5.46	2.18
Low	4	1.84	6.90	2.76
Low	>=5	2.21	8.28	3.31
Medium	1	0.82	2.18	0.90
Medium	2	1.53	4.09	1.68
Medium	3	2.10	5.60	2.30
Medium	4	2.75	7.34	3.01
Medium	>=5	3.34	8.90	3.66
High	1	0.90	2.48	1.13
High	2	1.84	5.06	2.30
High	3	2.44	6.71	3.05
High	4	2.96	8.14	3.70
High	>=5	3.64	10.01	4.55

Trip Attraction

The number of trips attracted to a zone is estimated based upon the trip attraction rates as shown in Table A2. Trip attraction is estimated for the same three different trip purposes as in trip production.

For illustration purposes, trip attractions are estimated for the same TAZ, TAZ no.24, which contains 256 households and 71 total jobs with 4 in basic, 67 in service and none in retail sectors. Trip attractions are estimated as:

HBW trip attractions = 1.45 x 67 = 97 trips

HBO trip attractions = 9.00 x 0 + 1.7 x 67 + 0.5 x 4 + 0.9 x 256 = 346 trips

NHB trip attractions = 4.1 x 0 + 1.2 x 67 + 0.5 x 4 + 0.5 x 256 = 210 trips

Table A2: Trip Rate for Trip Attraction

Input data	HBW	HBO	NHB
Total employment	1.45	N/A	N/A
Retail employment	N/A	9.00	4.10
Service employment	N/A	1.70	1.20
Basic employment	N/A	0.50	0.50
Household	N/A	0.90	0.50

Since trip production and attraction are calculated using different formula and input, the total trip production is most likely not equal to total trip attraction. A balancing effort is made by integrating internal generation with external generation based on the trip purposes.

External Trip Estimation

Internal trip generation deals with the trips generated by households and employments inside the study area. However, there are a number of trips on the transportation network that are generated by household and employment outside of the study area. These trips are regarded as external trips; they can be defined as the trips that have at least one end outside the study area. If both the origin and destination of a trip are outside the area, this trip is considered as a through trip or external-external trip. For example, a trip is made from Lexington to Frankfort and pass through Woodford; this trip is classified as an external-external trip. When only one trip end is outside the study area in either origin or destination, this trip is categorized as an internal-external or external-internal trip. For example, a trip is made from Lexington to Woodford; this trip is an external-internal trip.

External trip data is usually collected by conducting a roadside intercept travel survey. Unfortunately the study area does not implement this survey. In this study, external trips will be estimated based on observed traffic flow in the external stations and using procedure recommended by NCHRP Report 365. External stations are located outside Woodford County. External stations are located at the point along a route where the Woodford County boundary line was crossed. A total of 22 external stations were defined for the Woodford County area.

Figure A1 shows the location of external stations and connections between external stations and road network.

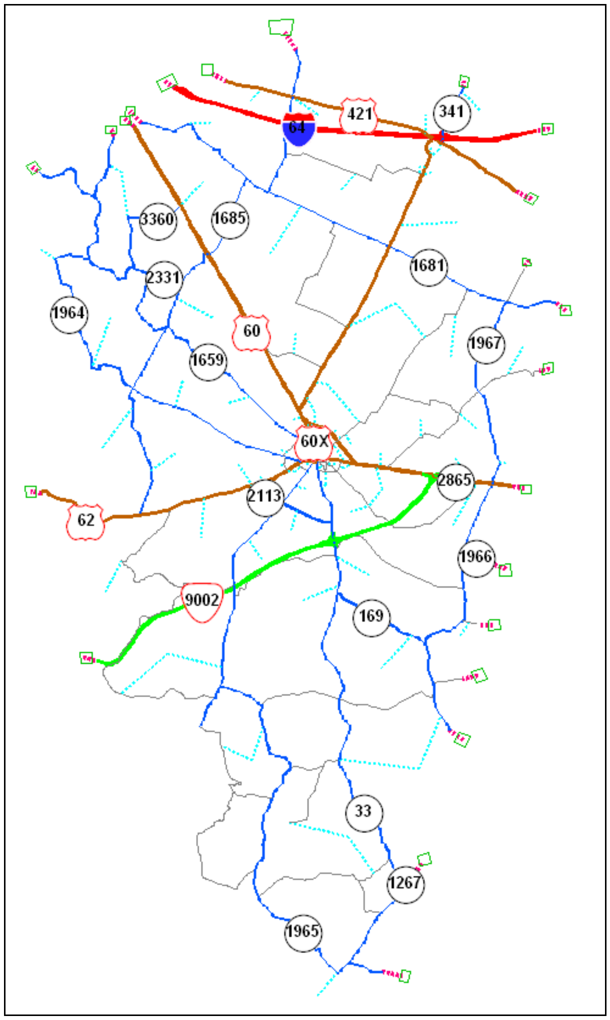

Figure A1: External Station Distribution

There are two major tasks included in the estimation of E-E, and E-I/I-E trips: (1) estimate percentage of E-E and E-I/I-E trips in the total traffic volume for each external station; and (2) estimate trip distribution of E-E between any two external stations.

Estimate E-E/E-I Trip Split

The percentage of E-E trips at an external station is calculated based on the characteristics of this external station, including the functional classification of the highway where this external station is located, the average daily traffic volume at this station, the population of the study area, the connectivity to other external stations, and the vehicle composition at this external station. Equation A1 is used to estimate the percent of E-E trips at each external station; results are displayed in Table A3.

$$TTY_i = 76.76 + 11.22 \times I - 25.74 \times PA - 42.18 \times MA + 0.00012 \times ADT_i \\ + 0.59 \times PTKS_i - 0.48 \times PPS_i - 0.000417 \times POP \tag{A1}$$

where

TTY_i : Percentage of E-E trips at external station i

I : Interstate (0 or 1)

PA : Principal arterial (0 or 1)

MA : Minor arterial (0 or 1)

ADT_i : Average daily traffic at external station i

$PTKS_i$: Percentage of trucks excluding vans and pickups at external station i

PPS_i : Percentage of vans and pickups at external station i

POP : Population of study area

The number of E-E trips at an external station can be obtained by multiplying the percentage of E-E trips with the average daily traffic (ADT) at this station. The E-I and I-E trips are then estimated by simply subtracting E-E trips from the average daily traffic for each external station. The E-E and E-I trips estimated for each external station are shown in Table A3.

Table A3: Through Trips and E-I Trips

External Station	ADT	Road Description	Percentage of E-E trips	E-E Trips	E-I/I-E Trips
81	871	KY 1685 W	20	174	697
82	4990	US 421 W	26	1297	3693
83	34900	I 64 W	93	32457	2443
84	853	KY 1681 W	20	171	682
85	15500	US 60 W	42	6510	8990
86	835	KY 1681 W	20	167	668
87	250	KY 1659 W	20	50	200
88	4130	US 62 W	25	1033	3098
89	17300	TR 9002 W	75	12975	4325
90	1080	KY 33 E	20	216	864
91	506	KY 1267 E	20	101	405
92	3100	KY 169 E	23	713	2387
93	783	CR 1108 E	20	157	626
94	135	CR 1107 E	20	27	108
95	1350	KY 1966 E	20	270	1080
96	45800	US 60 E	41	18778	27022
97	129	CR 1004 E	20	26	103
98	2420	KY 1681 E	24	581	1839
99	1319	CR 1013 E	20	264	1055
100	7770	US 62 E	28	2176	5594
101	30000	I 64 E	92	27600	2400
102	1641	KY 341 E	20	328	1313

Estimate E-E Trip Distribution

After the number of E-E trips at each external station is determined, the distribution of E-E trips between each two external stations is then estimated. Trip exchange between two external stations is highly associated with the functional classification of the highways connecting these two external stations, the percentage of E-E trips at the destination station, route continuity between origin and destination, and the average daily traffic at the destination station. The functional classification of the highway connecting to the destination station determines which formula will be used. For example, percent distribution of E-E trips from origin station i to destination station j is estimated using the following equation (A2) if the highway at destination zone j is classified as principal arterial.

$$TTY_{ij} = -7.40 + 0.55 \times PTTDES_j + 24.68 \times RTECON_{ij} + 45.62 \times \frac{ADT_j}{\sum_{j=1}^{n} ADT_j} \qquad \text{A2}$$

where

TTY_{ij} : Percentage distribution of E-E trips ends from origin station i to destination station j; or the percentage of E-E trips at origin station i that will end at destination j

$PTTDES_j$: Percentage of E-E trips in the traffic flow of destination station j

$RTECON_{ij}$: Route continuity between i and j : Yes =1 and No=0

ADT_j : Average daily traffic at the destination station j

After the percentage distribution of E-E trips is obtained by equation A2, the number of trip exchanges between each two external station can be calculated by multiplying E-E trips with percentage distribution. This will create the OD matrix associated with external trip distribution. However, the above equation does not guarantee the number of trips from i to j is equal to the number of trips from j to i. Since this travel demand model is established to reflect average daily travel, it is reasonably assumed that the OD matrix should be symmetrical. The next step is to produce a symmetrical OD matrix by averaging ij value and ji value.

In this new symmetrical OD matrix, it is very possible that row totals and column totals are not equal to E-E trips. The recommend solution is to apply the Fratar technique to adjust the OD matrix so that the total and column totals are consistent with E-E trips estimated in Table A3 (Martin et al, 1998; Caliper Corporation, 2004).

Estimate I-E/E-I Trips

I-E/E-I trips are related to the households and employments inside the study area. The trips generated in internal zones by the households and employments are categorized by trip purpose such as HBW, HBO, and NHB. The I-E/E-I trips have to be analyzed by these trip purposes in order to be consistent with the trips generated in internal zones. NCHRP Report 365 suggests a breakdown of I-E/E-I trips by trip purpose. Table A4 lists in the first row the percentages of trip production and attraction in E-I/I-E trips by trip purpose. Therefore, the number of trips by purpose and by production/attraction at each external station can be estimated, as shown in Table A4.

Table A4: Trip Production and Attraction of External Stations

NCHRP 365 Recommendations		0.10	0.15	0.23	0.27	0.17	0.08
STATION	I-E/E-I	HBW_P	HBW_A	HBO_P	HBO_A	NHB_P	NHB_A
1	697	70	105	160	188	118	56
2	3693	369	554	849	997	628	295
83	2443	244	366	562	660	415	195
84	682	68	102	157	184	116	55
85	8990	899	1349	2068	2427	1528	719
86	668	67	100	154	180	114	53
87	200	20	30	46	54	34	16
88	3098	310	465	713	836	527	248
89	4325	433	649	995	1168	735	346
90	864	86	130	199	233	147	69
91	405	41	61	93	109	69	32
92	2387	239	358	549	644	406	191
93	626	63	94	144	169	106	50
94	108	11	16	25	29	18	9
95	1080	108	162	248	292	184	86
96	27022	2702	4053	6215	7296	4594	2162
97	103	10	15	24	28	18	8
98	1839	184	276	423	497	313	147
99	1055	106	158	243	285	179	84
100	5594	559	839	1287	1510	951	448
101	2400	240	360	552	648	408	192
102	1313	131	197	302	355	223	105

Trip Balancing

Since trip production and attraction are estimated independently, there is no guarantee that the area-wide total production and attraction have the same numerical value as discussed in section 3.2.1. Balancing is needed to ensure these two values are the same for the study area. Since trip production estimate is usually more accurate than attraction because of more reliable data, the equation A3 recommended by NCHRP Report 365 is used.

$$CT_p = \sum P_i + \sum P_e - \sum A_e \qquad \text{A3}$$

where

CT_p : The control total of trip production

P_i : Trip production at each internal TAZ

P_e : Trip production at each external station

A_e : Trip attraction at each external station

Then, the balancing factor for trip attraction is computed as

$$Factor = \frac{CT_p}{\sum A_i}$$

where

A_i : Trip attraction at each internal TAZ

For each internal TAZ, trip attraction is then multiplied by the balancing factor according to trip purpose to obtain the balanced trip attraction.

Trip Distribution

Trip distribution is the second major step in the traditional four-step travel demand model. The trip distribution model estimates the number of trips between each two TAZs. The E-E trip distribution is calculated as described in the previous section.

The I-I, E-I, and I-E trip distributions are estimated using gravity model embedded with gamma function form. The gravity model for transportation planning is based on the Newton's law of gravitation, which states that trips between two zones are directly proportional to the number of trips generated in these two zones and inversely proportional to a function of spatial separation of these two zones (travel impedance between them). The gravity model can be described as equation A4.

$$t_{ij} = a_i \times P_i \times b_j \times A_j \times f(c_{ij}) \qquad \text{A4}$$

where

t_{ij}: Trips between origin i and destination j

P_i: Trips produced by zone i

A_j: Trips attracted to zone j

a_i: The balancing factor for row i

b_j: The balancing factor for column j

$f(c_{ij})$: The impedance function between zone i and zone j

c_{ij}: The impedance between zone i and zone j

For the Woodford model, the impedance is first measured by free-flow travel time between two zones. Once the traffic assignment is performed, the congested travel time will replace the free-flow travel time to re-run the model. The travel impedance function uses gamma function, which is often recommended in U.S. planning practice because it fits well with observed travel behavior (Caliper Corporation, 2004). The function is defined mathematically as equation A5:

$$f(c_{ij}) = \alpha \times c_{ij}^{-\varpi} \times e^{-\xi \times c_{ij}} \qquad \text{A5}$$

where

α, ϖ and ξ: derived parameters

The default parameters are used in this study as listed in Table A5. Note that parameter α is a scaling factor and can be omitted.

Table A5: Parameters for Impedance Function

Trip Purpose	ϖ	ξ
HBW	0.020	0.123
HBO	1.285	0.094
NHB	1.332	0.100

The trips from gravity model are measured by person trips. Auto occupancy factors shown in Table A6 are needed to convert person trips to vehicle trips.

Table A6: Auto Occupancy Factors

Trip Purpose	Auto Occupancy Factor
HBW	1.11
HBO	1.59
NHB	1.66

An overall trip matrix can then be created by combining all OD matrices for all trip purposes.

Traffic Assignment

Traffic assignment is the last major step in the traditional four-step travel demand model. It is the process of assigning interzonal trips to the physical roadway network. A detailed transportation network specification (such as node-link-path definition, capacity, and speed limit) is needed in addition to the OD matrix. The user equilibrium assignment method is performed based on the assumption that travelers are aware of the travel times (or costs) on all paths connecting their origins and destinations and they always choose the path that minimizes their individual travel time (or cost). A BPR function discussed in Chapter 4 is used to describe how travel time on a link varies with the traffic flow on this link.

Through solving the equilibrium assignment problem, the link flow pattern (and consequently travel time on each link) can be found. There is no guarantee that the model output (i.e., link flow) matches the measured traffic flow on the network after the first run of the assignment. Therefore, model calibration is needed.

Model Calibration

The main goal of calibration is to fit the estimated traffic volume with the observed traffic volume by adjusting some parameters. The observed traffic volumes on roadways in Woodford County come from two sources: HIS extract for all count stations, and additional counts on segments of US 60, US 60X, US62 and KY33 collected by a KYTC transportation study. A total of 110 road segments with traffic counts are used in the calibration process. The distribution of these traffic count stations is shown in Figure A2. The locations where KYTC collected additional counts are shown in Figure A3. Detailed counts from both sources are shown in the Appendix 3-B, along with estimated traffic volume.

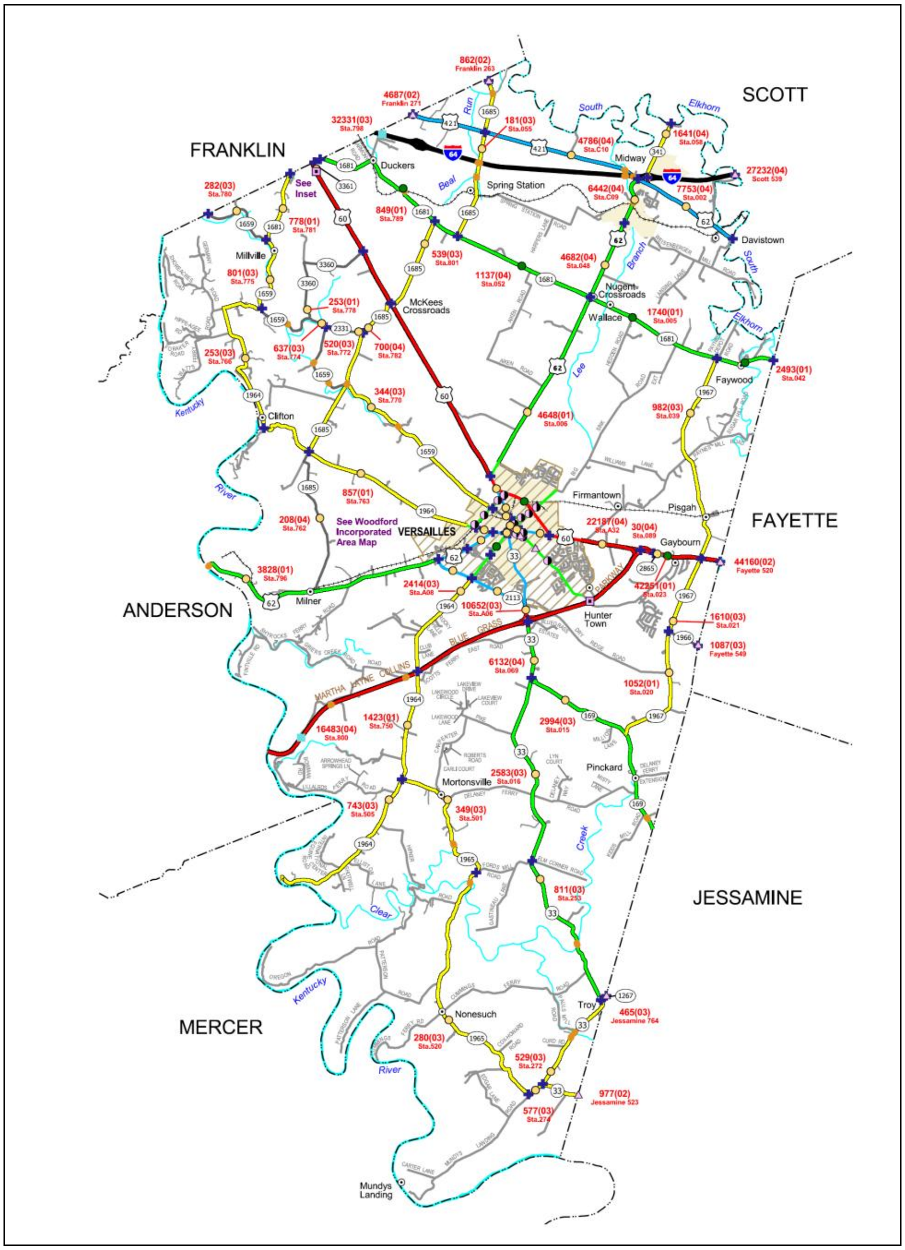

Figure A2: Locations of Traffic Count Stations

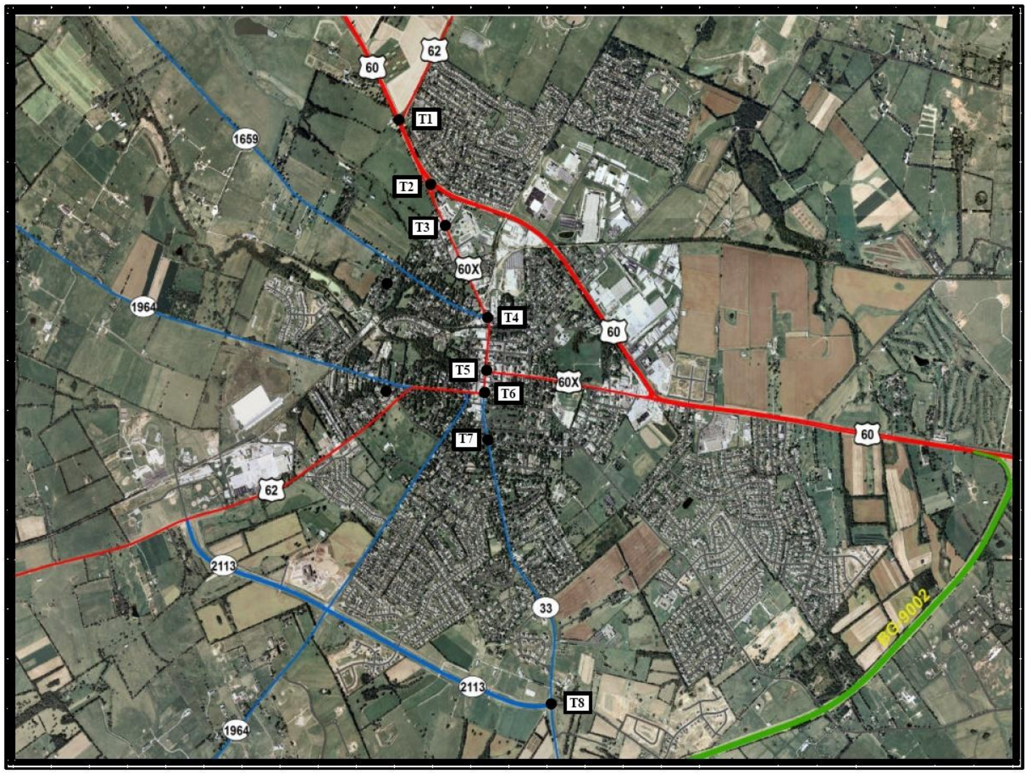

Figure A3: Locations of Additional Traffic Counts Collected by KYTC

Percent Root Mean Squared Error (PRMSE), as defined in equation A6, is used to measure the difference between the observed and estimated link traffic volumes.

$$PRMSE = \frac{\sqrt{\sum_{i}^{N} (\hat{v}_n - v_n)^2 / N}}{\sum_{n}^{N} v_n / N} \qquad \text{A6}$$

where

$\hat{v}_n$: Estimated traffic volume on traffic count station n

v_n : Observed traffic volume on traffic count station n

N : Total number of traffic count stations in the area

The acceptable PRMSE in common practice is under 30% (Kentucky Transportation Cabinet, 2004). The first run of the Woodford model produced a PRMSE of 44.7%. Therefore, some adjustments have been made. Those segments with a significant difference between the assigned and observed volumes are first identified. The following adjustments have been made in this traditional travel demand model.

The E-E trips are adjusted by increasing more through traffic percentage along the road whose assigned traffic volume is lower than observed. For example, the initial estimate of through trips on I-64 is too low; the through trip percentage on I-64 is increased.

Centroid connectors were moved to better represent the load point(s) of trips on the road network. For example, the centroid connector in TAZ 22 is first connected to where is far away from US60X, which leads to under-assignment of US60X because of higher travel time to it; the centroid connector in TAZ 22 that loads traffic to the network is revised close to US60X, which partially helps correct the under-assignment on US60X.

Free-flow speed for the local roads and/or roads in other lower functional classification is adjusted to avoid the inaccuracy caused by the default free-flow speed for local roads (i.e., 25 mph). For instance, the free-flow speed of CR1100 is originally set to 25 mph according to the facility type; it is then adjusted to 41 mph after consulting with local planners and referencing the HERS speed model of Kentucky State.

Turn penalties were revised to better represent the actual maneuver. For example, the turn penalties in the intersection of US60 with US62 were imposed to properly represent the traffic maneuver. A number of model runs have been performed in the calibration process, with a final PRMSE of 25.7%. The comparison between the assigned and observed volumes on each segment with counts is shown in the Appendix 3-B.

Appendix 3-B: Comparison between Assigned and Observed Volumes

ADT STATION	KYTC LOCATION	ROUTE	OBSERVED VOLUME	ESTIMATED VOLUME	DIFFERENCE
034520		120 US-60	22900	22901	1
034520		120 US-60	22900	22901	1
034547		120 KY-1966	1350	1351	1
037263		120 KY-1685	871	871	0
037271		120 US-421	4990	4991	1
057523		120 KY-33	1080	1079	-1
057764		120 KY-1267	506	506	0
105539		120 I-64	15000	15000	0
105539		120 I-64	15000	15000	0
120002		120 US-62	7770	7771	1
120005		120 KY-1681	1850	5665	3815
120006		120 US-62	5180	6706	1526
120015		120 KY-169	3100	3099	-1
120016		120 KY-33	2840	2678	-162
120020		120 KY-1967	1130	2047	917
120021		120 KY-1967	1690	2909	1219
120023		120 US-60	23550	20451	-3099
120023		120 US-60	23550	20274	-3276
120039		120 KY-1967	1319	1320	1
120042		120 KY-1681	2420	2420	0
120048		120 US-62	4730	8287	3557
120052		120 KY-1681	1170	1636	466
120055		120 KY-1685	189	2202	2013
120058		120 KY-341	1640	1640	0
120069		120 KY-33	6520	7302	782
120253		120 KY-33	854	2920	2066
120272		120 KY-33	545	1710	1165
120274		120 KY-1965	638	1160	522
120501		120 KY-1965	402	2118	1716
120505		120 KY-1964	822	0	-822
120520		120 KY-1965	304	207	-97
120750		120 KY-1964	1670	2987	1317
120762		120 KY-1685	204	681	477
120763		120 KY-1964	959	1448	489
120766		120 KY-1964	351	925	574
120770		120 KY-1659	354	386	32
120772		120 KY-2331	592	1359	767
120774		120 KY-1659	628	1767	1139
120775		120 KY-1659	873	1098	225
120778		120 KY-3360	258	257	-1
120780		120 KY-1659	250	250	0

ADT STATION	KYTC LOCATION	ROUTE	OBSERVED VOLUME	ESTIMATED VOLUME	DIFFERENCE
120781		120 KY-1681	835	835	0
120782		120 KY-1685	728	1882	1154
120789		120 KY-1681	853	1187	334
120796		120 US-62	4130	4130	0
120798		120 I-64	17450	17450	0
120798		120 I-64	17450	17450	0
120800		120 BG-9002	8650	8651	1
120800		120 BG-9002	8650	8651	1
120801		120 KY-1685	548	3054	2506
120A01	NT5_60X	120 US-60X	12000	10245	-1755
120A02		120 US-60X	7410	2106	-5304
120A05		120 US-62	8170	7119	-1051
120A06	ST8_KY33	120 KY-33	10000	8672	-1328
120A07		120 CS-1158	5080	3917	-1163
120A08		120 KY-1964	2880	3521	641
120A09		120 KY-1964	1870	2006	136
120A11		120 US-60	10950	10667	-283
120A11		120 US-60	10950	9742	-1208
120A12		120 CR-1028	7300	9133	1833
120A15		120 US-60X	10500	10528	28
120A23		120 CS-1023	6340	5641	-699
120A25		120 CS-1068	2830	201	-2629
120A27		120 KY-1964	4490	6441	1951
120A32		120 US-60	11750	11550	-200
120A32		120 US-60	11750	12071	321
120A37	ST2_US60	120 US-60X	7200	3411	-3789
120A38	WT3_CS1038	120 CS-1038	4000	4081	81
120A40		120 CS-1044	2620	1785	-835
120A41		120 CS-1045	2570	2358	-212
120A42	WT4_KY1659	120 KY-1659	2500	2874	374
120A43	NT4_US60X	120 US-60X	9220	7953	-1267
120A45		120 CS-1027	4350	1030	-3320
120A47		120 CS-1058	6330	4903	-1427
120A48		120 CS-1148	8590	5420	-3170
120A49		120 CS-1061	11200	7411	-3789
120A50	WT6_62	120 US-62	16000	16468	468
120A53		120 CS-1146	4890	450	-4440
120A55	ET3_CS1041	120 CS-1041	5000	2362	-2638
120C09		120 US-62	6460	6242	-218
120C10		120 US-421	4800	3135	-1665
120C11	WT8_KY2113	120 KY-2113	2730	1575	-1155
120C11	WT8_KY2113	120 KY-2113	2730	1321	-1409
120C12		120 KY-2113	1935	958	-977
120C12		120 KY-2113	1935	958	-977
120ES_1107		120 CR-1107	135	134	-1
120ES_1108		120 CR-1108	783	783	0
120P53		120 BG-9002	9700	8776	-925

ADT STATION	KYTC LOCATION	ROUTE	OBSERVED VOLUME	ESTIMATED VOLUME	DIFFERENCE
120P53		120 BG-9002	9700	8431	-1269
120P60		120 US-60	7700	7750	50
120P60		120 US-60	7700	7750	50
120ES_1004		120 CR-1004	129	127	-2
	NT8_KY33	120 KY-33	10000	7720	-2280
	ST7_KY33	120 KY-33	12000	9058	-2942
	ST6_KY33	120 KY-33	12000	9933	-2067
	NT1_US60	120 US-60	8000	9077	1077
	ST1_US60	120 US-60	11500	12139	639
	ET2_US60	120 US-60	8000	8728	728
	NT1_US60	120 US-60-10	8000	9043	1043
	ST1_US60	120 US-60-10	11500	12509	1009
	ET2_US60	120 US-60-10	8000	9273	1273
	NT3_US60X	120 US-60X	8000	7172	-828
	ST3_US60X	120 US-60X	8500	7729	-771
	ST4_US60X	120 US-60X	12000	10376	-1624
	ET5_US60X	120 US-60X	9000	6525	-2475
	ST5_US62	20 US-62	14000	13794	-206
	ET1_US62	120 US-62	9000	7083	-1917
	ET6_CS1061	120 CS-1061	10000	7644	-2356
	ET4_CS1027	120 CS-1027	3000	1376	-1624
	ET7_CS1070	120 CS-1070	5500	2978	-2522

Appendix 5-A: Household Distribution Results from the ITLUP model

TAZ	Total	Assumption	ITLUP-Observed	ITLUP-Modeled	TAZ	Total	Assumption	ITLUP-Observed	ITLUP-Modeled
1	7	4	3	59	40	37	21	16	19
2	22	13	9	57	41	3	2	1	16
3	42	24	18	58	42	32	18	14	20
4	34	19	15	59	43	4	2	2	29
5	173	99	74	60	44	37	21	16	24
6	105	60	45	57	45	118	67	51	28
7	49	28	21	58	46	123	70	53	30
8	83	47	36	56	47	49	28	21	34
9	113	65	48	62	48	111	63	48	29
10	294	168	126	63	49	200	115	85	25
11	65	37	28	61	50	273	156	117	18
12	143	81	62	58	51	60	34	26	21
13	61	35	26	64	52	52	30	22	29
14	124	71	53	61	53	46	26	20	353
15	86	49	37	50	54	8	4	4	22
16	0	0	0	54	55	12	6	6	49
17	0	0	0	43	56	549	314	235	40
18	0	0	0	43	57	43	25	18	20
19	43	25	18	70	58	52	30	22	31
20	1	1	0	89	59	16	9	7	45
21	3	2	1	62	60	7	4	3	17
22	314	179	135	61	61	82	47	35	54
23	329	188	141	55	62	65	37	28	24
24	271	155	116	37	63	57	33	24	24
25	115	66	49	31	64	206	118	88	18
26	319	182	137	55	65	111	63	48	15
27	4	2	2	48	66	52	30	22	14
28	2	1	1	26	67	40	23	17	19
29	223	127	96	86	68	29	16	13	14
30	38	22	16	95	69	9	5	4	184
31	10	6	4	30	70	23	13	10	144
32	512	293	219	30	71	2	1	1	98
33	603	345	258	219	72	118	67	51	111
34	24	13	11	50	73	194	111	83	25
35	24	13	11	130	74	142	81	61	23
36	802	459	343	33	75	261	149	112	18
37	166	95	71	23	76	158	90	68	27
38	461	263	198	27	77	77	44	33	23
39	127	73	54	22	78	130	75	55	21
R^2 = 0.56; α_h = 1.37e-07; β_h = 0.000298; c_h = -0.86; d_h = -0.01									

Appendix 5-B: Numerical Value of Variables

TAZ	Y_1	Y_2	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
1	12.48	340.54	0.00	0.00	0.00	0.36	0.00	0.00	0.09	0.55	0.39	0.04	0.53	20.25	23.42	44.33	34.15
2	62.82	173.49	0.00	0.12	0.00	0.30	0.00	0.00	0.00	0.58	0.52	0.14	0.54	20.34	23.67	44.75	33.46
3	65.05	152.87	0.02	0.48	0.04	0.27	0.00	0.00	0.00	0.19	0.68	0.10	0.53	20.30	23.84	44.23	33.07
4	48.36	55.68	0.00	0.63	0.02	0.06	0.00	0.00	0.00	0.29	0.49	0.59	0.88	20.53	23.75	42.58	31.94
5	53.70	5.24	0.01	0.57	0.02	0.01	0.05	0.00	0.32	0.03	0.40	0.43	1.69	20.56	23.53	40.49	28.27
6	104.42	6.27	0.15	0.54	0.08	0.12	0.00	0.00	0.07	0.05	0.67	0.17	0.58	20.38	23.71	44.89	32.46
7	11.28	102.96	0.00	0.15	0.07	0.11	0.12	0.38	0.12	0.06	0.63	0.32	1.38	20.17	23.39	39.70	31.43
8	51.62	96.71	0.29	0.66	0.00	0.03	0.00	0.00	0.01	0.00	0.41	0.24	1.22	20.37	24.05	40.11	30.57
9	52.12	11.20	0.03	0.54	0.01	0.00	0.00	0.00	0.02	0.41	0.47	0.20	0.85	20.36	23.69	42.90	31.30
10	79.85	73.82	0.15	0.73	0.01	0.03	0.00	0.00	0.03	0.05	0.46	0.62	1.40	21.09	24.26	38.13	28.28
11	44.05	22.74	0.06	0.93	0.01	0.00	0.00	0.00	0.00	0.00	0.16	0.12	1.01	20.83	24.00	42.84	31.33
12	28.58	204.45	0.12	0.38	0.00	0.05	0.06	0.00	0.08	0.32	0.72	0.91	2.09	21.22	23.86	32.64	26.22
13	30.60	0.00	0.00	0.65	0.00	0.00	0.02	0.00	0.29	0.05	0.28	0.68	1.67	20.80	23.44	42.19	29.09
14	64.68	99.75	0.17	0.65	0.01	0.10	0.00	0.00	0.06	0.01	0.50	0.36	1.28	20.52	22.91	41.71	30.66
15	40.45	13.32	0.27	0.06	0.00	0.48	0.02	0.00	0.00	0.17	0.70	0.16	1.56	20.40	22.84	41.57	30.08
16	0.00	55.51	0.00	0.00	0.28	0.72	0.00	0.00	0.00	0.00	0.33	0.14	1.94	20.20	22.59	41.84	30.82
17	0.00	78.42	0.00	0.00	0.00	0.47	0.36	0.00	0.00	0.18	0.57	0.20	2.39	20.09	22.91	39.70	29.41
18	0.00	0.81	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.23	2.34	20.14	23.05	39.80	29.27
19	53.38	1.30	0.11	0.55	0.15	0.00	0.00	0.00	0.00	0.19	0.65	0.27	1.96	20.12	23.32	41.01	29.65
20	0.13	81.96	0.00	0.00	0.06	0.12	0.54	0.08	0.14	0.07	0.52	0.79	2.79	20.94	24.14	30.92	25.65
21	0.17	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.47	3.36	19.68	25.25	29.99	22.20
22	43.66	77.76	0.04	0.39	0.02	0.16	0.00	0.36	0.02	0.01	0.51	0.48	1.67	20.52	24.76	35.21	27.73
23	62.19	64.79	0.04	0.92	0.00	0.01	0.00	0.00	0.00	0.04	0.20	1.28	1.98	21.41	24.70	33.82	25.71
24	84.50	23.43	0.11	0.88	0.00	0.00	0.00	0.00	0.01	0.01	0.21	1.54	2.72	22.18	25.35	30.93	23.12
25	22.67	1.46	0.00	0.21	0.00	0.01	0.00	0.50	0.28	0.00	0.20	1.47	1.97	22.04	25.21	39.34	27.73

TAZ	Y_1	Y_2	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
26	52.48	240.69	0.15	0.34	0.00	0.00	0.36	0.00	0.10	0.06	0.67	0.25	2.23	21.71	24.79	28.26	25.67
27	0.73	18.86	0.00	0.00	0.00	0.00	0.18	0.26	0.00	0.56	0.36	0.23	3.23	22.45	25.62	30.68	23.10
28	0.06	0.38	0.00	0.00	0.00	0.00	0.17	0.64	0.00	0.20	0.35	0.91	4.92	24.01	27.18	22.64	17.05
29	8.81	0.67	0.05	0.07	0.00	0.00	0.00	0.73	0.01	0.16	0.34	1.11	2.52	21.56	24.25	39.11	25.81
30	0.79	0.79	0.00	0.01	0.00	0.00	0.00	0.99	0.00	0.00	0.03	0.25	2.64	20.33	22.71	36.58	26.75
31	0.14	0.30	0.00	0.00	0.00	0.00	0.00	0.98	0.00	0.02	0.05	0.98	7.77	22.83	23.30	16.00	12.32
32	77.08	4.30	0.00	0.94	0.00	0.00	0.00	0.00	0.00	0.05	0.13	0.31	2.71	20.69	22.86	37.36	22.94
33	8.54	14.29	0.00	0.13	0.00	0.00	0.04	0.79	0.00	0.05	0.30	1.65	4.78	22.14	25.34	21.88	17.50
34	0.50	0.07	0.00	0.01	0.00	0.00	0.00	0.99	0.00	0.00	0.03	2.38	6.51	20.37	28.42	16.18	13.16
35	0.46	2.06	0.00	0.00	0.01	0.00	0.00	0.97	0.00	0.02	0.06	0.52	7.08	16.99	27.66	14.79	12.91
36	18.79	3.89	0.00	0.36	0.00	0.03	0.00	0.54	0.06	0.01	0.28	0.69	4.01	19.97	25.59	25.90	17.94
37	5.14	3.57	0.09	0.13	0.00	0.00	0.00	0.56	0.09	0.13	0.42	1.87	4.97	20.96	26.66	20.95	17.80
38	8.84	0.37	0.00	0.25	0.00	0.00	0.00	0.72	0.01	0.01	0.22	2.57	5.04	23.45	26.97	20.67	15.43
39	0.85	0.22	0.00	0.02	0.00	0.00	0.00	0.93	0.00	0.05	0.12	1.44	7.88	26.27	29.37	12.56	9.38
40	0.33	0.09	0.00	0.02	0.00	0.00	0.00	0.97	0.00	0.01	0.07	2.98	12.42	28.46	27.53	5.30	4.20
41	0.08	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	4.09	6.41	23.86	24.88	17.56	12.31
42	0.52	0.12	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.02	6.23	9.43	26.30	26.22	9.80	7.02
43	0.05	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	3.73	12.86	26.99	23.43	5.68	4.14
44	0.15	0.18	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	5.70	17.95	24.94	22.94	2.90	2.44
45	0.60	1.18	0.00	0.01	0.00	0.00	0.00	0.94	0.00	0.05	0.11	3.96	19.71	26.84	23.62	2.58	2.60
46	54.26	21.33	0.01	0.81	0.01	0.00	0.00	0.00	0.05	0.12	0.28	0.20	14.12	21.19	21.06	7.76	7.04
47	22.26	9.47	0.00	0.51	0.00	0.03	0.09	0.09	0.00	0.28	0.57	0.24	14.78	21.67	21.26	7.04	6.91
48	25.34	0.24	0.00	0.48	0.00	0.01	0.00	0.51	0.00	0.00	0.24	0.41	14.99	21.90	21.54	6.94	6.24
49	33.49	40.75	0.03	0.58	0.06	0.13	0.05	0.00	0.06	0.08	0.68	0.35	14.09	21.23	21.25	6.29	6.36
50	1.09	0.65	0.00	0.02	0.00	0.00	0.00	0.96	0.00	0.01	0.08	2.14	15.18	23.15	24.12	4.70	3.97
51	0.45	0.13	0.00	0.03	0.00	0.00	0.00	0.98	0.00	0.00	0.05	1.99	10.19	22.69	25.28	10.30	4.85
52	0.48	1.28	0.00	0.03	0.00	0.00	0.00	0.96	0.00	0.01	0.08	5.92	14.38	20.69	26.68	4.24	3.43
53	0.41	0.77	0.00	0.01	0.00	0.00	0.00	0.99	0.00	0.01	0.04	6.09	13.93	19.32	29.89	4.14	3.60
54	0.31	0.31	0.00	0.01	0.00	0.00	0.00	0.99	0.00	0.00	0.03	4.53	9.76	18.35	28.60	9.11	7.45
55	1.02	0.00	0.00	0.01	0.00	0.00	0.00	0.95	0.00	0.04	0.10	1.07	7.35	15.67	27.14	14.64	12.61
56	3.62	1.21	0.00	0.14	0.00	0.00	0.00	0.81	0.01	0.04	0.23	2.41	8.69	18.72	29.40	10.59	8.26

TAZ	Y_1	Y_2	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
57	0.61	0.03	0.00	0.03	0.00	0.00	0.00	0.96	0.00	0.00	0.07	6.11	13.25	20.16	31.50	4.63	4.55
58	0.47	0.36	0.00	0.04	0.00	0.00	0.00	0.95	0.00	0.00	0.08	1.78	6.85	23.80	28.48	13.93	11.71
59	0.70	2.46	0.00	0.12	0.00	0.00	0.00	0.88	0.00	0.00	0.14	4.06	6.83	24.87	28.04	14.15	11.87
60	0.70	0.10	0.00	0.05	0.00	0.00	0.00	0.93	0.02	0.00	0.09	6.56	11.91	28.87	32.04	5.37	4.69
61	0.58	0.14	0.00	0.04	0.00	0.00	0.00	0.94	0.02	0.01	0.09	4.10	13.18	31.10	32.97	4.44	3.45
62	0.38	0.07	0.00	0.06	0.00	0.00	0.00	0.84	0.01	0.10	0.22	6.44	14.48	29.28	27.46	3.80	2.79
63	0.39	0.40	0.00	0.01	0.00	0.00	0.00	0.85	0.00	0.14	0.18	9.43	15.91	30.17	28.36	2.99	2.25
64	1.07	1.18	0.00	0.06	0.00	0.00	0.00	0.80	0.03	0.11	0.23	7.56	18.07	28.66	24.72	2.17	1.68
65	0.56	0.31	0.00	0.02	0.00	0.00	0.00	0.87	0.01	0.11	0.19	3.83	15.09	28.58	24.03	4.09	2.86
66	0.49	2.70	0.00	0.01	0.00	0.00	0.00	0.50	0.00	0.49	0.24	1.16	10.95	26.17	21.62	7.54	5.82
67	0.44	0.10	0.00	0.02	0.00	0.00	0.00	0.98	0.00	0.01	0.06	6.64	16.02	25.79	20.17	3.33	2.52
68	0.48	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.02	0.60	19.15	26.40	21.19	2.51	2.28
69	0.13	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.01	0.00	0.00	1.27	19.30	26.93	21.72	2.44	2.28
70	0.35	0.90	0.00	0.01	0.00	0.00	0.00	0.95	0.00	0.05	0.10	0.72	16.68	23.36	21.75	4.73	5.04
71	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.83	0.17	0.00	0.02	2.69	19.80	25.34	23.66	2.70	2.79
72	0.94	0.04	0.00	0.15	0.00	0.00	0.00	0.82	0.01	0.02	0.20	6.50	12.16	23.68	32.18	5.08	4.42
73	0.87	0.56	0.00	0.07	0.00	0.00	0.00	0.92	0.01	0.01	0.13	2.83	8.09	24.51	29.19	10.78	9.07
74	0.68	0.29	0.00	0.07	0.00	0.00	0.00	0.92	0.01	0.00	0.12	9.56	15.28	30.55	33.72	2.81	2.57
75	0.68	0.21	0.00	0.04	0.00	0.00	0.00	0.94	0.01	0.01	0.10	11.23	18.16	31.54	34.71	1.76	1.69
76	0.41	0.17	0.00	0.02	0.00	0.00	0.00	0.94	0.01	0.03	0.10	14.20	21.89	34.08	38.76	0.83	0.89
77	0.40	0.50	0.00	0.02	0.00	0.00	0.00	0.95	0.00	0.03	0.10	11.38	16.71	26.28	35.65	2.12	2.19
78	0.46	0.12	0.00	0.02	0.00	0.00	0.00	0.94	0.00	0.04	0.11	18.92	26.07	35.81	40.49	0.41	0.48

REFERENCES

Abraham, J. E. and J. D. Hunt (1999). Policy Analysis Using the Sacramento MEPLAN Land Use-Transportation Interaction Model, Journal of Transportation Research Record, Vol. 1685, p.199-208.

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent.* Harvard University Press, Cambridge, MA.

Beckmann, M., C. B. McGuire and C. B. Winston (1956). *Studies in the Economics of Transportation.* Yale University Press New Haven.

Boyce, D. E. (1978). Equilibrium Solutions to Combined Urban Residential Location, Model Choice and Trip Assignment Models. *Competition among Small Regions*, p.246-264, Baden-Baden: Nomes Verlagsgesellschaft.

Boyce, D. E. and F. Southworth (1979). Quasi-Dynamic Urban-Location Models with Endogenously Determined Travel Costs. Environment and Planning A, Vol.11, p.575-584.

Boyce, D. E. (1980). A Framework for Constructing Network Equilibrium Models of Urban Location. Transportation Science, Vol.14, No.1, p.77-96.

Boyce, D. E., K. S. Chon., Y. J. Lee., K. T. Lin and L. J. Leblanc (1983). Implementation and Computational Issues for Combined Models of Location, Destination, Mode, and Route Choice. Environment and Planning A, Vol.15, p.1219-1230.

Boyce, D. E. (1986). Integration of Supply and Demand Models in Transportation and Location: Problem Formulation and Research Questions. Environment and Planning A, Vol.18, p.485-489.

Boyce, D. E., L. J. Leblanc and K. S. Chon (1988). Network Equilibrium Models of Urban Location and Travel Choices: A Retrospective Survey. Journal of Regional Science, Vol.28, No.2, p.159-183.

Caliper Corporation (2004). Travel Demand Modeling with TransCAD 4.7, Caliper Corporation, Newton, Massachusetts.

Cambridge Systematics, Inc. (2001). Land Use and Travel Demand Forecasting Models: Review of the Literature and Operational Models, Final Report, Puget Sound Regional Council, Seattle, WA.

Chang, J. S. and R. L. Mackett (2006). A Bi-level Model of the Relationship between Transport and Residential Location, Transportation Research B, p.123-146.

Chen, M. and D. H. Bernstein (2004). Solving the Toll Design Problem with Multiple User Groups. Transportation Research B, p. 61-79.

Clay, M. J. and R. A. Johnston (2006). Multivariate Uncertainty Analysis of an Integrated Land Use and Transportation Model: MEPLAN. Transportation Research D, p.191-203.

De La Barra, T., B. Perez and N. Vera (1984). TRANUS-J: Putting Large Models into Small Computers, Environment and Planning B, p.87-101.

De la Barra, T. (1989). *Integrated Land Use & Transport Modeling: Decision Chains & Hierarchies.* Cambridge University Press, New York, NY.

Domencich, T. A. and D. McFadden (1975). *Urban Travel Demand: a Behavioral Analysis.* Amsterdam/Oxford: North Holland Publishing Company.

Downs, A. (1992). *Stuck in Traffic: Coping with Peak-Hour Traffic Congestion.* The Brookings Institution and Cambridge, Washington, D. C.

Echenique, M. H., A. D. Flowerdew, J. D. Hunt, T. R. Mayo, I. J. Skidmore and D. C. Simmonds (1990). The MEPLAN Models of Bilbao, Leeds and Dortmund. Transport Reviews 10: p.309-322.

Erlander, S. and S. Scheele (1974). A Mathematical Programming Model for Bus Traffic in a Network. *Transportation and Traffic Theory*, New York: Elsevier, p.581-605.

Erlander, S. (1977). Accessibility, Entropy, and the Distribution and Assignment of Traffic. Transportation Research, 11, p.149-152.

Erlander, S., S. Nguyen and N. F. Stewart (1979). On the Calibration of the Combined Distribution-Assignment Model. Transportation Research B, p.259-267.

Erlander, S. (1980). Optimal Spatial Interaction and the Gravity Models, Lectures Notes in Economics and Mathematical System, Springer Verlag, Berlin.

Erlander, S. (1981). Entropy in Linear Programs. Mathematical Programming, 21, p.137-151.

Evan, S. P. (1976). Derivation and Analysis of Some Models for Combining Trip Distribution and Assignment. Transportation Research 10, p37-57.

Handy, S. L. (1993). Regional Versus Local Accessibility: Implications for Nonwork Travel. Transportation Research Record 1400, p. 58-66.

Hunt, J. D. and D. C. Simmonds (1993). Theory and Application of an Integrated Land-Use and Transport Modeling Framework. Environment and Planning B, Vol. 20, p.221-224.

Hunt, J. D. (1994). Calibrating the Naples Land-Use and Transport Model. Environment and Planning B, Vol. 21, p.569-290.

Hunt, J. D., D. S. Kriger and E. J. Miller (2005). Current Operational Urban Land-use-Transport Modeling Frameworks: A Review. Transport Reviews, Vol.25, No.3, p.329-376.

Jun, M. J. (1999). An Integrated Metropolitan Model Incorporating Demographic-economic, Land-use and Transport Models. Urban Studies, Vol, 36, p.1399-1408.

Kentucky Transportation Cabinet, (2004). Simpson County Travel Demand Model Report, Frankfort, Kentucky.

Kim, T. J. (1989). *Integrated Urban Systems Modeling Theory and Application*. Kluwer Academic Publishers, London.

Kim, T. J. (1990). *Advanced Transport and Spatial Systems Models: Application to Korea*. Springer-Verlag, New York

Kockelman, K. M. (1997). Travel Behavior as a Function of Accessibility, Land Use Mixing, and Land Use Balance: Evidence from the San Francisco Bay Area. Transportation Research Record 1607, p.117–125.

Kriger, K. S., E. J. Miller. and J. D. Hunt (1999). Integrated Urban Models: Improving the State-of-the-Practice. Presented at Proceedings of the Seventh TRB Conference on the Application of Transportation Planning Methods, Boston, MA.

Krishnamurthy, S. and K. M. Kockelman (2003). Propagation of Uncertainty in Transportation-Land Use Models: An Investigation of DRAM-EMPAL and UTPP Predictions in Austin, Texas. Transportation Research Record 1831, p.219-229.

Landis, J. D. and M. Zhang (1998a). The Second Generation of the California Urban Futures Model, Part 1: Model Logic and Theory. Environment and Planning B, Vol25, p.657-666.

Landis, J. D. and M. Zhang (1998b). The Second Generation of the California Urban Futures Model, Part 2: Specification and Calibration Results of the Land-Use Change Submodel. Environment and Planning B, Vol25, p.795-824.

Larsson T., Lundgren J.T., Rydergren C., Patriksson M. (2001). Most Likely Traffic Equilibrium Route Flows: Analysis and Computation. *Equilibrium Problems: Nonsmooth Optimization and Variational Inequality Methods*, p.129–159, Kluwer, Dordrecht.

Lee, D. B., Jr. (1973). Requiem for Large-Scale Models. Journal of the American Institute of Planners, Vol39, p.163-178.

Lemp, J. D., B. Zhou, K. M. Kockelman and B. M. Parmenter (2007). Visioning Vs. Modeling: Analyzing the Land Use-Transportation Futures of Urban Regions. Presented at the 86th Annual Meeting of the Transportation Research Board, Washington, D.C.

Leontief, W. (1967). *Input-Output Economics.* Oxford University Press, New York.

Lowry, I. S. (1964). *A Model of Metropolis*, RM-4035, Rand Corp., Santa Monica, CA.

Mackett, R. L. (1983). The Leeds Integrated Land-Use Transport Model (LILT). Supplementary Report SR 805. Crowthorn: Transport and Road Research Laboratory.

Marchel, F. (2005). A Trip Generation Method for Time-Dependent Large-Scale Simulation of Transport and Land-Use. Network and Spatial Economics, Vol5, p.179-192.

Martin, W., N. Mcguckin, and Barton-Aschman Associates, Inc. (1998). Travel Estimation Techniques for Urban Planning. NCHRP Report 365, Transportation Research Board, National Research Council, Washington, D.C.

Martinez, F. (1991). Transport Investments and Land Use Values Interaction: The Case of Santiago City. Proceedings of the PTRC Summer Annual Meeting, London.

Martinez, F. (1992a). Towards the 5-Stage Land-Use Transport Model. Land Use, Development and Globalization. Proceedings of the Sixth World Conference on Transport Research, Lyon. St.-Just-La-Pendue.

Martinez, F. (1992b). The Bid-Choice Land Use Model: An Integrated Economic Framework. Environment and Planning A, Vol24: p.871-885.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. *Frontiers in Econometrics*, Ed. P Zarembka, Academic Press, New York.

McFadden, D. (1978). Modeling the Choice of Residential Location. *Spatial Interaction Theory and Residential Location.* North-Holland, Amsterdam.

McFadden, D. (2001). Economic Choice. The American Economic Review, Vol 91, No.3, p.351-378.

Meng, Q., H. Yang and S. C. Wong (2000). A Combined Land-Use and Transportation Model for Work Trips. Environment and Planning B, Vol27, p.93-103.

Miyagi, T. (1989). A Combined Residential Location and Transport Network Equilibrium Model. Proceedings of the Fifth World Conference on Transport Research, Yokohama, Ventura.

Miyamoto, K and K. Kitazume (1989). A Land-Use Model Based on Random Utility/Rent-Bidding Analysis (RURBAN). Proceedings of the Fifth World Conference on Transport Research, Yokohama. Ventura.

Miller, E. J., D. S. Kriger, and J. D. Hunt (1999). Integrated Urban Models for Simulation of Transit and Land Use Policies: Guidelines for Implementation and Use. TCRP report 48. TRB, National Research Council, Washington, D.C.

Miller, E. J. (2004). Land Use-Transportation Modeling (Draft, Chapter 5). Transportation Engineering Handbook: Planning Methods and Applications. CRC Press LLC. FL.

Parsons Brinckerhoff Quade & Douglas, Inc. (1998). Land Use Impact of Transportation: A Guidebook. NCHRP report 423A. TRB, National Research Council, Washington, D.C.

Prastacos, P. (1986a). An integrated Land-Use-Transportation Model for the San Francisco Region: 1. Design and Mathematical Structure. Environment and Planning A, Vol18, p.307-332.

Prastacos, P. (1986b). An integrated Land-Use-Transportation Model for the San Francisco Region: 2. Empirical Estimation and Results. Environment and Planning A, Vol18, p.511-528.

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

Putman, S. H. (1988). Results from Implementation of Integrated Transportation and Land Use Models in Metropolitan Regions. *Network Infrastructure and the Urban Environment*, Springer, New York.

Putman, S. H. (1991). *Integrated Urban Models 2: Policy Analysis of Transportation and Land Use*. Pion, London. Britain.

Putman, S. H. (1995), EMPAL and DRAM Location and Land Use Models: A Technical Overview. Proceedings of the Land use Modeling conference, Dallas, TX.

Putman, S. H. Associates, Inc. (2001). Integrated Transportation and Land Use Forecasting: Sensitivity Tests of Alternative Model systems Configuration. TMIP. Publication FHWA-EP-01-025. U.S. Department of Transportation.

Rosner, B. (2005). *Fundamentals of Biostatistics*, sixth edition, Cengage Learning Press, Belmont, CA.

Sheffi, Y. (1992). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods,* Prentice-Hall, Englewood Cliffs, NJ.

Southworth, F. Center for Transportation Analysis, Energy Division (1995). A Technical Review of Urban Land Use--Transportation Models as Tools for Evaluating Vehicle Travel Reduction Strategies, http://ntl.bts.gov/DOCS/ornl.html. Accessed October, 2004.

Tayman, J. (1996). The Accuracy of Small-Area Population Forecasts Based on a Spatial Interaction Land-Use Modeling System. Journal of the American Planning Association, Vol. 62, p 85-98.

Vanderbei R. J. (1998). LOQO User’s Manual-Version 3.10. Statistics and Operations Research, Princeton University (Technical Report No.SOR-97-08).

Waddle, P. (2002). UrbanSim: Modeling Urban Development for Land Use. Journal of the American Planning Association, Vol68, p.297-314.

Waddell, P. and G. F. Ulfarsson (2003a). Accessibility and Agglomeration: Discrete-choice Models of Employment Location by Industry Sector. Presented at 82nd Annual Meeting of the Transportation Research Board, Washington, D.C.

Waddell, P., J. Franklin. and J. Britting (2003b). UrbanSim: Development, Application and Integration with the Wasatch Front Regional Travel Model (Technical Report CUSPA-03-01). Wasatch Front Regional Council.

Waddle, P., G. F. Ulfarsson, J. P. Franklin and J. Lobb (2007). Incorporating Land Use in Metropolitan Transportation Planning, Transportation Research A, Vol41, p.382-410.

Waltz, R. A, T. D. Plantenga, and Ziena Optimization Inc. (2004). Knitro User's Manual, downloaded from http://www.ziena.com/docs.

Wardrop, J. G. (1952). Some Theoretical Aspects of Road Traffic Research, Proceeding of the Institute of Civil Engineering, Part II, p.325-378.

Wegener, M. (1994). Operational Urban Models, State of the Art. Journal of the American Planning Association, Vol60, p.17-29.

Wilson, A. G. (1970). *Entropy in Urban and Regional Modeling*. Pion, London.

Williams, H.C.W.L. (1977). On the Formulation of Travel Demand Models and Economic Evaluation Measures of User Benefit. Environment and Planning A, Vol9, p.285-344.

Zargari, Shahriar A., A. Morteza and M. Abolfazl (2008). Combined Trip Distribution and Assignment Model: Applications to Captive Trips. Presented at the 87th Annual Meeting of the Transportation Research Board, Washington, D.C.

Zhao, Y. and K. M. Kockelman. (2004). The Random-Utility-Based Multiregional Input-Output Model: Solution Existence and Uniqueness, Transportation Research B, Vol38, p.789-807.

VITA

Meiwu An
203 Southpointe Dr Apt B, Edwardsville, IL 62025
Phone: (765)409-2696; E-mail: anmeiwu@engr.uky.edu

DATE OF BIRTH
November 30, 1976

EDUCATION
Ph.D., Civil Engineering, University of Kentucky, Anticipated: 2010
M.S., Civil Engineering, Beijing University of Technology, Beijing, P. R. China, 2002
B.S., Civil Engineering, Beijing University of Technology, Beijing, P. R. China, 1999

PROFESSIONAL EXPERIENCE
Transportation Engineer II, Pikes Peak Area Council of Governments, 2009 to Present
Transportation Planner, Bureau of Land Planning and Administration, 2002 to 2003

RESEARCH EXPERIENCE
Graduate Research Assistant, University of Kentucky, 2004 to 2008
Graduate Research Assistant, Beijing University of Technology, 1999 to 2002

REFEREED JOURNAL PUBLICATIONS
An, M. and M. Chen. Estimating Non-Motorized Travel Demand. Transportation Research Record: Journal of the Transportation Research Board, No.2002. TRB, National Research Council, Washington, DC, 2007, pp.25-28

CONFERENCE PRESENTATION
An, M. Integration of the Regression Based Land Use Model and the Combined Trip Distribution-Assignment Model. Present at the Transportation Research Board 89th Annual Meeting, Washington, DC, 2010

An, M. and M. Chen. Estimating Household and Employment Distribution. Presented at the University of Kentucky Graduate Interdisciplinary Conference, 2008

An, M. and M. Chen. Estimating Non-Motorized Travel Demand. Presented at the Transportation Research Board 86th Annual Meeting, Washington, DC, 2007

Meiwu An

12-21-2009