A Study of Alternative Land Use Forecasting Models

Final Report

Contract No. BD015-10

Submitted to

FDOT Systems Planning Office
605 Suwannee Street, MS 19
Tallahassee, FL 32399

Prepared by

Fang Zhao, Ph.D., P.E.
Associate Professor and Deputy Director
Lehman Center for Transportation Research
Florida International University
10555 W. Flagler Street
Miami, Florida 33174
Phone: (305) 348-3821
Fax: (305) 348-2802
E-mail: zhaof@fiu.edu

and

Soon Chung, Ph.D.
Research Associate
Lehman Center for Transportation Research
Florida International University
10555 W. Flagler Street
Miami, Florida 33174
Phone: (305) 348-4103
Fax: (305) 348-2802
E-mail: soon.chung@fiu.edu

April 2006

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
Final Report for BD015-10		

4. Title and Subtitle	5. Report Date
A Study of Alternative Land Use Forecasting Models	April 2006
	6. Performing Organization Code
7. Author(s) Fang Zhao and Soon Chung	8. Performing Organization Report No.
9. Performing Organization Name and Address Lehman Center for Transportation Research, Department of Civil and Environmental Engineering, Florida International University, Miami, Florida 33199	10. Work Unit No. (TRAIS) 11. Contract or Grant No. BD015-10
12. Sponsoring Agency Name and Address Research Office Florida Department of Transportation 650 Suwannee Street, MS 30 Tallahassee, Florida 32399-0450	13. Type of Report and Period Covered Final Report May 2004 – April 2006 14. Sponsoring Agency Code 99700-3596-119

15. Supplementary Notes

16. Abstract

FSUTMS requires future land use forecasts as input data to predict future travel demand and transportation needs. Given that the performance of the FSUTMS models relies heavily on the accuracy of land use forecast, there is a strong desire by planners to improve model input, especially for future forecast years. The purpose of this study is to survey the state-of-the-art and the state-of-the-practice, as well as to investigate the potential of UrbanSim as a land use model for Florida Applications. UrbanSim is a promising model for its spatial disaggregation (use of parcels to model land development), temporal disaggregation (one-year time steps), dynamics (disequilibrium model); detailed disaggregations of households and firms, and support to activity-based travel models. In this project, UrbanSim is applied to Volusia County, Florida, based on five scenarios of growth and transportation improvements. The model was validated by comparing the simulation results to the socioeconomic and demographic data adopted in the 2020 LRTP and the 2005 InfoUSA employment data. This report describes the implementation process of the land use simulation, including data collection and processing, model estimation and validation, and scenario building and testing. The most significant efforts in this project are related to data imputation and quality control, and model parameter estimation. The UrbanSim model produced reasonable results in a reasonable amount of time. It will be important to develop tools to support data imputation and processing. The modeling results also suggest that the existing "consensus building" processes adopted by many local governments may need improvements to allow community visions to be better reflected through the model.

17. Key Word	18. Distribution Statement

19. Security Classification. (of this report)	20. Security Classification. (of this page)	21. No. of Pages	22. Price
Unclassified.	Unclassified.		

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

DISCLAIMER

The contents of this report reflect the views of the authors and do not necessarily reflect the official views or policies of the Florida Department of Transportation. This report does not constitute a standard, specification, or regulation.

ACKNOWLEDGEMENTS

We would like to thank the System Planning Office and the Research Center of the Florida Department of Transportation for supporting this study. We also thank the project manager, Terry Corkery of the Systems Planning Office and Gary Kramer of the West Florida Regional Planning Council and the chair of the Land Use Committee of the Florida Model Task Force for their guidance and assistance in the project. We are grateful to Mr. Michael Neihart, Transportation Planner of the Volusia County MPO, and Ms. Sandy Hulvey of the Volusia County Property Appraiser's Office for providing the transportation and land use data used in this project. Ms. Shanshan Yang, Graduate Research Assistant of the Lehman Center for Transportation Research, provided assistance in summarizing and comparing results from UrbanSim, as well as producing the graphs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Travel demand models require socioeconomic and demographic data to predict future travel demand and transportation needs. Future socioeconomic data are estimated based on future land use forecasts, sometimes with the aid of a land use model. Currently, there are many land use models developed and applied to urban areas. Most models, however, either lack the rigor of economic theories or do not have the flexibility to allow customization to reflect the local characteristics and visions.

An extensive literature review reveals that UrbanSim, a micro-simulation land use model, is promising for the following reasons:

- It is based on theories of market economy and discrete choice behavior; therefore, it captures both the impacts of market forces, as well as individuals' choices on land development processes.
- It is spatially disaggregated by using small grid cells and parcel data to model land use. The simulation is household and job-based, hence, making it more realistic.
- It is temporally disaggregated by simulating land use changes on an annual basis.
- It models the dynamics of the land use and transportation interactions and the disequilibrium between the two systems, which is caused by the time lags before one system fully responds to changes in the other.
- It is designed for integration with a travel demand model. Its disaggregated nature also lends itself to activity-based travel demand modeling.

UrbanSim has been applied to Volusia County, Florida for the purpose of assessing the model's accuracy and investigating issues related to implementation. The UrbanSim model was run jointly with a FSUTMS/TRANPLAN model. Interfacing UrbanSim with FSUTMS requires an additional program to convert the output from one model to the input of the other, which is accomplished by developing an ArcView conversion program. Although the conversion program was not designed to provide full automation to help feedback between the two models, it simplified the data processing.

Five scenarios of land use development and transportation improvements are tested. The results are described in terms of V/C ratios, traffic volumes, accessibility, households, population, and employment. They are compared to the Volusia County adopted 2020 Long Rang Transportation Plan (LRTP). The following are the findings from this study:

- UrbanSim has been found to simulate land use changes reasonably well, although detailed analyses will help further understanding of the behavior of the UrbanSim model and reasons for the differences between UrbanSim predictions and projections by Volusia County. It is necessary to point out that differences are expected, since these projections are produced using different methods. Therefore, it is appropriate to evaluate the possible causes of the differences as well as review the assumptions.

- Many urban areas go through a "consensus building" process when allocating growth to different municipalities. In the current version of UrbanSim, such a process cannot be modeled. However, users are allowed to specify land development projects, including

location, type, intensity, and implementation schedule as part of a scenario. Thus community visions need to be put in more concrete terms of possible developments so that they may be included in UrbanSim. This also points to the possibility that the existing "consensus building" processes adopted by many local governments may need to be improved to allow community visions to be better reflected through the model.

- Consultations with local government agencies are desirable when developing model specifications and estimating model parameters. Location choice models and developer model reflect the behavior of local activities. Consultation with local agencies will help improve model performance.

- Feedback from the travel model to UrbanSim influences the land development patterns. Therefore, it will be useful to measure the sensitivity of UrbanSim to accessibility to determine the necessary frequency of the feedback. Through feedback, UrbanSim also has the potential of testing the effects of different project schedules on both land use and transportation.

- The most significant efforts in this project are related to data imputation and quality control, and to model parameter estimation. The main problems with the data include outdated or missing information in the address database (used to locate businesses to parcels), missing information on the number of housing units for multi-family dwellings (including condominiums and apartments), and missing information on properties. While these problems may be addressed separately by the government agencies that created the data, a GIS-based database tool is necessary to facilitate data compilation. This tool needs to provide simple statistics functions to allow examination of the data. ArcGIS is a more suitable GIS platform for this purpose than ArcView, the latter is limited in its database management capabilities.

- UrbanSim is designed to be integrated with life-style travel models. As a result, the output from UrbanSim, which is based on life-style household structure, needs to be summarized to support the classic travel demand models. On the other hand, it will be a natural fit to life-style models and activity-based models.

- High performance computers are required. For this study, a computer of 3.4 GHz Pentium with 2 GB of RAM is used. The computer time for a 10-year simulation is approximately three hours. The running time may vary by the study area.

- Development of an UrbanSim model requires expertise in both GIS and statistics, the latter for estimating discrete choice models. Some MPOs may not have in-house expertise and may need to rely on services provided by consultants.

- A detailed user manual on data processing is needed. The UrbanSim User manual and technical reports are provided on-line (http://www.urbansim.org/docs/), but information on data processing is inadequate. This may be because UrbanSim continues to be further improved and it is not yet a commercial product.

- A TAZ-based UrbanSim model will reduce the amount of data processing involved. The University of Washington has been in the process of developing such a model. However, TAZs are usually much larger spatial units than the current grid system used in UrbanSim (150 meters by 150 meters), so it is unclear whether model accuracy would be affected by the reduced spatial resolution.

1. INTRODUCTION

The Florida Standard Urban Transportation Modeling Structure (FSUTMS) requires socioeconomic data to predict future travel demand and transportation needs. Future socioeconomic data are estimated based on future land use forecasts, sometimes with the aid of a land use model. These forecasts are made based on an array of factors such as changes in population, demographics, structure of local economy, real estate market, land use, growth policies, environmental constraints, and so on. The accuracy of a land use forecast greatly affects that of a demand forecast model. As travel demand models continue to improve, there is a strong desire of planners to improve model input, especially for future forecast years due to the lack of reliable statistics on economy and plenty of uncertainties.

Future land use forecasts have gradually evolved from mainly a manual process to a greater reliance on land use forecast models. As analytical tools, land use models provide an understanding of the causes and consequences of land use changes and thus the functioning of the land use system. They may be used to analyze different scenarios of future land use changes and support land use planning and policy analysis.

Existing land use models range from rule-based programs that provide information and guidance on the process of allocating growth to different subareas, to sophisticated models that incorporate economic theories and market mechanisms. Currently, urban areas in Florida use different methods for land use forecasting. The Metropolitan Planning Organization of Orlando (MetroPlan) has tested the DRAM/EMPAL and concluded that the model has a number of weaknesses that prevent it from being useful. Another land use model developed under the sponsorship of the Florida Department of Transportation is the Urban Land Use Allocation Model (ULAM), which is a population/employment allocation process specifically designed for creating FSUTMS zonal inputs based on population control totals, projected by the Bureau of Economic and Business Research (BEBR) at the University of Florida for Florida's urban areas. It is a "top-down" aggregated approach based on trends and heuristic rules. While rule-based land use models are easy to understand and apply, they lack the rigor of models that are built based on economic theories and are useful for exploring the complex mechanisms of the socioeconomic forces that influence the rate and spatial pattern of land use change.

Existing land use models employ a wide range of approaches, such as spatial interaction, spatial input-output, linear programming, micro-simulation, discrete choice modeling, cellular automata, and rule-based (Waddell 2004). Spatial interaction, spatial input-output, and linear programming models were used in the early operational urban models of the 1960's and 1970's. Although developed in the 1960's, micro-simulation was not applied to urban modeling until the 1980's. The 1980's saw discrete choice models and cellular automata becoming the newest modeling approaches. In the 1990's, several land use models implemented a rule-based set of procedures to allocate population, employment, and/or land use on the GIS (Geographic Information System) platform.

UrbanSim was originally developed by Paul Waddell of the University of Washington (UrbanSim 2001 and Waddell 2004). UrbanSim simulates land use changes by considering factors such as changes in population, demographics, structure of the local economy, real estate markets, land use and growth policies, environmental constraints, and so on. It is based on a disaggregate activity-based model that considers individuals' choices on residential and employment locations, as well as the real estate market. Additionally, it is a freeware while most of the other operational land use models are either proprietary or unavailable. While UrbanSim is a new land use forecasting model that promises to overcome many of the limitations of the existing land use models and has attracted much attention as well as support from the National Science Foundation, there is a lack of experience in its implementation.

This project studies UrbanSim for its potential of applications in Florida. The objectives of this research include:

1. Understanding the state-of-the-art and state-of-the-practice of land use models;

2. Determining the data requirements and identifying application issues;

3. Investigating the need for developing additional computer programs for data processing and interfacing the FSUTMS; and

4. Identifying future research issues.

In the remainder of this report, literature on proposed and operational land use models is reviewed and summarized in Chapter 2. The emphases of the literature review are underlying theories, methodologies, and applications of the models.

Chapter 3 describes the study area, Volusia County. Selection criteria include that the study area has recent household survey data and up-to-date GIS data including parcel-level property data, and is relatively self-contained. Data collection and processing are also described in Chapter 3. Data collected include property parcels, business establishments, census, urban area boundaries, environmental and political planning boundaries, and Traffic Analysis Zones (TAZs). In Chapter 4, scenarios are developed based on alternatives of the Volusia County transportation improvement plan and assumptions of future demographic and economic growth.

Chapter 5 discusses model specification and estimation. Model parameters are estimated using an econometric software package, Limdep. Chapter 6 describes the simulation process of UrbanSim for Volusia County.

Chapter 7 and 8 present results from the simulations based on scenarios developed in Chapter 4. Chapter 7 compares results from UrbanSim based on demographic and socioeconomic data adopted in the 2020 Volusia County Long Range Transportation Plan (LRTP) and business establishments observed in 2005. Results by scenario are summarized in Chapter 8 and comparison between scenarios is also presented.

Finally, conclusions and recommendations regarding future adoption of UrbanSim in Florida are provided in Chapter 9.

2. LITERATURE REVIEW

There have been many efforts to develop new land use models or apply existing land use models during the last several decades. Reviews of existing operational models may be found in (Southworth 1995), (Wegener 1995), NCHRP 8-32 (Parsons Brinckerhoff 1998), TCRP Project H-12 report (Miller et al. 1998), U.S. EPA (2000), NCHRP 25-21 (Dowling et al. 2000), and a report by Cambridge Systematics, Inc. (2001). The Federal Highway Administration (FHWA) groups models used for land use and transportation planning based on their forecasting methodologies at the FHWA *Toolbox for Regional Policy Analysis* website (2004). Timmermans (2003) summarizes three generations of land use models. The models in the first generation are based on aggregate data and principles of gravitation and entropy-maximization. Second generation models are based on the principle of utility-maximization. More recently, models have been developed based on micro-data and activity-travel patterns. In this report, land use models and their methodologies are reviewed according to the classification proposed by Waddell (2004). If a model employs a combination of different methodologies, the model is classified based on the approach on which the model put more emphasis. For instance, UrbanSim employs a discrete choice approach and is designed using the micro-simulation technique and classified as such.

2.1 Spatial Interaction Models

This section describes four spatial interaction models: DRAM/EMPAL/ITLUP/METROPILUS, LILT, HLFM II+, and LUTRIM. Models based on the spatial interaction approach utilize the gravity theory from physics. In a gravity model of land use changes, the distribution of population and employment is a function of attractiveness and travel costs associated with places. One of the limitations of this type of models is that many behavioral factors influencing location choices are not represented. Another is that the role of real estate markets and prices are not considered. These models also tend to be limited in the degree of spatial detail (Waddell 2004).

2.1.1 DRAM/EMPAL/ITLUP/METROPILUS

The most popular models in the United States are perhaps the Disaggregate Residential Allocation Model (DRAM) and the Employment Allocation Model (EMPAL). These models expand the premises of the Lowry model with a maximum entropy formulation. The Lowry model is an urban model developed in 1964 that combines the economic base multiplier model and the gravity model. DRAM and EMPAL were developed by Putman and his colleagues in the early 1970s and have been improved over time. Putman (1995) claimed that they were the "most widely applied models" in the U.S.

DRAM is a modified version of the standard singly constrained spatial interaction model. It forecasts the number of households in a zone by household categories defined by annual income. The equation of DRAM is given below (Putman 1983):

$$N_i^n = \sum_j \left(\sum_R a_{Rn} E_j^R\right)\left[\frac{W_i^n f^n(c_{ij})}{\sum_k W_k^n f^n(c_{ij})}\right]$$

where

N_i^n = estimated number of households in zone i;
E_j^R = employment size of category R in zone j;
a_{Rn} = region-wide coefficient relating the number of type R employees to type n households;
W_i^n = attractiveness measure for zone i; and
$f^n(c_{ij})$ = cost of travel function for type n residents moving from i to j.

The attractiveness of a zone (W_i^n) is a function of the area of vacant developable land in zone i (L_i^v), the proportion of developable land in the zone that has already been developed (x_i), the area of residential land use in zone i (L_i^r), and the number of residents of type n' in zone i ($N_i^{n'}$). W_i^n is computed as shown below, where q^n, r^n, s^n, and $b_{n'}^n$ are empirically derived parameters:

$$W_i^n = (L_i^v)^{q^n}(1+x_i)^{r^n}(L_i^r)^{s^n}\prod_{n'}\left[\left((1+\frac{N_i^{n'}}{\sum_n N_i^n})\right)^{b_{n'}^n}\right]$$

DRAM allocates more households to a zone if it is accessible (i.e., taking less time to travel from other zones), has more vacant land, is dominantly residential, and has the right type of employment.

The employment of type R (E_j^R) in zone j is estimated by either the EMPAL model or other models. EMPAL forecasts the employment size of four to eight types (sectors) in relation to an attractiveness measure and the employment size of the same type from the previous modeling time period. EMPAL is also a modified, singly-constrained spatial interaction model, as shown below (Cambridge Systematics 1994, Southworth 1995):

$$E_{j,t}^R = \lambda^R[\sum_i P_{i,t-1}A_{i,t-1}^R W_{j,t-1}^R f^R(c_{ij,t})] + (1-\lambda^R)E_{j,t-1}^R$$

where
$E_{j,t}^R$ = employment of type R in zone j at time t;
$P_{i,t-1}$ = total number of households in zone i at time $t-1$;
$A_{i,t-1}^R$ = "balancing term" in zone i at time $t-1$;
$W_{j,t-1}^R$ = attractiveness of type R activity in zone j at time $t-1$;
$f^R(c_{ij,t})$ = function of cost of travel for type R activity moving from i to j at time t;
$E_{j,t-1}^R$ = employment of type R in zone j at time $t-1$; and
λ^R = empirically derived parameter.

The attractiveness of zone j at time $t - 1$, ($W^R_{j,t-1}$), and a balancing term of zone j at time $t - 1$, ($A^R_{i,t-1}$), may be estimated using the following equations:

$$W^R_{j,t-1} = (E^R_{j,t-1})^{a^R} L_j^{b^R}$$

$$A^R_{j,t-1} = [\sum_j W^R_{j,t-1} f^R(c_{ij,t})]^{-1}$$

where

L_j = total area of zone j, and

a^R , b^R = empirically derived parameters.

EMPAL allocates more employment to a zone if it has a larger area and, in the previous modeling time period, better accessibility, more households, and more employment.

The DRAM/EMPAL model had been calibrated for more than forty metropolitan regions (Cambridge Systematics 1994). These regions vary significantly in geographic and demographic size. The largest is the Los Angeles metropolitan region, with a population of more than 14.5 million; the smallest is the Colorado Springs metropolitan region, with a population of under 400,000. The number of zones or sub-areas used in previous calibration efforts has ranged from 772 for Los Angeles, to 100 for Portland, Oregon. Application sites include: Southern California; Atlanta Region; Boston, Massachusetts; Northeast Illinois; North Central Texas; Houston-Galveston, Texas area; Sacramento, California; Seattle, Washington; San Diego, California; Orange County, California; Kansas City; Orlando/Kissimmee, Florida; Phoenix, Arizona; Portland-Vancouver, Oregon; Colorado Springs, Colorado; and San Antonio, Texas.

Metroplan Orlando has used DRAM/EMPAL to project and allocate growth in the past and has encountered a number of problems. One of the problems is the difficulty of modeling comprehensive plans that govern land use policy, which are regulated through three levels of government – local, regional, and state. Additional problems include that the model does not adequately support the process to arrive at realistic growth allocation by consensus among local jurisdictions, and that transportation infrastructure impact on growth cannot be predicted easily (Canin Associates 2000).

According to the U.S. DOT (2002), DRAM has been capable of capturing more than 85% of the variation in land use in calibration. However, the EMPAL calibration has not been as successful except for service employment.

The strengths and limitations of DRAM/EMPAL are presented in Table 2.1 (U.S. EPA 2000).

Table 2.1 Strengths and Limitations of DRAM/EMPAL

Strengths
– Continues to be used by numerous metropolitan areas. – Is robust. – Has the ability to introduce constraints or other influences, particularly to account for local knowledge. – Data sources are generally available. – Is relatively easy to calibrate.
Limitations
– Focuses on statistical, aggregate choice behavior rather than on individual choice behavior. – Uses a reduced form of logit for location choice. – Little or no scope to introduce planning policies other than zoning except by specific constraints or attractiveness functions. – Absence of any mechanism for simulating the land market clearing process underlying multi-year infrastructure change. – The impact of zoning policies is not well represented. – Inability to represent monetary and non-monetary incentives to guide land-use development. – Possible underestimates of the full impact of some infrastructure improvements due to a limited number of independent variables. – Limited spatial resolution of the zones, mainly due to the unavailability of data. – Sensitivity analyses not possible. – Training and experience required to run model correctly and efficiently. – Not an off-the-shelf product and requires initial consultant involvement.

The Integrated Land Use Transportation Package (ITLUP) is the first fully operational integrated transportation and land use package (Putman 1983) and has DRAM and EMPAL as components. A third program, CALIB, produces maximum likelihood parameter estimates for DRAM and EMPAL. In the 1990's, METROPILUS was developed based on DRAM and EMPAL by combining employment and residence location and land consumption in a single comprehensive package. The structure of the individual components METROPILUS is based on logit or nested logit formulation. METROPILUS is also embedded in a Geographic Information System (GIS) environment.

2.1.2 LILT

The Leeds Integrated Land Use-Transportation modeling package (LILT) was developed by Mackett (1979, 1983) at the University of Leeds, Leeds, England. LILT combines a Lowry type location model with a four-stage aggregate travel model and a car ownership model. Based on entropy-maximizing principles, forecast future population, new housing units, and jobs are allocated to zones according to accessibility functions and the attractiveness of zones.

In the application of the model for Leeds, employment was classified into three categories of primary, secondary, and tertiary based on the degrees of accessibility. Employment in the primary category was allocated in proportion to the existing distribution of land use. For example, agricultural jobs were allocated in proportion to the amount of agricultural land and the number of jobs in each zone at the previous time point. Employment in the secondary category such as manufacturing, transportation, and communications was allocated based on the previous employment distribution and the ratio of accessibility to the supply of labor and other economic activities at the current time to the accessibility at the previous modeling time point. This implies that increased accessibility in a given zone will increase activity level in that zone, and vice versa. Tertiary category jobs were related to the population distribution and were allocated by taking into account the relative cost of travel by each mode (Webster et al. 1988).

LILT has been applied to several metropolitan regions including Dortmund, Germany; Tokyo, Japan (Mackett 1990), Harrogate, England, and Athens, Greece (Webster et al. 1988). LILT has also been applied in the studies of the Channel Tunnel and commuter rail in Hertfordshire England (Mackett and Nash 1991).

The strengths and limitations of LILT are summarized in Table 2.2.

Table 2.2 Strengths and Limitations of LILT

Strengths
– Flexible for testing policies for both land use and transportation. – Produces some aggregate tables, as well as basic output files.
Limitations
– Statistical modeling process, not based on economic theory. – Does not model the property market. – Does not consider property prices. – Does not attempt to produce any measure of benefit. – Relatively complex calibration process, involving standard maximum likelihood techniques for the non-linear functions.

2.1.3 HLFM II+

HLFM II+, which stands for Highway Land Use Forecasting Model, is a simple spatial interaction model developed by Alan Horowitz (Dowling et al. 2000). Extending the Lowry model, HLFM II+ is a full equilibrium model that considers accessibility and land availability as the key explanatory variables of location choice. It is similar to the DRAM/EMPAL model but somewhat simplified. HLFM II+ is tightly integrated with the Quick Response System (QRS), a four-step travel model[1] for highway and transit forecasting.

HLFM II+ forecasts the employment and population likely to be in each zone within an urban area based on information on the highway system, existing and proposed land use, demographics, and socioeconomics. Beginning with the location and amount of "basic industry" employment in the region, the model computes the conditional probabilities for worker resident locations and for service employment locations. Basic industries are those that choose their locations primarily based on their proximity to needed natural resources and urban infrastructure. The conditional probabilities are computed using singly constrained trip distribution equations with an exponential deterrence function, as shown below (Dowling et al. 2000):

[1] A four-step travel model includes four main modeling steps: trip generation, trip distribution, mode split, traffic/transit assignment.

$$a_{i,j} = \frac{w_i \times \exp(-\beta \times t_{i,j})}{\sum_i w_i \times \exp(-\beta \times t_{i,j})}$$

where

- $a_{i,j}$ = conditional probability that an individual working in district *j* will live in district *i*;
- w_i = attractiveness of district *i*;
- $t_{i,j}$ = travel cost, time, or disutility of travel between districts *i* and *j* (from travel model); and
- β = calibration parameter.

Three conditional probability matrices are included in HLFM II+: **A**, **B** and **H**. Matrix **A** contains the probability of a person working in district *j* and residing in district *i*. Matrix **B** stores the probability that an individual living in district *j* is served by an employee working in district *i*. Matrix **H** gives the probability that an employee working in district *j* is served by another employee working in district *i*. The attractiveness, w_i, of a district *i* is specified in terms of the net developable area when computing the residential location probabilities matrix **A**. Net developable area for the service industry is used for computing the other two location probability matrices **B** and **H**. The employment size in each district may be obtained from the following equation:

$$\mathbf{E} = (\mathbf{I} - \mathbf{GBQA} - \mathbf{HF})^{-1}\mathbf{E_B}$$

where

- **E** = vector of total employment size in each district;
- $\mathbf{E_B}$ = vector containing the basic employment in each district;
- **F** = diagonal matrix containing the ratio of service employment to all employment for each employment district, usually set uniformly to the regional average;
- **G** = diagonal matrix containing the ratio of total employment to population for each residential district, usually set uniformly to the regional average; and
- **Q** = diagonal matrix containing the ratio of population to total employment in each residential district.

F, **G**, and **Q** matrices contain information typically associated with base multipliers for a region.

Although HLFM II+ was originally designed for smaller MPOs with small budgets and staff, this model has been applied by the Indian Nations Council of Governments in Oklahoma, the Baltimore Regional Council in Maryland and the Capital District Transportation Commission in Albany, NY. Another example of its application may be found in Vancouver, Canada.

The strengths and limitations of HLFM II+ are summarized in Table 2.3 (PBSJ 1999).

Table 2.3 Strengths and Limitations of HLFM II+

Strengths
– Easy to use. – Gives a good indication of the global trends in urban development. – Substantially less data required than for DRAM/EMPAL. – Integrates with QRS, thus sensitive to the impacts of both highway and transit systems on land use, as well as the effects of traffic controls. – Low priced ($300), adequate documentation, and available for the Windows platform.
Limitations
– Not suitable for detailed zonal level land use information. – No disaggregation of households by type (e.g., by income or stage of life cycle). – Inconsistent travel impedance between the gravity model in the land use model and that in the trip distribution model. – Inadequate representation of zonal attractiveness with only vacant land as the measure of attractiveness. – Little behavioral content in the model, and does not lend itself to a wide variety of policy analyses.

2.1.4 LUTRIM

The Land Use-Transportation Interaction Model (LUTRIM), developed by William Mann (1995), addresses land use and transportation interactions. In LUTRIM, a land use model is considered as the fifth step following the traditional 4-step transportation planning process. LUTRIM may be used as a land use model, a travel forecast model, or both to measure the impacts of transportation improvements on land use.

LUTRIM calculates accessibility to jobs and households based on gravity model parameters for local trip distribution and forecasts households and basic and household-serving employment based on accessibility. LUTRIM assumes that changes in the accessibility to households will cause changes in the distribution of household-serving employment and that changes in the accessibility of employment will result in changes in the distribution of households. The model input include a previously adopted land use forecast, friction factors, and socioeconomic bias factor, or K factor, from the gravity model calibration for trip distribution and travel time matrix from travel models (PBSJ 1999).

In the NCHRP Report 8-32 (PBSJ 1999), it is pointed out that this model uses the unusual approach of calibrating the model not to historical or current data, but to a previously produced land use forecast. This means that the model is trained to reproduce the land use forecast with which the model is calibrated. This invalidates the use of LUTRIM for generating new forecasts.

The strengths and limitations of LUTRIM are given in Table 2.4 (PBSJ 1999).

Table 2.4 Strengths and Limitations of LUTRIM

Strengths
– Simple and easy to use. – Integration with travel model. – Input easy to obtain.
Limitations
– Aggregate, not based on discrete behavior. – Not based on economic theories. – Calibrate to land use forecast instead of actual land use conditions.

2.2 Spatial Input-Output Models

The spatial input-output framework was originally developed based on the input-output model representing the structure of the U.S. economy. The framework is designed to address spatial patterns of location of economic activities within regions and the movement of goods and people between zones. Monetary flows are converted to flows of goods and services by type of vehicle and to flows of commuting and shopping trips by mode. Real estate and labor markets are considered, and travel demand modeling is part of the modeling process. Spatial input-output models generate a static equilibrium solution to changes in one or more inputs (Waddell 2004). Models in this category that are reviewed include MEPLAN, TRANUS, and DELTA.

2.2.1 MEPLAN

MEPLAN is a proprietary software package developed by Marcial Echenique and Partners Ltd. in the United Kingdom. Hunt (1997) describes its framework as the interaction between two parallel markets: a land market and a transportation market. Behavior in each system is modeled as a response to price or price-like signals (including travel disutility). Each market moves towards equilibrium, but a complete equilibrium is not reached because there are time lags in the system, which are caused by the fact that building stock and transportation infrastructure cannot be changed instantaneously, as well as the fact that information exchange is not perfect.

As shown in Figure 2.1, MEPLAN consists of three main modules: land use/economic module (LUS), transportation module (TAS), and economic evaluation module (EVAL) and an interface program (FRED). The LUS models the spatial location of activities such as employment and population and produces trades between zones. TAS examines modal split, route assignment, and capacity restraint. FRED interfaces with TAS and LUS and deals with two-way interactions between these two modules. FRED estimates the number and distribution of trips or flows directly from the results of the land-use model. FRED also calculates the reverse interaction, which is how changes in transportation affect the pattern of land uses in the next time period. These modules run iteratively with a typical length of time period of five years. EVAL combines the results of LUS, TAS, and FRED and compares them with alternative plans or to a base-case scenario.

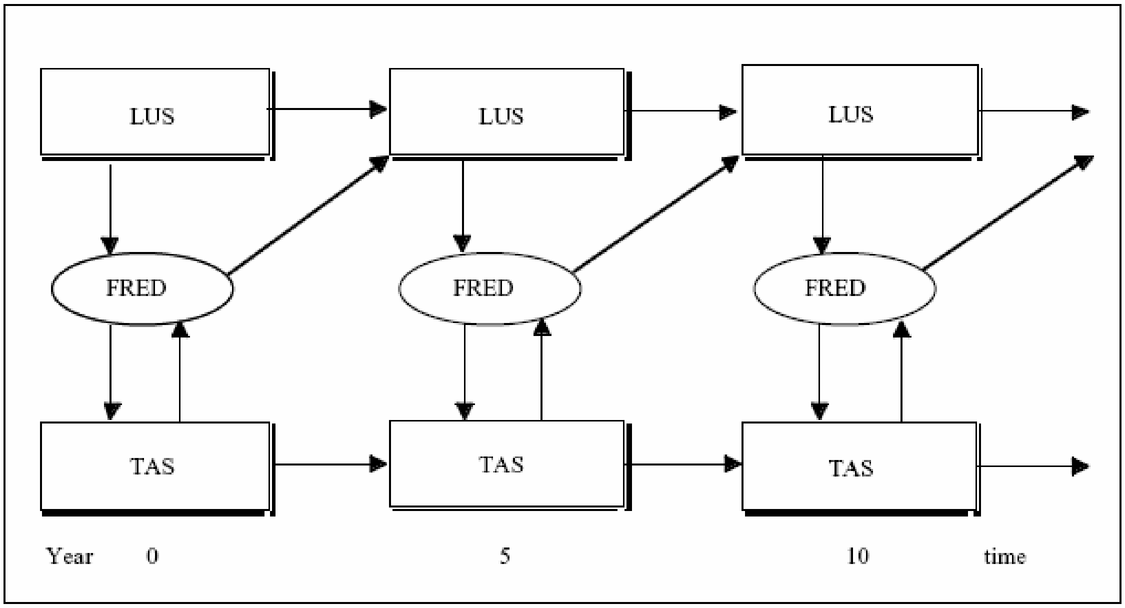


Figure 2.1 Dynamic Operation of MEPLAN (Source: Maffii and Martino 1999)

MEPLAN estimates the effects of transportation on the location choices by residents, employers, developers, and others, and determines how land use and economic activities induce travel demand. It also evaluates impacts from planning decisions on land use and transportation. MEPLAN provides the following outputs: employment by sector, population by income group, households by car ownership group, land area by activity, floor space by activity, price by floor space/land type.

Since MEPLAN is designed to flexibly meet user needs, the input required to operate the model is user-defined and may be altered for each run of the model. The input includes land use, floor space, supply and demand for land and buildings, prices for space, pattern of prices, availability of public transportation, ownership of cars, road and rail infrastructure, trip types, and other related information.

Hunt (1997) points out that while MEPLAN requires a small amount of data to run for forecasting purposes, calibrating MEPLAN is complex and demanding, as "an extremely large and rich set of observed data would be required in order to perform a 'full and complete' calibration of all model components over several periods in a typical application." Particularly, for each time period modeled, model data must be available. However, the effort may be worthwhile because of its ability to incorporate the interactions of different components in land use and transportation markets and policy alternatives and effects.

MEPLAN has been applied to many metropolitan areas and countries including London, England; Southeast England; Cambridge, UK; Santiago, Chile; Sao Paulo, Brazil; Bilbao, Spain; Tokyo, Japan; Helsinki, Finland; Caracas, Venezuela; Sacramento, USA; Naples, Italy; Bolzano, Italy; Madrid, Spain; San Sebastian, Spain; Basque Region, Spain; Colombia (national model); Chile (national model); Sweden (national model); Sao Paulo state, Brazil; and the central region of Chile.

Table 2.5 summarizes the strengths and weakness of MEPLAN (U.S. EPA 2000).

Table 2.5 Strengths and Limitations of MEPLAN

Strengths
– Comes close to modeling interrelated variables describing both land use and transportation. – Allows analysis of different kinds of policies. – Highly synthetic, allowing most of the description of even the base situation to be estimated within the model, reducing the reliance on observed data. – May be implemented with small amount of data except for the base year model. – Representation of the impact of zoning policies by including zoning restrictions on floor space in the spatial choice formulation as well as development costs. – Projection of increase in floor space partially based on development costs, thus allowing policies that offer monetary incentives to developers to build in targeted zones or at specified densities to be represented in terms of decreased development costs.
Limitations
– Data-intensive. Calibration process may be difficult and time consuming if base year observed data are lacking or inconsistent.

2.2.2 TRANUS

The TRANUS model is an integrated land use-transportation modeling package. Compared to transportation-only models, TRANUS provides future projections based on not only the growth and location of activities, but also the effects of transportation policies on the location and the land market, which influences the accessibility or functionality of the location.

TRANUS is designed to simulate and evaluate transportation, economics, and/or environmental policies. It assesses the implications of transportation policies on the location and interaction of activities and their effects on the land market. TRANUS converts economic flows (in annual dollars) by economic sector to daily travel demand in appropriate units. For example, economic flows from households to industries are converted to commute trips. Available economic flows in the model include industry-to-household flows, household-to-industry flows, industry-to-industry flows, household-to-household flows, internal-to-internal flows, and internal-to-external flows.

TRANUS requires the following user input: network nodes, links, and routes; travel time in the previous time period; activity location and land use data by zone; activity location; and land use variables. TRANUS produces many outputs, as listed below:

- All paths between each O-D pair for each travel mode and combination of modes.
- General assignment results for each link: total volume in demand units (passengers or tons), vehicles, equivalent vehicles, V/C ratio, and level of service.
- Detailed assignment results for each link: volume by each mode and route, speed and waiting times on the link under congestion conditions, demand/capacity ratio for transit vehicles.
- Indicators of the performance of the transportation system concerning users, operators, and administrators to be used in the evaluation processes. For users, TRANUS reports global demand by mode, total and average travel times, distance, cost, and disutility by

mode. For operators, TRANUS reports the number of passengers, passenger-kilometers, passengers' income (from tariffs), operating cost per passenger, and revenue in the simulation period. For administrators of the transportation infrastructure, TRANUS reports maintenance costs and income from tolls, road pricing, or any other charges.

- A database file containing all results from the transportation model, allowing the model output to be extracted and processed by using a standard database program.
- Transit route profile with demand-supply information for each route on each link: number of passengers boarding by route, waiting time for boarding passengers, demand and capacity of transit vehicles, available seats, and so on.
- Activity location and land use consumption outputs.

TRANUS has been used for many projects such as the Detailed Transport Demand Study for the Bogota Metro Systems, Columbia; Land Use and Transportation Model for the City of Valencia, Spain; Land Use and Transportation Model for the Baltimore Metropolitan Areas; and an Input-Output and Transport Model for the State of Oregon.

Table 2.6 presents the strengths and weaknesses of TRANUS (U.S. EPA 2000).

Table 2.6 Strengths and Limitations of TRANUS

Strengths
– One of the very few integrated land use and transportation models commercially available, backed by a sound history of practical applications in many countries. – User-friendly, with a powerful graphical Windows-based interface, supported by an object-oriented database and GIS interface capabilities. – Applicable to a large variety of cases, ranging from simple urban or regional models to highly sophisticated national or regional input-output models. – Site-licensing available. – Extensive email support for licensed programs. – Backed by a continued research and development process. New versions of TRANUS released annually, with frequent upgrades.
Limitations
– Large zones. – Requires a GIS program such as ArcView, TransCAD, or MapInfo to map the results of the model. – High price for the software.

2.2.3 DELTA

DELTA is both an urban and a regional model. At the urban level, it projects changes in the location of households, population, and employment and the amount of real estate development in an urban area. At the regional level, it provides projections of changes in the regional economy and migration between urban areas. The DELTA urban model consists of six sub-models that address the development process, demographic change (e.g., household formation), economic growth, location and relocation of households and jobs in the property market, car-ownership choices, changes in employment status (working/non-working) and commuting patterns, and changes in the quality of residential areas. The DELTA regional model contains three additional models for migration between different urban areas, the location of investment/disinvestment, and the pattern of production and trade.

DELTA is designed to interact with a transportation model to forecast land use changes. Based on accessibility changes estimated by the transportation model, DELTA assesses the impact based on a variety of variables, including the location of different activities (e.g., households and employment) and the value of buildings.

The input data requirement is flexible to allow the user to alter the input for each run of the model. The input to the DELTA urban and/or regional models includes information on the location of households and jobs, car ownership levels, floor space supply, and rent for a base year and the proceeding years. Variables that define the economic and demographic scenarios to be modeled and coefficients to describe the behavior of households, businesses, developers, etc., must be provided.

DELTA has been applied in Greater Manchester and the Trans-Pennine Corridor, England; Edinburgh, Scotland; Yorkshire, England; Sardina, Italy; and Uruguay.

Table 2.7 summarizes the strengths and weaknesses of DELTA (U.S. EPA 2000).

Table 2.7 Strengths and Limitations of DELTA

Strengths
– Unique capability to forecast changes over a series of short periods. – Allows the user to generate specific conditions for input into the model. – Provides an integrated software package that may be used as a stand-alone package or set up to interact with a wide-range of transportation models.
Limitations
– Unavailable as an off-the-shelf product. Licensing is on a project specific basis and includes the services of the model developer.

2.3 Linear Programming Models

Linear programming (LP) is more manageable, understandable, and computationally easier than other optimization techniques such as dynamic programming models, goal programming, hierarchical programming, linear and quadratic assignment, nonlinear programming models, and utility maximization models. LP models consist of one or more objective functions and a set of constraints. Land is allocated to each land use type to optimize one or more objectives. Some possible objectives are, for instance, maximization of a household's or individual's rent-paying ability, minimization of environmental impacts, maximization of population income, minimization of the cost of development (or maximization of the benefits of development), etc. Waddell and Ulfarsson (2004) argue that this approach is suited more to exploration of alternative land use configurations to optimize transportation flows than to describing realistic behavioral responses to changes in the transportation system or in land use policies. It may also be difficult to describe some of the regulatory policies and decision-making processes that constrain land development and transportation investment in a precise mathematical language. Behavior and uncertainty are also difficult to model when using the LP formulation.

The first model that used this technique in the analysis of land use is the Herbert-Stevens Linear Programming Model, which was designed for the Penn-Jersey Transportation Study (Herbert and Stevens 1960). Other models include TOPAZ (Dickey and Leiner 1983) developed in Australia and POLIS for the Bay Area (Prastacos 1985). These three models are described in the following sections.

2.3.1 Herbert-Stevens Linear Programming Model

The Herbert-Stevens Linear Programming Model (Herbert and Stevens 1960) was developed to locate land use activities for the Penn-Jersey Transportation Study. It is designed to obtain the optimal distribution of future households to forecasted residential land use in the metropolitan area. The model operates iteratively. During each iteration, the amount of available residential land and the number of households in the study region are forecasted exogenously. The Herbert-Stevens Linear Programming Model then allocates households to the available residential land in the study region, which is subdivided into smaller zones. Basic assumptions of the model include:

- Households choose their locations on the basis of an available total budget, a "market basket", and the costs of obtaining objectives such as maximization of (household or individual) rent-paying ability, minimization of environmental impacts, maximization of population income, minimization of the cost of development (or maximization of the benefits of development) and so on. The "market basket" is a unique combination of a residential bundle (a house, an amenity level, a trip set, and a site of a particular size) and a bundle of all other commodities consumed annually by a given household group, which is a collection of households with similar residential budgets and tastes regarding housing.
- For each household group, a number of market baskets exist, to which households in the group are indifferent.
- A household tends to optimize its condition by selecting from the set of market baskets one that maximizes its savings, which are defined as the rent-paying ability of the household for a particular site in a particular area.

The objective to be maximized is the aggregate rent-paying ability, which corresponds to the maximization of savings for each household. In mathematical form it is expressed as:

$$\max Z = \sum_{k=1}^{U} \sum_{i=1}^{n} \sum_{h=1}^{m} X_{ih}^{K} (b_{ih} - c_{ih}^{K})$$

subject to:

$$\sum_{i=1}^{n} \sum_{h=1}^{m} s_{ih} X_{ih}^{K} \leq L^{K}$$

$$\sum_{K=1}^{U} \sum_{H=1}^{m} -X_{ih}^{K} = -N_i, \qquad all X_{ih}^{K} \geq 0$$

where

U = total number of zones of the study region, $K = 1, \ldots, U$;
n = household groups, $i = 1, \ldots, n$;
m = residential bundles, $i = 1, \ldots, m$;

b_{ih} = residential budget allocated by a household in group *i* to purchase a residential bundle *h*;
c_{ih}^{k} = annual cost to a household of group *i* of the residential bundle *h* in area *K* – exclusive of site cost;
s_{ih} = number of acres in the site used by a household of group *i* if it uses residential bundle *h*;
L^{k} = *L* number of acres available for residential use in area *K* in a particular iteration;
N_i = number of households of group *i* that are to be located in a zone during a particular iteration; and
X_{ih}^{k} = number of households of group *I* using residential bundle *h* located by the model in area *K*.

The Herbert-Stevens model may be considered a land use change model in the sense that although it does not directly assess changes in the use of land, it allocates households to available residential land on the basis of particular behavioral assumptions.

Table 2.8 summarizes the strengths and limitations of the Herbert-Stevens Linear Programming Model.

Table 2.8 Strengths and Limitations of the Herbert-Stevens Linear Programming Model

Strengths
– The fine level of aggregation with respect to households allows the inclusion of households of different behavioral characteristics and permits a more realistic land allocation process. – The simulation of the market clearing mechanism involves a simple linear programming action. – Policy constraints are considered for the amount of available land. – An operational form makes its real world application possible.
Limitations
– LP formulation imposes the linearity assumption on both the objective function and the constraints, which may not always be the case in the real world. – The (good quality) data requirements of the model are heavy. – The iterative nature of the model ensures that the allocation of households within the given iteration period will be optimal, but it does not ensure that the allocation will be optimal in the aggregate.

2.3.2 TOPAZ/TOPMET

The Technique for Optimal Placement of Activities in Zones (TOPAZ) was developed by Drs. Brotchie, Sharpe, and Toakley from the Division of Building Research of the Commonwealth Scientific and Industrial Research Organization, Australia, in 1970. It is an optimizing program that maximizes a user-specified non-linear objective function subject to constraints to generate patterns of activity locations. TOPMET is a version of TOPAZ particularly tailored to a more detailed level of planning (Cambridge Systematics, Inc. 1991).

TOPAZ needs the exogenous forecasts of both employment by sector and total population for the entire region. With these forecasts, the TOPAZ model executes the procedures below:

(1) Allocate employment and housing to zones to minimize a weighted sum of the costs of the urban infrastructure and the incurred transportation costs.
(2) Generate trips, split them into two modes (road and rail), assign them to the network, and calculate the travel time and cost and the related land prices.
(3) Aggregate the data to the urban or regional level.

Several assumptions are made in the model:

- The pattern of households and employment locations may be described as allocations of new land uses in such a way as to optimize an objective function that consists of transportation costs and activity establishment costs.
- Constraints are intended to ensure that zones are not filled beyond capacity and all activities are allocated.
- Travel demand may be forecasted from spatial active locations using entropy-maximizing principles.
- Welfare economics principles are used to set an objective function in terms of maximizing social benefit.

TOPAZ needs input data on zonal employment, total or zonal population, land allocation by zone and by activity, transportation network (links, speeds, and capacities), trip matrix, travel time by mode, and trip lengths for a base year. It also requires data on employment by sector and by zone, land allocation by activity, constraints on land use, car ownership, and establishment costs of activities for the future. Future travel time by mode and trip lengths may be calculated endogenously and input into the model.

TOPAZ and TOPMET are both useful planning tools to determine what should happen under given objectives. The TOPAZ produces employment by sector, total population, land allocation by activity, vacant land, and the location of houses. Trips by mode and by journey type, travel energy used by origin and purpose, and (optional) air pollution as a consequence of travel are generated by the model. Additionally, the model produces two types of indicators: planning indicators and economic indicators. Planning indicators include accessibility by zone by trip type and by mode, and economic indicators are developers' costs, travel costs by mode and by trip type, and the marginal cost of incrementing activity levels.

TOPAZ was originally developed for the Melbourne region, Australia. It has also been studied in Blacksburg, Virginia; Gosford-Wyong, Australia; and Darwin, Australia. Since it is applicable at different geographic scales, ranging from regional land use configurations to the organization of individual buildings, it has been applied to areas such as New River Valley, Australia, and on a smaller scale, to arrange buildings on the university campus of the Virginia Polytechnic Institute and State University, as well as at the location of rooms in hospitals in Sydney, Australia.

2.3.3 POLIS

The first version of the Projective Optimization Land Use System (POLIS) was developed for the City of Cologne beginning 1969 and later applied to the cities of Vienna and Darmstadt. In the 1980s, the Association of Bay Area Governments (ABAG) applied POLIS to the San Francisco Bay Area.

POLIS is a structured mathematical programming optimization model. It simulates the spatial distribution of population, employment, buildings, land use, and transportation through iteration of a number of time periods until a planning horizon is reached. While the spatial distribution of activities is optimized with respect to an objective function, planning constraints are also satisfied. The form of the objective function in POLIS is derived from the random utility theory and describes the behavior of individuals in selecting from a set of alternatives one that maximize their utility. The constraints of the model describe housing and land supplies, development policies of different cities, and employment/housing to be allocated among all the zones within a county (Southworth 1995).

Prastacos (1985) distinguished the POLIS model from Lowry-type models such as ITLUP in three key aspects: use of microeconomic behavioral principles, formation as a mathematical programming problem, consideration of job location, basic and non-basic employment, residence selection, and trip making in an integrated manner. The model has the following form:

$$
\begin{aligned}
&MaxZ(T_{ijm}, S_{ij}^{k}, \Delta E_{j}^{n}, \Delta H_{i}) \\
&= -(1/\beta)\sum_{ijm}[\ln(1/W_i \sum_{m} T_{ijm}) - 1] - (1/\lambda)\sum_{ijm} T_{ijm}[\ln(T_{ijm})] \\
&- \sum_{ijm} T_{ijm} c_{ijm}^{W} - \sum_{k \in K}(1/\beta_{k}^{S})\sum_{ij} S_{ijk}[\ln(S_{ij}^{k}/W_{j}^{k}) - 1] - \sum_{ijk} S_{ij}^{k} c_{ij}^{k} + \sum_{i,n \in K}(f_{i}^{n})\Delta E_{i}^{n}
\end{aligned}
$$

where

T_{ijm}	=	number of work trips from zone *i* to zone *j* by mode *m* (private or public transport);
S_{ij}^{k}	=	number of trips in the "retail" or local service sector *k*;
ΔH_i	=	number of new households located in zone *i*;
c_{ijm}	=	interzonal travel cost by mode *m* (all service sector travel assumed to be by automobile);
c_{ij}^{k}	=	travel costs in the "retail" or local service sector *k*;
W_i	=	attractiveness of zone *i* to be chosen for residence;
W_j^k	=	attractiveness of zone *j* as a center for retail or local service activity;
f_i^n	=	an agglomeration potential function specific to zone *i*;
α^n	=	exponent of this agglomeration function (a model parameter to be estimated);
ΔE_i^n	=	number of additional jobs in the basic employment sector *n* ($n \in K_{bas}$) to be located in zone *j*; and
β, β_k^s, and λ	=	spatial interaction and modal split sub-model parameters to be estimated.

The model is subjected to a set of constraints, including trip production and attraction to be consistent with housing and employment availability in a zone, employment of related sectors to be allocated jointly, the limitations of housing and employment allocation, and the exogenous location of employment and housing (policy constraints).

The model seeks to jointly maximize the locational benefits associated with multimodal travel to work and retail, local service sector travel, and, significantly and jointly, the agglomeration benefits accruing to basic-sector employers.

2.4 Micro-Simulation Models

Micro-analytic simulation, or "micro-simulation" for short, refers to computer models that simulate the behaviors of individuals of a representative population, the cumulative effects of which form the overall behavior of the system, and draw conclusions that apply to higher levels of aggregation of the entire population. For instance, if a response is a specific travel choice, the summation of all individual responses provides the aggregate travel demand in the system for planning studies. These models distinguish themselves from aggregate models in that the explanatory variables reflect the characteristics of individuals and their decision-making processes, whereas in an aggregate model, the explanatory variables represent the collective properties of the objects (such as population groups or households of different types) being modeled. Aggregate models therefore model the collective effects of individual behaviors but cannot explain them, thus resulting in the inability to deal with certain policy issues. Microsimulation models are relatively easy to understand and implement since the decision-making process may be modeled.

With advances in the field of computer technology and greater availability of detailed data, micro-simulation has become an increasingly popular analysis tool because micro-simulation allows modeling at an individual level, where behavioral theory is clearer and individual-level analysis is more effective than cross-classification of households using multiple characteristics. It has been applied in the formulation of urban models such as NBER/HUDS (Kain and Apgar 1985), MASTER (Mackett 1992), IRPUD (Wegener 1985), and UrbanSim (Waddell 2002), which are described in the following sections.

2.4.1 NBER/HUDS

This section describes two simulation models developed in the 1970s: the NBER (National Bureau of Economic Research) and the HUDS (Harvard Urban Development Simulation) model. The HUDS model is a direct descendent of the NBER model, and both model the demand for housing, supply of housing, and a housing market clearing mechanism.

The NBER and HUDS models simulate the housing market clearing process annually by employing a disequilibrium framework. Housing is modeled in terms of structure type, neighborhood quality, and quantity of structure services. HUDS is designed to evaluate the impacts of spatially concentrated housing improvement programs (Kain 1986). As shown in Figure 2.2, the structure of the basic NBER model consists of six submodels corresponding to the demand and the supply side of the housing market.

The first submodel is the Filtering submodel, which provides a distribution of expected housing prices by dwelling type and zone and an aged and renovated housing stock. The second submodel is the Employment-Location submodel, which translates changes in employment levels and composition (by industry) at each workplace into changes in employee's household characteristics such as the age of the head of household, family size, income, education, and race for each of the 96 household types. The third submodel is the Movers' submodel, which generates demand in the housing market, as well as supply as migration makes some housing units available. The fourth submodel is the Demand-Allocation submodel, which deals with demand allocation and allocates housing demanders of the 96 household types at each workplace to one of the 50 housing submarkets defined as 50 housing bundles based on five neighborhood quality levels and 10 structure types. The fifth submodel is the Supply submodel, which determines the supply by estimating demolitions, conversions, and new construction for each zone. The sixth and last submodel is the Market-clearing Assignment model, which assigns each housing submarket participant to residence zones based on a linear programming algorithm.

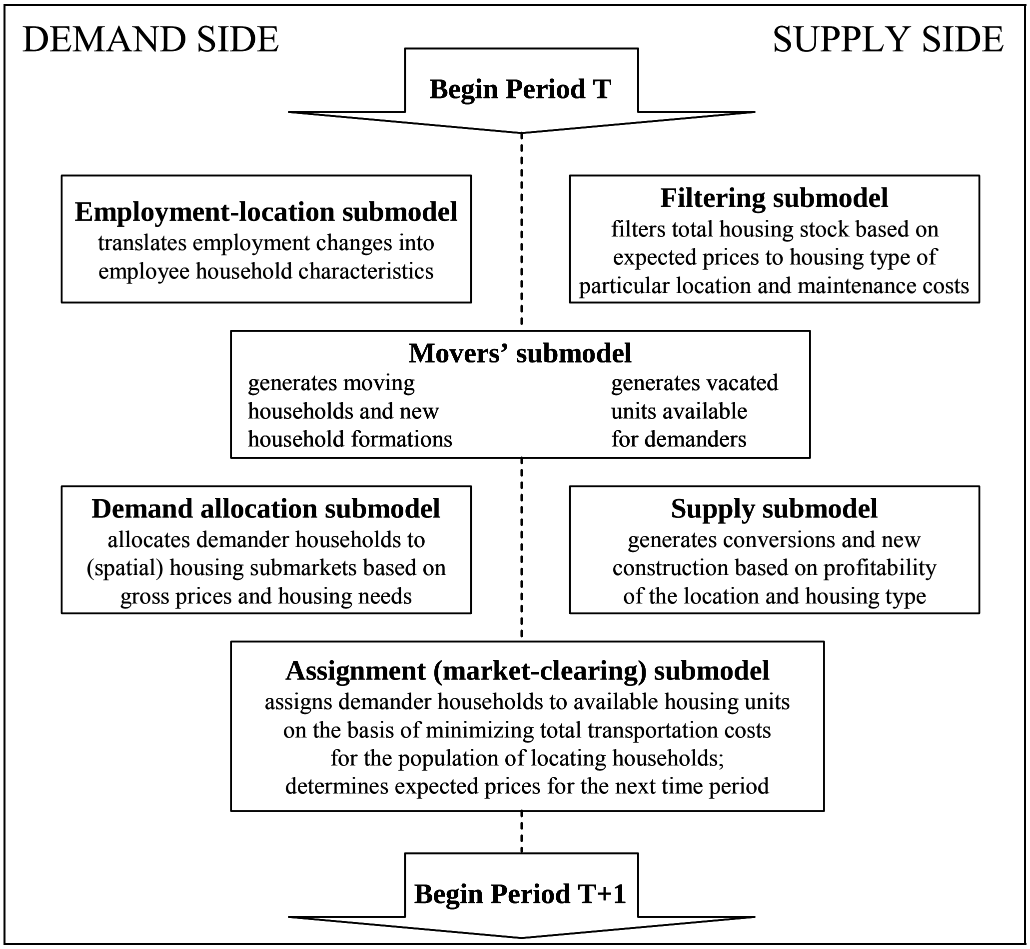


Figure 2.2 Structure of the Basic NBER Model (Source: Kain 1986)

Table 2.9 summarizes the strengths and limitations of the NBER and HUDS models.

Table 2.9 Strengths and Limitations of the NBER and HUDS Models

Strengths
– Suited to analyzing impacts of housing policies.
Limitations
– Applications restricted to urban areas. – Lacks a host of other explanatory factors (environmental, social, cultural, political, and institutional) in the analysis of land use change.

2.4.2 MASTER

The Micro-Analytical Simulation of Transport, Employment and Residence (MASTER) model, developed in the United Kingdom by Mackett (1990a, 1990b), is an integrated land use-transportation model that operates at the household level. Population growth and household structure are modeled based on the lifecycle including birth, aging, death, marriage, divorce, and migration. The choice of residential location is based on the weighted function of generalized work-related travel costs for the head of a household (only work trips are modeled), while housing type choice is based on household size and composition. The supplies of housing and jobs are exogenous data input into the model, and housing vacancy is tracked for each zone. Household members' choice of jobs, employment and unemployment, retirement, education level, sex, social group of the head of household, job vacancies, and salary ranges are all modeled in MASTER.

The transportation processes modeled include acquisition of a driver's license, auto ownership, car availability, and work trip mode choice, all of which are functions of age, sex, household income, household composition, and travel costs. The travel costs are estimated based on travel distance without considering congestion since the model does not assign traffic to each route. While incorporating a traffic assignment routine seems to be straightforward, Southworth (1995) questions Mackett's suggestion that 1% sample of households is necessary for model calibration. Instead, Southworth suggested that up to a 100% sample might be necessary if trips are to be assigned to specific routes.

2.4.3 IRPUD

The IRPUD model, formally called Dortmund, is a simulation model of intraregional locations and mobility decisions in a metropolitan area. The IRPUD model was developed for Dortmund, Germany by Wegener and his colleagues (Wegener 1982, 1983, and 1985). The IRPUD model consists of six interlinked submodels operating in a recursive fashion on a common spatiotemporal database. The submodels are: Transport, Ageing, Public Programs, Private Construction, Labor Market, and Housing Market (Wegener 1998).

The Transport submodel estimates work, shopping, service, and school trips for four socioeconomic groups, as well as three travel modes of walking/cycling, public transport, and car use. The Ageing submodel computes all of the changes of the stock variables such as employment, population, and households/housing in the model, which are assumed to result from biological, technological, or long-term socioeconomic trends originating outside the model. The Public Programs submodel processes public programs (e.g., infrastructure investments and public housing programs) specified by the model user regarding employment, housing, health, welfare, education, recreation and transportation. The Private Construction submodel models the regional land and construction market by considering investment and location decisions by private developers. The Labor Market submodel models intraregional labor mobility, such as decisions of workers to change their job location in the regional labor market. The Housing Market submodel simulates the intraregional migration decisions of households, such as search processes in the regional housing market. The housing search is modeled in a stochastic micro-simulation framework.

Figure 2.3 depicts the recursive process of the six submodels. While the Transport submodel is an equilibrium model referring to a specific point in time, all other submodels are incremental and refer to a period of time.

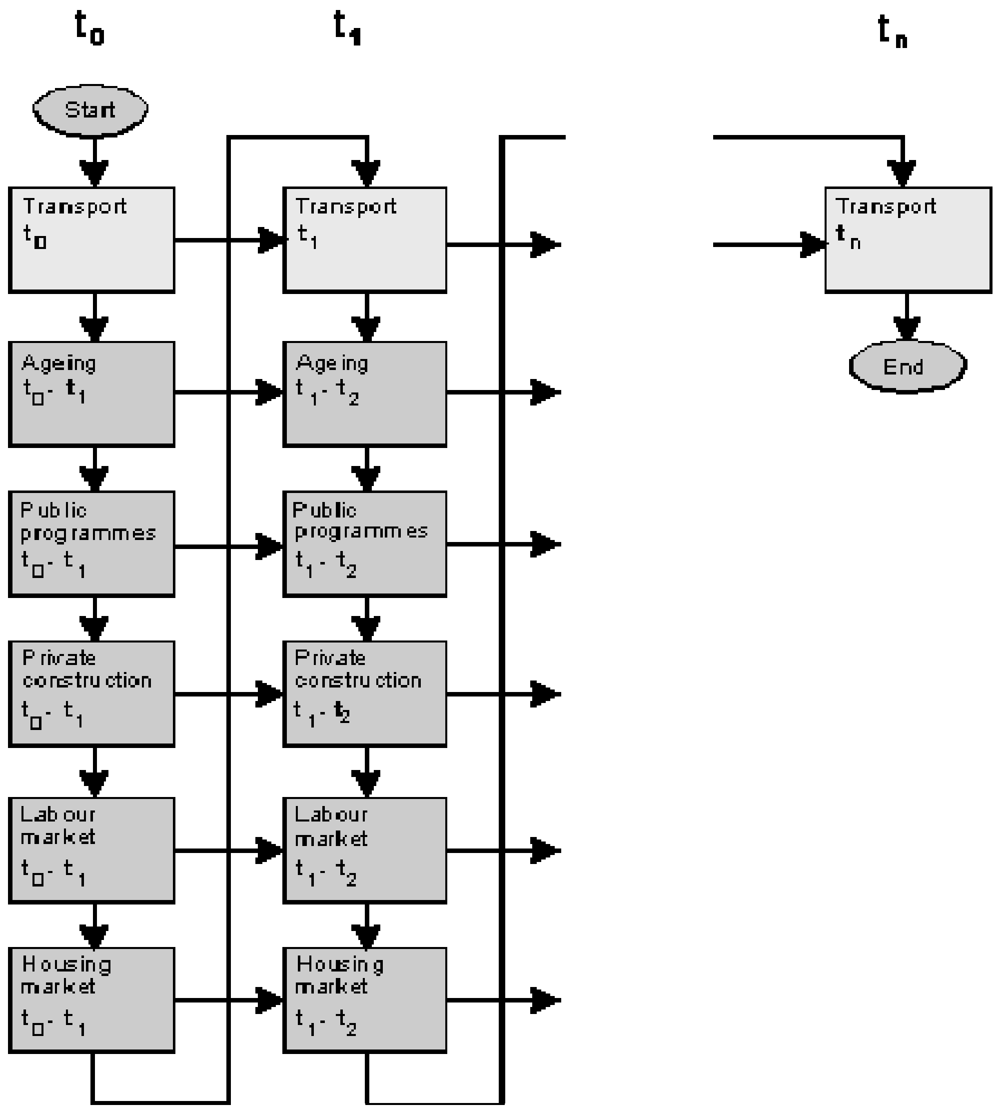


Figure 2.3 Recursive Processing of Submodels (Source: Wegener et al. 2001)

The IRPUD model requires four types of data as input: model parameters, regional data, zonal data, and network data. Model parameters are exogenously estimated and include demographics, household, housing, technical, monetary, and preference parameters. Regional data include employment, immigration, and emigration. IRPUD must have zonal information such as population, labor force/unemployment, households, dwellings, households/housing, employment/workplaces, public facilities, land use and rent/prices. The IRPUD model considers transportation networks associated with travel modes such as automobile, public transportation, and walking/bicycling. Network data include link type, direction, length, and travel time (public transportation) and base speed (road). For each public transportation line, additional information is required such as lists of nodes and peak-hour headway.

The IRPUD model has been applied to the metropolitan region of Dortmund, Germany. Table 2.10 summarizes the strengths and limitations of the IRPUD model (U.S. EPA 2000).

Table 2.10 Strengths and Limitations of the IRPUD Model

Strengths
– Capable of dealing with the impacts of policies related to industrial development, housing, public facilities, and transportation. – Addresses global policies (i.e., those that affect urban development in the whole region) and local policies (i.e., regulatory or direct zone-specific investment projects). – High temporal resolution and full integration of land-use transportation interaction in each simulation period, making it a truly dynamic model (compared with other approaches, such as cross-sectional equilibrium approaches). – Introduces urban modeling assumptions about human spatial behavior drawn from time-space geography based on time and cost budgets and satisfying behavior, which makes the model uniquely suitable to model elastic trip generation behavior (responsible for much of the growth in mobility in metropolitan regions).
Limitations
– Coarse spatial resolution.

2.4.4 UrbanSim

Developed by the Urban Planning Department and the Computer Science Department at the University of Washington, UrbanSim is a simulation model for integrated planning and analysis of urban development, incorporating the interactions between land use, transportation, and public policy. UrbanSim is a public domain software package, and is intended for use by metropolitan planning organizations and other planning organizations to interface existing travel models with new land use forecasting and analysis capabilities. The University of Washington received three new grants in September 2001 from the National Science Foundation to continue the development of UrbanSim (UrbanSim 2001).

The development of UrbanSim has been motivated by the need of planning organizations to test policies that deal with environmental, sociological, and economic concerns (UrbanSim 2001). Examples of such policy issues include preserving prime agricultural lands, forests, wetlands, open space, redevelopment, infill, and inner-city decline. Possible strategies developed based on such policies may range from urban growth boundaries at the regional or metropolitan scale to street design, mixing of uses, and pedestrian access at the neighborhood or site-specific scale.

UrbanSim is a discrete choice model based on the random utility theory. With a modular design, it models household location choice, employment location choice, and real estate development. It also interfaces with travel models and incorporates urban dynamics into the models. Households are classified in a disaggregate manner by income, persons, workers, and the presence of children. Employment is classified into 10-20 sectors, while real estate into 24 development types. Real estate measures include acres, housing units, and floor space. Real estate prices are also modeled. The model currently operates at a geographic scale of 150-meter grid cells (Waddell 2001), but the development of a TAZ version of the model is underway, according to Dr. Waddell.

UrbanSim has eight core models:

- Given control totals, the Demographic Transition Model simulates births and deaths in the population of households. Distribution of income groups, age, size, and presence or absence of children may also be specified. Iterative proportional fitting (Beckman et al. 1995) is used to add or delete households. Newly created households are added to the household list to be assigned to housing units by the Household Location Choice Model later in this report.

- The Economic Transition Model simulates job creation and loss with a control total and distribution of business sectors specified.
- The Household Mobility Model simulates household decisions such as where to move based on probabilities determined from historical data.
- The Employment Mobility Model determines which jobs will move from their current locations during a particular year using a similar approach to that in the Household Mobility Model.
- The Household Location Model chooses a location for each household in the housing list. To do this, a list of vacant housing is maintained and a multinomial logit model calibrated to observed data is used to select a housing from a random sample of the vacant housing units, which are described by attributes of the housing in the grid cell (price, density, age), neighborhood characteristics (land use mix, density, average property values, local accessibility to retail), and regional accessibility to jobs.
- The Employment Location Model is responsible for determining a location for each job that has no location. Alternatives are selected through random sampling, which are described by real estate characteristics in the grid cell (price, type of space, density, age), neighborhood characteristics (average land values, land use mix, employment in each sector), and regional accessibility to population.
- The Real Estate Development Model simulates developers' choices about location and type of construction to undertake. Each year, the model iterates over all grid cells where development is allowed and creates a list of possible transition alternatives (representing different development types), including the alternative of not developing. The probability for each alternative being chosen is calculated with a multinomial logit model. Variables in the developer model include characteristics of the grid cell (current development, policy constraints, land and improvement value), characteristics of the site location (proximity to highways, arterials, existing development, and recent development), and regional accessibility to population.
- The Land Price Model simulates the land price of each grid cell as the characteristics of locations change over time based on urban economic theory. The model is calibrated from historical data using a hedonic regression to include the effects of the site, neighborhood, accessibility, vacancy rates, and policy effects on land prices. The model variables are similar to those in the Development Model.

The model requires input on population and employment estimates, regional economic forecasts, transportation system plans, land use plans, and land development policies such as density constraints, environmental constraints, and development impact fees. The user is allowed to create "scenarios" as input to UrbanSim by specifying alternative forecasts of population and employment, land-use policy assumptions, transportation infrastructure assumptions, etc. The model then provides output regarding future year distributions of population, households by type (e.g. income, age of head, household size, presence of children, and housing type), units of housing by type, businesses by type (e.g., industry and number of employees), land use by type

(user-specified), square footage of nonresidential space by type, densities of development by type of land use, and prices of land and improvements by land use (UrbanSim 2001).
Input data required to run UrbanSim are listed below.

- Parcel data in ArcView shape file. The attribute table should have information on lot size, land use, housing units, square footage of building space, year built, zoning, land use plan, assessed land value, and assessed improvement value.
- Business establishments, which should be geocoded.
- Household data from the census STF3A.
- Environmentally sensitive areas such as wetlands, floodplains, high slopes, fault zones.
- Urban Growth Boundaries or other policy boundaries.
- TAZ layer as an ArcView shape file.
- Travel impedance from travel models (peak times and logsums).

Calibration of the model requires knowledge of statistical software to perform multiple regression and logit model estimation using external econometric software such as Alogit or Limdep.

UrbanSim has been validated for the Eugene-Springfield, Oregon area (population 375,823) using data from 1980 to 1994, and has been applied in Honolulu, HI and Salt Lake City, UT. Recently, Puget Sound, WA adopted UrbanSim. Other metropolitan areas are beginning to utilize it as well. The UrbanSim software is distributed as open source software under the GNU General Public License, which allows anyone to use, modify, and redistribute the source code at no cost. The source code of UrbanSim is available at www.urbansim.org.

The strengths and limitations of UrbanSim are presented in Table 2.11 (U.S. EPA 2000).

Table 2.11 Strengths and Limitations of the UrbanSim Model

Strengths
– Dynamic behavioral foundation, which makes the model more transparent and explainable to users and decision-makers. – Reflects real-world processes, which make the model easier to evolve and to interface with other process models such as environmental models. – High degree of spatial resolution, currently using spatial grid of 150 meters, for interface with environmental data. – A visualization component that provides integrated 2- and 3-dimensional mapping in addition to charts and graphs for interpreting and comparing model results, and for diagnosis during model development and testing. – Program and source code that are free and available for use and modification.
Limitations
– High data requirements; data mining and synthetic data cleaning tools are currently being designed to facilitate working with messy data. – Experience limited to current applications in Hawaii, Oregon, Utah, and Washington. – Rapid evolution, with the first major release based on a complete redesign of the software architecture in the second quarter, 2000.

2.5 Random Utility/Discrete Choice Models

Discrete choice models are designed to model an individual's choices by taking into account the characteristics of that choice. Discrete choice modeling techniques have long been used in travel demand modeling, especially in the analysis of mode choice. For example, a discrete choice model may be used to predict mode choice based on the travel cost, travel time, and other characteristics of each mode. Discrete choice models may also be applied for land use planning, such as making decisions for locations of households and firms. For instance, employment location choice may be modeled as a function of the characteristics of business (such as industry size), potential zones (such as accessibility, density, and employment levels), and space (quantity and cost) with a discrete choice framework.

Daniel McFadden's work on Random Utility Theory and his derivation of the generalized models, including multinomial and nested logit models, have resulted in discrete choice models that have a firm foundation in the area of econometrics and have become a standard method for developing models to predict individual choices among a finite set of alternatives (Waddell and Ulfarsson 2004). In this class of models, there are METROSIM (Alex Anas), Boyce (David Boyce), 5-LUT (Francisco Martinez), OMPO (Paul Waddell), and RURBAN (Miyamoto and Kitazume). This chapter describes one of the models, METROSIM, which was recently developed and implemented.

METROSIM is an urban simulation model developed by Alex Anas and Associates (1998) that uses an economic approach to forecast the interdependent effects of transportation and land use systems, as well as land use and transportation policies for metropolitan areas. NYMTC-LUM is a customized version of METROSIM, developed for the New York Metropolitan Region, which comprises the land-use model component of METROSIM linked to the existing MTC transport model.

METROSIM forecasts travel flows, employment changes, congestion levels, new construction of residential and commercial buildings, land use changes, etc. METROSIM also provides benefit-cost ratios for transportation projects or policy interventions. METROSIM is capable of obtaining a one-shot long run equilibrium forecast for transportation and land use in a metropolitan area or producing yearly changes in transportation and land use from the base condition to a steady state, when convergence is achieved.

METROSIM employs the discrete choice method with a market clearing mechanism. METROSIM attempts to reach equilibrium between three major market sectors including labor market (jobs), housing market (dwellings), and commercial floor space. METROSIM iterates between these markets and the transportation system until land use and transportation flows reach an equilibrium state.

METROSIM is designed to use standard U.S. Census data sources and parcel data from tax assessors in metropolitan areas. Additionally, the following input data are required to run the model:

- CTPP elements 1, 2, and 3, and Bureau of the Census STF1A and STF3A files
- Transportation network

- Land use by type of land use
- Regional input/output model (optional)
- Land and property values by zone and land use type (optional)

The outputs available from the model are listed below:

- Basic industry distribution by zone and by type of basic industry
- Non-basic industry distribution by zone and by type of non-basic industry
- Residential real estate distribution by type and zone
- Non-residential real estate distribution by type and zone
- Vacant land distribution by type and zone
- Households
- Travel (commuting and non-work)
- Traffic assignment on the network
- Rents and market prices for each type of real estate by zone
- Vacancy rates for each type of real estate

This model has been applied to a number of metropolitan areas in the U.S. including Chicago, IL; Houston, TX; Harlem Line Corridor, New York City, NY; New York City, NY Region; Pittsburgh, PA; Staten Island, NY; and San Diego, CA.

The strengths and limitations of METROSIM are presented in Table 2.12 (U.S. EPA 2000).

Table 2.12 Strengths and Limitations of the METROSIM Model

Strengths
– Firmly rooted in economics and recognizes how market forces operate in shaping and changing land use. – Deals with land use policy and land use change explicitly. – Fast execution on computer and does not rely on approximate solutions.
Limitations
– Currently no GIS interface, which may be easily developed at a small fee. – Proprietary software (license fee \$20,000 - \$30,000, \$2,500 for three initial runs, user support \$5,000 - \$10,000/year, training \$10,000 one time).

2.6 Cellular Automaton Models

Cellular Automata (CA) has been used in many fields: physics, chemistry, biology, philosophy, sociology, and geography. Nowadays CA is applied to city planning as an important modeling and simulation tool. Following the first study using the CA approach in geographical planning, Couclelis (1996) and Takeyama (1996) generalized modeling language, enabling integrated, dynamic, and spatial modeling at different scales and using GIS. Batty and Xie (1994) developed a CA-based model for not only land use changes but also urban growth and form.

In CA-based urban models, cells simulate four types of settlements including trade, industrial, residential, and empty areas. Interactions among cells vary depending on the distances between them. The approach is particularly useful for representing the interactions between a location and its immediate environment, but tends to reflect a fairly abstract representation of agents, decisions, and behavior, since the models focus on simulating changes at individual cell levels.

To date, applications have been limited mostly to research purposes rather than operational planning or policy planning, although efforts are underway to make these models useful for planning purposes (Waddell 2004). Additionally, challenges remain in reconciling the emergent behavior of cells acting on localized rules with more systemic or macro-scale behavior, in validating these models using observed data, and in computational requirements. The most ambitious use of the CA approach to date is the TRANSIMS traffic microsimulation system, which has been tested in Portland, Oregon (LANL 2002). In this chapter, SLEUTH, a CA-based land use model, is described.

2.6.1 SLEUTH

The SLEUTH (Slope, Land use, Exclusion, Urban, Transportation, Hill shading) model is commonly known as the Clarke Cellular Automata Urban Growth Model or as the Clarke Urban Growth Model. This model simulates the changes from non-urban land-use such as agricultural, forest, wetlands, water, preservation, park land to urban land-use such as residential, commercial, mixed use, industrial, and other land uses based on a grid of cells (cellular automaton) to understand how urban areas extend to their surrounding land and the environmental impact brought by this extension on the local environment (Clarke et al. 1996, U.S. EPA 2000).

An underlying assumption of the model is that historical growth trends will continue and that the future may be projected based on these trends. Under this assumption, all the cells are updated synchronously in discrete time steps (one year) and the state of each cell depends on the previous state of its surrounding neighbors. In the model, each cell is used to model land use changes and the land use state of each cell is predicted based on local factors (e.g., roads, existing urban areas, topography), temporal factors, and random factors. Urban land is defined as residential, commercial, mixed use, and industrial land uses.

Six data inputs are required to run the SLEUTH model. They include slope, land use, urban, exclusion, transportation, and hill shading. Since spatial framework of the model is raster-based, the input data are required to be in a raster format. Slope data in GIF format may be derived from a Digital Elevation Model (DEM). Excluded areas where urbanization cannot occur are also provided in GIF format. These areas include, for instance, water bodies, roads/transportation network, urban extent, and land cover. The main sources of data that have been used in case studies include the U.S. Geological Survey (USGS), the Regional Planning Association, and the U.S. Census Bureau.

The model outputs are provided as a set of GIF image files that may be merged into an animation or brought into a GIS as data layers. The resolution of output images depends on the resolution of the input data. The model output includes a snapshot of a particular year, a cumulative image produced by multiple runs that shows a probability of urbanization for a given year, a set of best fit metric between modeled and real data for calibrating the model, actual values of model output for control years averaged over the number of model simulations, the standard deviations of the average actual values, final coefficient values, and the start and stop times for an entire model execution (U.S. EPA 2000).

SLEUTH has been applied to several metropolitan areas in the U.S, including Baltimore-Washington, DC; Chester County, PA; Orange County, CA; Santa Barbara, CA; San Francisco, CA; Sterling Forest, NY; Utah Front Range, UT; Chicago-Milwaukee; Detroit, MI; Greater New York Area; Mid-Atlantic Interstate Area; Middle Rio Grande Basin, NM; and Philadelphia-Wilmington, PA.

The strengths and limitations of the SLEUTH model are presented in Table 2.13 (U.S. EPA 2000).

Table 2.13 Strengths and Limitations of the SLEUTH Model

Strengths
– Concurrently simulates four types of growth (spontaneous, diffusive, organic, and road influenced) – Provides both graphical and statistical outputs. – Incorporates momentum of booms and busts using a threshold multiplier with subsequent temporal decay. – Allows for relatively simple alternative scenario projection.
Limitations
– Does not explicitly deal with population, policies, and economic impacts on land use change, except in terms of growth around roads. – Not based on economic theories but relies on historical trends. – The growth assumption may not hold.

2.7 Rule-Based Models

Rule-based land use models are useful tools for MPOs and counties for long-range scenario testing because they are easy to apply and the data required to operate the models are generally available. Although rule-based land use models are developed based on economic theories and market rules, rules are not comprehensive or flexible enough to model the complex economic and market processes in detail. In general, there is a risk that model users would interpret the models as having a more behavioral basis than their rules actually contain.

Several land use models have been developed in recent years implementing a set of rule-based procedures to allocate population, employment, and/or land use based on GIS platform. This chapter describes some of the rule-based models including the CUF model (CUF-1 and CUF-2), SAM, UPLAN, What If?, SLAM, and ULAM.

2.7.1 CUFM: CUF-1/CUF-2

The California Urban Futures (CUF) Model, also known as CUF-1, was developed based on the Bay Area Simulation System (BASS II) by Landis (1994) at the University of California at Berkeley to simulate the impacts of alternative regulatory and investment policy initiatives on urban development in the Northern California Bay Region.

The CUF-1 model is designed to simulate the effects of growth and development policies on the location, pattern, and intensity of urban development at various levels of government such as state government, local government, and special districts. The CUF-1 model uses two primary units of analysis: political jurisdictions (incorporated cities or counties) and developable land units (DLUs).

The CUF-1 model consists of four related submodels: the bottom-up population growth submodel, the spatial database that updates the geometry, location, and attributes of each DLU, the spatial allocation submodel, and the annexation-incorporation submodel. Using these submodels, the CUF-1 model first projects population growth based on city population growth trends and development potential by DLUs. The projected population growth is updated on the map layers in the spatial database. The CUF-1 model then simulates the growth of an area and allocates new development to each DLU per model period based on the population growth of each city or county, the profitability potential of each DLU, and user-specified development regulations and/or incentives (Landis 1994, U.S. EPA 2000).

The California Urban Futures Model Second Generation (CUF-2) model performs many of the functions of the CUF-1 model. However, as a second-generation model, it addresses some of the theoretical holes in the first model. Modifications made for the CUF-2 model are listed below (U.S. EPA 2000):

- The employment projection is a new component of CUF-2, although CUF-2 projects the future population and households as does CUF-1.
- In CUF-2, DLUs are defined by one-hectare grid-cells, not irregularly-shaped polygons as in the CUF-1.
- Spatial bidding is allowed in CUF-2 for sites between four types of new development land uses and three types of redevelopment.

CUF-1 has been applied to Solano and Sonoma counties in Northern California, while CUF-2 has been used in San Francisco Bay Region, CA.

Table 2.14 and 2.15 list the strengths and limitations of the CUF-1 and CUF-2 models, respectively.

Table 2.14 Strengths and Limitations of the CUF-1 Model

Strengths
– Easy to use and visualize. – Alternative policy scenarios prepared quickly (in hours) and in easy-to-read maps. – Modular system of related but independent submodels that may be updated. – Alternative development futures simulated based on specific policy changes.
Limitations
– Limited to residential development and no methods for projecting and/or allocating future industrial, commercial, and public activities. The most profitable sites to develop reserved for residential development (unless explicitly prohibited). – Lack of "infill" development and redevelopment by assuming almost all population growth will occur at the urban edge. – Growth allocation primarily depends on development profitability thus insensitive to other factors that impact growth patterns and locations (e.g., infrastructure investments). – Rules for allocating future development were not based on historical experience. – Currently unavailable as an "off the shelf" product.

Table 2.15 Strengths and Limitations of the CUF-2 Model

Strengths
– Easy to use and visual: The CUF-2 model allows users to prepare and evaluate alternative policy scenarios quickly (a typical simulation can be completed in a matter of hours) and in easy to read map form at almost any level of spatial detail. – Expandable: The CUF-2 model is designed as a modular system of related but independent submodels that can be updated to include new information and theories. – Policy approach: The CUF-2 model simulates alternative development futures based on specific policy changes. – Calibrated to past local experience.
Limitations
– Availability: The CUF-2 model is currently unavailable as an "off the shelf" product. – Data intensive: The CUF-2 model requires much more data than the original CUF-1 model. – Model calibration requires detailed knowledge of statistics. Results may be spatially auto-correlated.

2.7.2 SAM/LAM/SAM-IM

The Subarea Allocation Model (SAM) is a GIS-based land use allocation and forecasting model. The SAM forecasts future distribution of land uses within a given study region using demographic and geographic information. SAM was developed for the Maricopa Association of Governments (MAG), Arizona to support official adopted land use planning for the transportation and air quality programs in Maricopa County. There is another version of this model, formally known as the Land Use Analysis Model (LAM), which was developed for the Middle Rio Grande Council of Governments in Albuquerque, New Mexico. Both versions of the model were developed by Planning Technologies, LLC (U.S. EPA 2000).

The modeling of land use, population, and socioeconomics at the MAG consists of three tier models (Walton 2004). The first tier model is a demographic model used to produce county control totals. This model projects births, deaths and net migration in each county for a fifty-year time horizon. This model considers population by age, sex, birth rates, death rates, and net migration trends. The model takes into account short-term economic conditions, but not long-range employment trends. As the second tier model, the MAG is using DRAM/EMPAL to allocate the county control total population and employment to sub-regions. DRAM/EMPAL projects the spatial patterns of households and employment in the MAG region. The third tier SAM allocates population and employment from the Regional Analysis Zones (RAZ) to one-acre grids, which are then aggregated to Traffic Analysis Zones (TAZ).

The original SAM model was implemented as a set of ARC/INFO AMLs executed on UNIX, using the ARC/INFO GRID module to convert feature coverages to grids. The model was re-implemented in ArcView after Environmental Systems Research Institute, Inc. (ESRI) released the Spatial Analyst extension of ArcView in 1997. The new model was named Subarea Allocation Model - Information Manager, SAM-IM (Walton 2004). Features provided in the SAM-IM model include (U.S. EPA 2000):

- Analyzing land use plans: SAM-IM provides a number of measures of relationships between different types of land use, such as measures of job-housing balance for a region or for individual communities in a region.
- Creating and editing land use plans: SAM-IM provides the planner with a toolbox of editing and drawing tools to create a database to reflect the growth management strategy.
- Creating site evaluations: The application is equipped with a toolbox that makes it easy to create site evaluation themes that describe the characteristics of land and its suitability to be developed.
- Projecting future land use patterns: SAM-IM disaggregates or allocates a spatial growth forecast for a region by giving the planner a way to represent the value of land and the probability if developed.
- Supporting other urban model systems: The SAM-IM platform includes a "geographic calculator" that allows users to easily create and format data sets for other modeling systems, such as the transportation model EMME/2.

The geographic input information needed by the model includes an existing land use map and a proposed land use map. The demographic information includes the dwelling unit density and employment density for each land use type to generate a regional forecast. There are also optional inputs that the user may choose to incorporate into the model, such as proposed project information and other areas of concern (e.g., protected habitats, stream buffers) represented in GIS coverages. The result of a SAM-IM simulation is a land use map that depicts land use throughout the region for a given future year.

Table 2.16 presents the strengths and limitations of SAM-IM (U.S. EPA 2000).

Table 2.16 Strengths and Limitations of the SAM-IM Model

Strengths
– Editor feature: allows users to edit land use files while maintaining planar polygon topology by adding deleting, copying, pasting, splitting, and recoding the use and density of land use polygons. – Toolbox of functions: significantly extends ArcView's capabilities for cellular representation of geography in the Spatial Analyst by letting users build cellular land use grids, import grids from other conventional polygon themes on any attribute, compute new grids, compute new grids through look-up tables, measure proximities, feature buffers, and neighborhood sums, etc. – Land use "scenarios": supports scenarios representing alternative land use plans, growth policies, and proposed projections for target years and time-series of long-range forecasts over extended periods of time, at five-year intervals. – Adherence to land use policies and constraints for a region. – Geographic calculator: allows generation of new datasets used by other urban models such as transportation models and computation of new socioeconomic variables from equations from dwelling unit descriptions and based on assumptions about dwelling unit vacancy rates and persons per household. – Measures of relationships between different types of land use: e.g., job-housing balance for a region or for individual communities in a region. – "Microscopic level": possible to provide a user with a forecast for an area smaller than an acre. – Use of community performance indicators: evaluates alternative land use scenarios against local performance indicators developed by the community.
Limitations
– Requires mature GIS support capabilities, systems, and databases and a significant degree of GIS expertise on-staff. – Too complicated for a community to use without assistance from the model developer. – Steep initial learning curve, which may require much technical expertise and large resources to use this model. – Cumbersome in handling mixed uses and redevelopment due to considering a wide variety of land use types.

2.7.3 UPLAN

The UPLAN Urban Growth Model is a simple rule-based model developed by Johnston et al. (2003) based on a platform of ArcView GIS. Since input datasets for UPLAN are generally available and the model is easy to use by ArcView users, Johnston et al. suggest that MPOs and counties start with UPLAN and advance to a more complex model type as they gain expertise and gather more data.

UPLAN provides land use evaluations and change analyses based on general land-use plans, population and employment projections, characteristics of housing, and other user-defined conditions. UPLAN allocates increments of urban growth in user-specified discrete categories consumed in future years. User-specified discrete categories may be industrial, high-density commercial, low-density commercial, high-density residential, medium-density residential, and low-density residential, but are subject to change to match categories used in a MPO.

County or regional land consumption is calculated endogenously, based on user-specified assumptions, and population for the county or the entire region is projected. The user needs to input demographic and land use density factors that are converted to per hectare of land consumed for each land use. The user provides persons per household, percent of households in each density class, and average parcel size for each density class to determine the area of land needed for future housing. The area of land consumed for industry and commerce may be determined in a similar way based on workers per household, percent of workers in each employment class, and average land area per worker. These calculations yield a table of land demand for each land use type, from which the model operates its land allocation routine (Timmermans 2003).

Grid (raster) is adopted in the UPLAN because it takes less model runtime than a vector data model. Each grid cell roughly matches the development parcel size. Grid cells of 200 meters are used for low-density residential to represent the average parcel size (≈ 4 ha), while 50-meter grid cells (≈ 0.25 ha) are used for all other land uses. The two categories of the desired inputs are attraction grids and exclusion grids:

- Attraction Grids: freeway ramps, highways, major arterials, minor arterials, cities, passenger rail stations, and, for industrial allocation only, airports and ports.
- Exclusion Grids: land use plans, buffered rivers, buffered lakes, vernal pools (seasonal wetlands; buffered), floodplains, slopes, public lands, existing urban lands, permanent open spaces, and farmlands.

UPLAN has been applied to the Sacramento, CA region and Espanola region of New Mexico. Table 2.17 summarizes the strengths and limitations of UPLAN (U.S. EPA 2000).

Table 2.17 Strengths and Limitations of the UPLAN Model

Strengths
– Easy to use: allows users to prepare and evaluate alternative suitability, growth, and allocation scenarios with specific prompts generated by the program. – Customizable: allows the system to be customized to many different geographic areas and conditions. – Integrated: incorporates user-provided GIS and other data as a foundation and applies various evaluation/decision-tools (e.g., land use projection) to the underlying data. – The six default land-use types (industrial, commercial hi-density and low density, and three residential densities) permit the evaluation of the impacts of future land use patterns on runoff, water pollution, habitats, and costs from flooding and wildfires. Data grids may be as small as the data permit, generally of 25-meter squares.
Limitations
– Lack of sophisticated modeling: does not provide sophisticated modeling capability and/or theoretical basis to examine the interrelated factors of fiscal policies, and other planning decisions on the amount and type of future development and land use change that will occur. The attractiveness criteria are pseudo-economic, in that they represent land value and accessibility.

2.7.4 What If?

What If? is an interactive GIS-based planning support system developed in 1997 to support communities to create alternative visions for their area's future by mapping alternative development patterns determined by local land development policies. As suggested by its name, the model uses electronic spreadsheets to identify what would happen if a scenario's underlying assumptions are correct. The model provides a range of potential futures based on a range of alternative scenarios rather than a single "exact" prediction of the future.

What If? is designed to reflect the movement "planning with the public" instead of "planning for the public." The model provides a set of computer-based tools that support the open and continuous processes of community learning, debate, and compromise. In contrast to other models that are for closed and unsupervised "objective" analyses by technical experts, the model directly involves the public in the planning process. Modeling outputs are represented in easy-to-understand maps and reports to support community-based collaborative planning efforts (Klosterman 1999).

What If? projects the regional growth by aggregating the projected values of the uniform analysis zones (UAZs) in the region, a "bottom-up" approach. The UAZs are homogeneous polygons in terms of all variables in the model. For example, all points within a UAZ will have the same slope, be located in the same municipality, have the same zoning designation, be within the same distance of an existing or proposed highway, and so on. The UAZs may be generated by using GIS. What If? consists of three modules: Suitability Module, Growth Module, and Allocation Module. The Suitability Module determines land use suitability. The Growth Module projects future land-use demands. The Allocation Module allocates projected demands to the most suitable location (Klosterman 1999, U.S. EPA 2000).

General input to the What If? model includes natural features, infrastructure plans, existing land-use patterns, and approved comprehensive plans or zoning ordinances in GIS coverage, which are combined to create the UAZs. The following information is desirable for better analyses and may be entered into the system manually (U.S. EPA 2000):

- Growth projections for number of households, assumed vacancy and loss rates, assumed housing densities per land use, employment by type, assumed employment density
- Alternative development scenarios that are pre-defined by the community using What If?
- Land-use classifications that are pre-defined by the community using What If?
- Infrastructure plans that are pre-defined by the community using What If?

What If? has been used in three counties in Ohio: Hamilton, Medina, and Summit.

The strengths and limitations of What If? are summarized in Table 2.18 (U.S. EPA 2000).

Table 2.18 Strengths and Limitations of the What If? Model

Strengths
– Easy to use: allows users to prepare and evaluate suitability, growth, and allocation scenarios by solely using -Windows standard buttons, check boxes, and text boxes. – Integrated system: provides an integrated software package that incorporates user-provided GIS and other data as a foundation and applies various evaluation/decision tools to the underlying data. – Self-contained: requires no additional GIS or non-GIS software. – Flexible data requirements: fairly easy to use with minimum data requirements for existing land-use data to provide the basics for running What If? scenarios.
Limitations
– Lack of sophisticated modeling: no sophisticated modeling capability and/or theoretical basis to examine the interrelated factors of transportation infrastructure, fiscal policies, and other planning decisions on the amount and type of future development and land-use changes that occur. – No measures of spatial interaction. – Not behavior-based: No random utility or discrete choice theory to explain and project the behavior of various urban actors. Does not represent the interlinked markets for land, housing, nonresidential uses, labor and infrastructure, or provide any procedures for "market clearing" and price adjustment in the face of changes in demand and/or supply.

2.7.5 SLAM

The Simplified Land Allocation Model (SLAM) was developed by the Corradino Group in the early 1980's as part of the "Volusia Land Use Study" to provide future year ZDATA forecasts for the Florida Standard Urban Transportation Modeling System (FSUTMS) (Kaltenbach 2003). SLAM is coded in TransCAD and is a tool to help allocate the population and employment forecast in each township to individual TAZs by estimating the number of households, basic employment, service employment, retail employment, and other employment (Corradino 2004).

2.7.6 ULAM

The Urban Land Use Allocation Model (ULAM), developed by Transportation Planning Services, Inc., is a land use planning package that allocates future population and employment forecasted at the county level to traffic analysis zones (TAZs) (ULAM 2004). The model generates ZDATA1 and ZDATA2 files, which are input files for FSUTMS.

In addition to the future population and employment, zonal input data that are required by the model include the vacant buildable land acreage, allowable land use densities, the land of existing and approved development, population per dwelling units, percentage of vacant or seasonal units, auto ownership information, variables for the life style trip generation model, and the concurrency restrictions.

The model begins with allocating approved development to the vacant land and updates the vacant land data by subtracting land needed for approved development. Since the vacant land data may or may not include the vacant land required for approved development, the model provides an option to allow the user to choose whether to update the vacant land data or not. The model then combines the vacant land and the development index computed based on approved development, historical trends, and the market index. The future population and employment is allocated to TAZs based on the development index, considering the availability of vacant, buildable land and the concurrency restrictions.

The model provides a GIS interface that runs on an ArcView GIS platform and allows users to edit the input files and visualize the model output. The model has been used in many Florida counties including Bay, Leon, Citrus, Hernando, Pasco, Pinellas, Hillsborough, Charlotte, Indian River, St. Lucie, Martin, Palm Beach, and Broward.

2.8 Other Models

In this chapter, models that do not fall into any classifications described in the previous chapters are described. These models include a Markov model, INDEX, LUCAS, and Smart Places.

2.8.1 Markov Model of Residential Vacancy Transfer

The Markov Model of Residential Vacancy Transfer, developed by Emmi and Magnusson (1995), simulates the intersectoral transfer and absorption of vacant housing opportunities as a function of vacancy creations by using a matrix of vacancy transition probabilities. Vacancies are created by two housing events: housing demographic events and housing inventory change events. Vacancies created by housing demographic events include local migration, household death, and household dissolution. Vacancies created by housing inventory change events include new construction, conversion of single family units to multiple occupancy, conversion of non-residential and secondary residential units to primary residential occupancy, and reintroduction of units previously withdrawn from the building stock.

In the model, the urban housing market is assumed to be delineated from its "rural hinterland." The model also assumes that the number of vacancies created and transferred into a sector is equivalent to that transferred from and absorbed out of that sector. The model is also dependent on the assumptions of all Markov models (U.S. EPA 2000):

- Markovicity: The current state of the system depends on the immediately previous state and none earlier.
- Stationarity: Transition probabilities do not change over time.
- Homogeneity: All changes within a given sector are subject to a statistically identical set of transition probabilities.

The Markov model is used to measure the impacts of housing construction programs, out-migration, household death, and dissolution on intraurban residential relocation. This model is also used to simulate the impacts of new housing opportunities on the prospects for mobility among various population subgroups. The model is useful for planning new residential zoning and development based on existing demographics and population pressures, or to identify where certain residential sectors or areas might decline without coordinated efforts to accommodate demographic changes. The model is particularly useful for small towns and cities on the metropolitan fringe. Application sites of Markov model include three Swedish cities and 42 U.S. metropolitan areas.

Historical and current residential addresses are required for the model's input, which needs to be classified into "internally homogeneous" housing sectors (e.g., single-family residences, retiree apartment complexes, single-parent public housing, etc.) within a study area. The model also requires the past and present addresses of the households. This information may be derived from either sequential census records or a survey of recent household creations, conclusions, and moves.

The model's outputs include the following:

(1) A simulation of intra-urban household moves between residential sectors across a census or projection period;
(2) Probabilities of residential mobility by household type as a function of the sectoral distribution of vacancy initiations, the pattern of housing sector interaction and the sectoral distribution of households; and
(3) A measure of housing sector interaction in terms of the probability of a vacancy introduced into sector "I" being associated with a residential move in sector "J".

The strengths and limitations of the Markov Model of Residential Vacancy Transfer are presented in Table 2.19 (U.S. EPA 2000).

Table 2.19 Strengths and Limitations of the Markov Model

Strengths
– Simulates impacts of new vacancies on urban residential relocations and the accommodation of new entrants (immigrants and newly formed households) into the housing market. – Simulates impacts of newly created vacancies on the residential mobility for various urban sub-groups (e.g., single professionals, young families, and wealthy empty nesters). – High level of projection accuracy.
Limitations
– Depends on a stable, semi-closed system of residential moves between census years. – Does not explicitly simulate land-use changes. – Examines only discrete sectors of the residential housing market.

2.8.2 INDEX

INDEX is a GIS-based planning support system developed by Criterion Planners/Engineers in 1994 to measure the characteristics and performance of land-use plans and urban designs with "indicators" derived from community goals and policies (e.g., measuring the degree of transit orientation in a proposed residential subdivision). In the model, indicators are used to benchmark existing conditions, evaluate alternative courses of action, and monitor change over time.

INDEX is primarily designed for static time scale applications that focus on a built environment's measurements at the regional, community, and neighborhood levels, as opposed to operational integrated urban models, such as UrbanSim and TRANUS, which are predictive tools for a detailed dynamic analysis of complex urban systems. INDEX is capable of modeling dynamic land-use/transportation changes over time. However, it is only suitable for "what if" sketch planning because the function that models the dynamics of the urban system depends on a simplified gravity method of spatial growth allocation that does not consider land economics (Allen 2001).

INDEX consists of nine modules including areas, studies, cases, elements, create, score, compare, visualize, and link. Figure 2.4 shows a structure of these modules and features associated with modules. Modules of INDEX are described below (Allen 2001).

- The Areas module provides a geographic and topical hierarchy for organizing and accessing software runs.
- The Studies module defines studies classified by geographic area, such as neighborhood or district, and by topical focus, such as housing infill or transit-supportive development.
- The Cases module is used to create planning scenarios or cases. Cases may describe real or proposed conditions in a study area.
- The Elements module contains the database used to support indicator calculations, including land-use, housing, employment, recreation, environment, transportation, and infrastructure.
- The Indicators module offers a user a menu of indicators.
- The Rating and Weighting module provides an option for stakeholders to weight the importance of indicators relative to each other and to establish acceptability ratings for indicator scores.
- The Case Comparison module compares cases using charts and maps to highlight their differences.
- The Visualization Tools and Internet Links are used to visually communicate proposals and outcomes to stakeholders.
- The Customize allows accommodations for changing community conditions and stakeholder interests; this module allows users to add new indicators as needed over time.

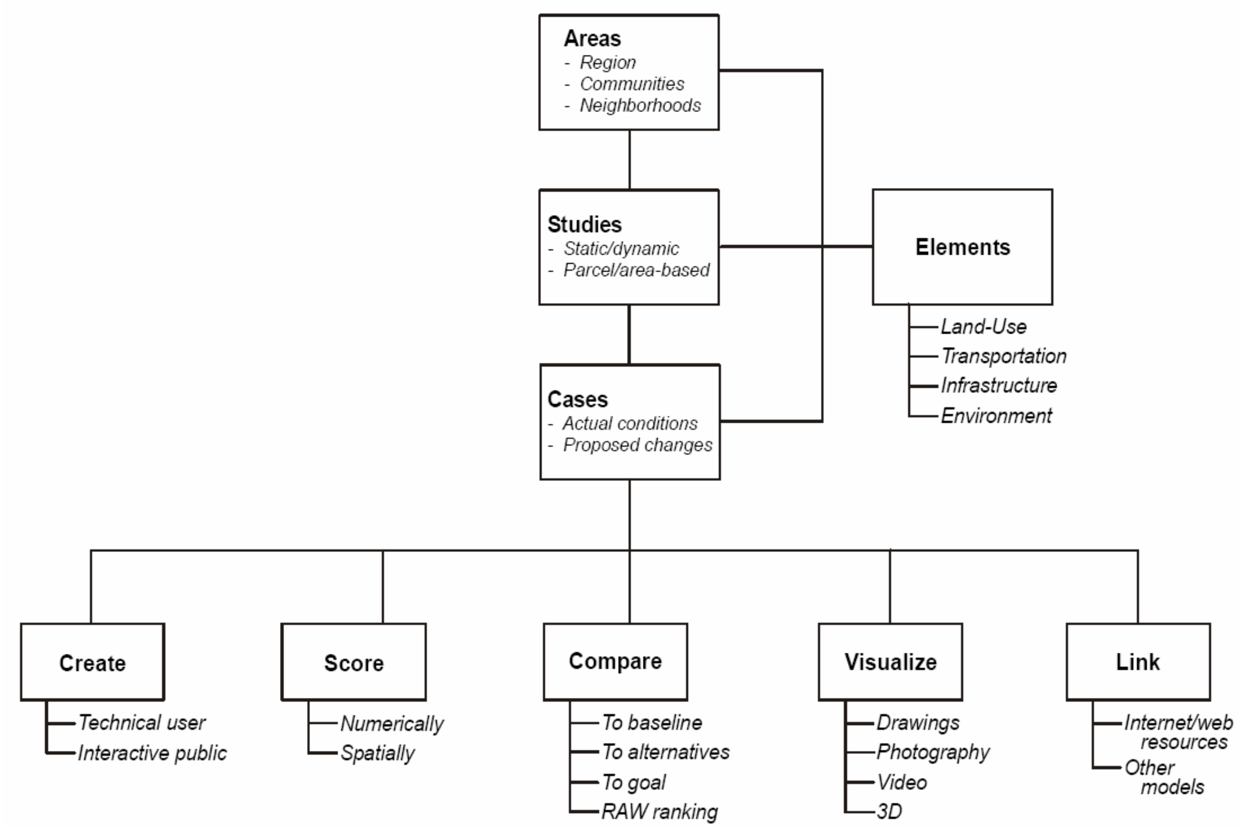


Figure 2.4 Generalized Structure and Functions of INDEX (Source: Allen 2001)

The outputs from INDEX are provided in an ArcView shape file format and Access database and include: jobs/housing ratio, street connectivity, residential density, transit orientation, employment density, parking supply, land-use mix, imperviousness, proximity to community amenities, pedestrian route directness, residential water consumption, park and open space availability, criteria air pollutant emissions, and greenhouse gas emissions. The number of land use categories in each community's unique land-use planning system determines the number of land-use categories addressed by INDEX.

The number and type of inputs required to run INDEX depends on each community and may include parcels, street centerlines, land-uses, dwelling units by type, employment by type, transit routes and stops, sidewalks, bicycle routes, off-street parking areas, building footprints, and significant environmental features.

Since its introduction in 1994, INDEX has been licensed to over 70 organizations in 25 states. Approximately half of these users are city and county planning departments, 25% are regional planning agencies, and the remaining 25% are divided among state and federal agencies, advocacy organizations, and academic institutions.

The strengths and limitations of the INDEX model are presented in Table 2.20 (U.S. EPA 2000).

Table 2.20 Strengths and Limitations of the INDEX Model

Strengths
– Each copy is customized for a community's unique set of conditions and priorities. – Integrates the explanatory power of GIS mapping with a comprehensive set of urban impact measurements. – Provides communities with a consistent, efficient tool for evaluating incremental development proposals and monitoring the implementation of long-range land-use plans.
Limitations
– Requires detailed GIS data and user expertise. – Must be in tandem with local four-step travel demand models in order to provide comprehensive land-use/transportation impact estimates. – The land-use plan or urban design being evaluated must be created exogenously.

2.8.3 LUCAS

The Land-Use Change Analysis System (LUCAS) was developed in 1994 to study the effects of land use on landscape structures in regions such as the Little Tennessee River basin in western North Carolina and the Olympic Peninsula of Washington state. LUCAS is a prototype computer application specifically designed to examine the impact of human activities on land use and the subsequent impacts on environmental and natural resource sustainability. The main functions of LUCAS include storing, displaying, and analyzing map layers derived from remotely-sensed images, census and ownership maps, topographical maps, and outputs from econometric models using the Geographic Resources Analysis Support System (GRASS), a public-domain GIS (Berry 1996).

LUCAS was developed under the premise that landscape properties, such as fragmentation, connectivity, spatial dynamics, and the degree of dominance of habitat types, are influenced by market processes, human institutions, landowner knowledge, and ecological processes. The structure adopted for LUCAS consists of three subject modules: the socioeconomic model, the landscape-change model, and the impacts model. The socioeconomic model derives transition probabilities for changes in land cover. These probabilities are computed based on variables such as transportation costs, slope and elevation, ownership, land cover, and population density. The landscape-change model produces the land-cover map, based on which the impact model estimates the impacts on selected environmental and resource-supply variables. The environmental variables are the amount and spatial arrangement of habitats for selected species and changes in water quality caused by human land use. Potential resource-supply variables include timber yields and real estate values.

The map layers used by LUCAS are derived from remotely-sensed images, census and ownership maps, topographical maps, and outputs from econometric models. LUCAS generates new maps of land cover representing the amount of land-cover change so that issues such as biodiversity conservation, assessment of the importance of landscape elements to meet conservation goals, and long-term landscape integrity can be addressed. The input data needed to operate LUCAS include transportation networks (access and transportation costs), slope and elevation (indicators of land-use potential), ownership (land holder characteristics), land cover (vegetation), and population density (Berry 1996).

Table 2.21 summarizes the strengths and limitations of the LUCAS model (U.S. EPA 2000).

Table 2.21 Strengths and Limitations of the LUCAS Model

Strengths
– Provides a graphical user interface that is intuitive and easily understood by users with a wide range of technical abilities and experience. – Provides a flexible and interactive computing environment for landscape management studies.
Limitations
– Intended to be used by a researcher working with a resource manager. – As a non-commercial GIS package, many bugs still exist in the GRASS software. Some of the features of GRASS are not well-documented. – Requires training and experience to calibrate. It is not a commercial, off-the-shelf product.

2.8.4 Smart Places

Smart Places was developed by the Electric Power Research Institute (EPRI) based on ESRI's ArcView software. Smart Places is an interactive computer tool that provides decision-making insights for target marketing, economic development, land-use planning, transportation systems, facilities management, environmental remediation and protection, energy forecasting, water allocation, and resource control. Smart Places offers innovation in the planning process through the exploration, design, modification, illustration, and evaluation of alternative planning scenarios.

Advanced features found in Smart Places include: selectable sophistication levels, automatic calculation of attributes for new and existing geographic features, automatic constraint compliance that checks new and existing geographic features, user modifications of attribute calculation formulae and constraint target values, storage and retrieval of entire design scenarios, comparisons of multiple land use scenarios, configurable links to empirical resource analysis models, and automatic results for visualization and report generation. Smart Places is designed for customization for specific locations and may be applied to small rural towns, as well as large urban areas.

The input data required to run the model depends on the goals and objectives of the user. Generally, information on natural features, infrastructure plans, existing land use patterns, and approved comprehensive plans or zoning ordinances provided in the ArcView shape file format or data file format are needed. Smart Places generates reports and ArcView shape files for the results of evaluations that compare - scenario results with assigned goals and boundaries for each evaluation parameter.

Smart Places has been used in Denver, CO and licensed to 38 other sites. Table 2.22 summarizes the strengths and limitations of Smart Places (U.S. EPA 2000).

Table 2.22 Strengths and Limitations of the Smart Places Model

Strengths
– Easy to use: the graphical user interface supports four distinct user levels, from new users to programmers/developers. Users may prepare and evaluate alternative land use scenarios using simple, narrative, pull-down menus. – Customizable: intended for customization by the user, allowing the system to model specific locations and design objectives. – System integration: provides the user the ability to link to models written in the ArcView script language Avenue or other programming languages, including C++ and Visual Basic.
Limitations
– Not a self-contained system: is an extension of ESRI's ArcView GIS. Users must have ArcView to use Smart Places. – Lack of sophisticated modeling: provided as an open system that allows users to evaluate the impacts of alternative development scenarios. The users must provide data inputs and evaluation models. – Expertise: background in GIS is desirable for all user levels beyond "new users."

2.8.5 Summary

Land use models, which are classified into eight groups, have been described. The eight groups are spatial interaction models, spatial input-output models, linear programming models, micro-simulation models, random utility/discrete choice models, cellular automaton models, rule-based models, among others.

Micro-simulation is a newer approach that models the decisions of an individual or firm in the real world regarding the choice for a location. Examples of micro-simulation models include NBER/HUDS, MASTER, IRPUD, and UrbanSim. Due to advancements in computer technology and the availability of detailed data, micro-simulation has become an increasingly popular analysis tool. Discrete choice models are designed to model an individual's choices, taking into account the characteristics of that choice, as well as the characteristics of the individual. Discrete choice modeling techniques have long been used in travel demand modeling, especially in the analysis of mode choice, and in land use planning when modeling location choice decisions of households and firms. By adopting Daniel McFadden's random utility theory, discrete choice models have a firm foundation within econometrics and become standard methods in developing models for predicting individual choices among a finite set of alternatives.

In cellular automata based urban models such as LEUTH, cells simulate four settlements such as trade, industrial, residential and empty areas, and they interact among themselves. The strength of the interactions may vary based on the distances between them. The applications of this approach are available mainly for research purposes rather than operational planning or policy purposes.

Rule-based land use models are useful tools for MPOs and counties to test long-range scenarios, mainly because they are easy to apply and the data required to operate the models are generally available. Their main weakness is the inability to fully model the market and real estate mechanism underlying the land use changes. Several land use models that implement a rule-based set of procedures to allocate population, employment, and/or land use include the CUF model (CUF-1 and CUF-2), SAM, UPLAN, What If?, SLAM, and ULAM.

Table 2.23 provides a comparative matrix of the different models reviewed in this report, based on selected attributes including:

- Whether the model integrates land use and transportation;
- Ease of use;
- Availability of a graphic user interface;
- GIS capability;
- Whether the model is footed in economic theories;
- Modeling of demographic change in population;
- Modeling of market mechanism;
- Consideration of income;
- Learning curve;
- Proprietary software (as opposed to public domain software);
- Availability of technical support; and
- Continued update.

Table 2.24 also provides a summary of the models' features regarding the model type, methodologies implemented, and model output.

After reviewing various land use models, Miller (Miller et al. 1998) and Dowling (Dowling et al. 2000) concluded that UrbanSim is the best model for the following reasons: spatial disaggregation (use of parcels to model land development); temporal aggregation (use of one-year time steps); dynamics (disequilibrium model); detailed disaggregations of households and firms; and use of activity-based travel models.

Dowling et al. (2000) point out that UrbanSim is a public domain program and is accessible to a broad group of practitioners who can contribute to the improvement of UrbanSim. Thus, UrbanSim is useful for evaluating alternative polices and management strategies, as well as in assessing the application of instruments to achieve these policies or strategies.

Table 2.23 Summary of Features of Existing Land Use Models

	Integration	Ease of Use	Graphic Interface	GIS Capability	Economic Theory Base	Demographic Change	Market Mechanism	Income	Learning Curve	Proprietary	Technical Support	Continued Update
DRAM/EMPAL/ ITLUP/METROPILUS	N	N	Y[2]	Y[3]	N	N	N	Y	Y	Y	Y	Y
LILT	Y	N	N	N	N	N	N	N	Y	N	N	N
HLFM II+	Y	Y	Y	N	N	N	N	Y	N	Y	N	N
LUTRIM	Y	Y	N	N	N	N	N	Y	N	N	N	N
MEPLAN	Y	N	Y	Y	Y	N	Y	Y	N	Y	Y	Y
TRANUS	Y	N	Y	Y	Y	N	Y	Y	N	Y	Y	Y
DELTA	N	N	N	N	Y	Y	Y	Y	Y	C	Y	Y
Herbert-Stevens Linear Programming Model	N	Y	N	N	N	N	Y	Y	N	N	N	N
TOPAZ/TOPMET	Y	Y	N	N	Y	N	Y	N	N	U	U	Y
POLIS	Y	Y	N	N	Y	N	Y	N	N	N	U	U
NBER/HUDS	N	U	U	N	Y	Y	Y	Y	U	U	U	U
MASTER	Y	U	U	U	Y	Y	Y	Y	U	U	U	U
IRPUD	Y	N	Y	Y	N	Y	Y	Y	Y	C	Y	Y

[2] Only METROPILUS provides Graphic Interface.
[3] Only METROPILUS has GIS Capabilities.

	Integration	Ease of Use	Graphic Interface	GIS Capability	Economic Theory Base	Demographic Change	Market Mechanism	Income	Learning Curve	Proprietary	Technical Support	Continued Update
UrbanSim	N	N	Y	N	Y	Y	Y	Y	Y	N	Y	Y
METROSIM	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y
SLEUTH	N	N	Y	N	N	N	N	N	Y	N	N	Y
CUFM: CUF-1	N	Y	N	N	Y	N	Y	N	N	C	N	Y
CUFM: CUF-2	N	Y	Y	Y	Y	N	Y	N	N	C	N	Y
SAM/LAM/SAM-IM	N	Y	Y	Y	Y	Y	N	N	Y	Y	Y	N
UPLAN	N	Y	Y	Y	N	Y	N	N	Y	N	N	Y
What If?	N	Y	Y	Y	N	N	N	N	N	Y	Y	Y
Markov Model of Residential Vacancy Transfer	N	Y	N	N	N	Y	N	Y	Y	N	N	N
INDEX	Y	N	Y	Y	N	N	N	N	Y	Y	Y	Y
LUCAS	N	N	Y	Y	N	N	Y	N	Y	N	N	N
Smart Places	N	Y	Y	Y	Y	N	N	N	N	C	N	Y
ULAM	N	Y	Y	Y	N	N	Y	N	Y	Y	Y	Y

Notes: Y – Yes
N – No
C – Contact developer
U – Unknown

Table 2.24 Summary of Main Features of Existing Land Use Models

Models	Model Type	Methodologies	Output
DRAM/EMPAL/ ITLUP/METROPILUS	• Urban Statistical • Spatial Interaction • Aggregate Logit	• Multinominal Logit • Regression	• EMPAL employment projections in each zone by economic sector • Number of households projected in each zone by income level or any other user-defined level • Land consumption projections in each zone
LILT	• Spatial Interaction • Travel Demand • Entropy-Maximizing Principles	• Regression	• Allocation of future population, new housing units, and jobs
HLFM II+	• Entropy-Maximizing Principles	• Matrix Adjustment Methods • Conditional probabilities	• Forecasts of employment and population by zone
LUTRIM	• Spatial Interaction	• Regression	• Forecasts of households and basic and household-serving employment
MEPLAN	• Travel Demand • Urban Economic/Land Use Market • Hedonic	• Multinominal Logit • Network Analysis	• Employment by sector • Population by income group • Households by car ownership group • Land area by activity • Floor space by activity • Price by floor space/land type
TRANUS	• GIS • Urban Impact • Travel Demand • Urban Economic/Land Use Market • Hedonic	• Causal Inference • Multinominal Logit • Network Analysis • Time-Series • Discrete Choice Analysis • Decision Theory • Random Utility Theory • Input-Output Analysis • Algorithms	• All paths between each O-D pair for each travel mode and combination of modes • General assignment results for each link • Detailed assignment results for each link • Indicators of the performance of the transportation system • Results from the transportation model • Transit route profile, with demand-supply information for each route on each link • Activity location and land use consumption outputs
DELTA	• Urban Economic/Land Use Market	• Markov Chains • Multinominal Logit • Cobb-Douglas Utility Functions • Elasticity-based Responses • Matrix Adjustment Methods	• Forecasts of numbers of households by type and location • Jobs by sector and location • Floor space by category and location

Herbert-Stevens Linear Programming Model	• Linear Programming Model	• Objective Functions	• Allocation of households
TOPAZ/TOPMET	• Linear Programming Model	• Non-linear Objective Functions	• Employment per sector, total population, land allocation per activity, vacant land and the location of houses • Trips per mode and per journey type, travel energy used per origin and purpose and (optional) air pollution • Planning indicators (accessibility per zone per trip type and per mode) • Economic indicators (developers cost, travel costs per mode and per trip type and marginal cost of incrementing activity levels)
POLIS	• Linear Programming Model	• Random Utility Theory	• Distribution of population, employment, buildings, land use, and transportation
NBER/HUDS	• Micro-Simulation	•	• Evaluation of the impacts of housing improvement programs
MASTER	• Travel Demand • Micro-Simulation	• Monte Carlo Simulation	• Population growth and household structure • Household members' choice of jobs, employment and becoming unemployed, retirement, education level, sex, social group of head of household, job vacancies, and salary ranges

IRPUD	• Travel Demand • Urban Economic/Land Use Market	• Markov Chains • Multinominal Logit • Microsimulation • Variant of Inclusive Value Method	• Percent foreign population • Trips by trip purpose (work, shopping, education, other) • Percent population 0–5, 6–14, 15–29, 30–59, 60+ years • Trips by mode • Households • Mean travel time • Total employment • Mean travel cost • Non-service, service, retail employment • Car-km per capita per day • Unemployment rate • CO_2 emissions by car per capita per day • Job-labor ratio • CO_2 emissions by transport per capita per day • Total dwellings • Transport expenses per household per month • Percent single-family dwellings • Public transport expenses per household per month • Housing floor space per capita • Car ownership (cars per 1,000 population) • Mean housing rent per square mile
UrbanSim	• Random Utility Logit • Urban Economic/Land Use Market • GIS • Hedonic	• Expert Systems • Multinominal Logit • Random Utility Theory • Regression • Monte Carlo Simulation	• Households by type (income, size, age of head, children, workers) and zone • Businesses and employment by type (sector) and zone • Acres by land use and zone • Housing units and building square footage by type and zone • Prices of land, housing and commercial space by type and zone • Simulated development by type and zone

METROSIM	• Travel Demand • Markov Chain • Urban Economic/Land Use Market • Hedonic • Discrete Choice Method	• Markov Chains • Multinominal Logit • Network Analysis • Regression • Time-Series • Dynamic Economic General Equilibrium Analysis	• Basic industry distribution by zone and by type of basic industry • Non-basic industry distribution by zone and by type of non-basic industry • Residential real estate distribution by type and zone • Non-residential real estate distribution by type and zone • Vacant land distribution by type and zone • Households • Travel (commuting and non-work) • Traffic assignment on the network • Rents and market prices for each type of real estate by zone • Vacancy rates for each type of real estate
SLEUTH	• Cellular Automata	• Cellular Automata • Time-Series • Monte Carlo Imaging	• Snapshot of a particular year • Cumulative image that results from multiple runs and shows a probability of urbanization for a given year
CUF-1	• Urban Growth	• Regression	• Acreage tabulations and total • Maps of newly-developed areas
CUF-2	• Land Use Change	• Multinominal Logit • Regression	• New development an redevelopment acreage total by land use type • Maps of existing and projected development by land use type
SAM/LAM/SAM-IM	• GIS	• Cellular Automata • Multinominal Logit • Regression	• Distribution of future land uses
UPLAN	• GIS • Urban Impact	• Not Specified	• Grid Maps (attraction grids, exclusion grids, general plan grids, and existing urban grids) • Analysis Report • Assumptions Report • Image Files
What If?	• GIS	• Mapping	• Suitability Analysis Map • Suitability Analysis Results Report • Suitability Analysis Assumptions Report • Growth Analysis Results Report • Growth Analysis Assumptions Report • Allocation Map • Allocation Analysis Results Report • Allocation Analysis Assumptions Report

Markov Model of Residential Vacancy Transfer	• Markov Chain	• Linear Programming • Markov Chains/Transition Matrices • Multinominal Logit • Regression	• Projection of urban household moving between residential sectors • Probabilities of residential mobility by household type • Probability of a vacancy introduced into sector "I" being associated with a residential move in sector "J"
INDEX	• GIS • Urban Impact	• Causal Inference • Correlation • Linear Programming • Network Analysis • Time-Series	• Jobs/Housing ratio • Residential density • Employment density • Land-use mix • Proximity to community amenities • Residential water consumption • Criteria air pollutant emissions • Street connectivity • Transit orientation • Parking supply • Pedestrian route directness • Park and open space availability • Greenhouse gas emissions
LUCAS	• GIS	• Time-Series	• Area by land use type • Amount of edge by land use type • Edge/area ratio by land use type • Mean patch size by land use type • Number of patches by land use type • Cumulative frequency distribution of patches by size by land use type • Proportion of land cover by land use type • Amount of total edge • Standard deviation of patch size by land use type • Size of largest patch by land use type • Mean patch shapes by land use type
Smart Places	• GIS	• Causal Inference	• Effects on land-use patterns • Effects on travel demand, local government fiscal conditions, availability of open space, environmental quality, school quality, crime, and other quality-of-life conditions
ULAM	• Rule-Based Model	•	• Allocation of population and Employment • ZDATA files for FSUTMS

3. DATA PROCESSING

This chapter provides an overview of the data preparation and compilation for applying UrbanSim to Volusia County, Florida. A description of the study area and a detailed list of required data are provided. Extensive data collection efforts were made for this study because UrbanSim requires many kinds of data, including land use, parcel, tax appraiser record, environmental and political layers, etc.

After the data collection, the parcel data and the employment data required further preparation, such as address standardization and data imputation to improve data quality. A GIS program was developed to convert the data to grid cell data. Section 3.1 summarizes the preparation of each data set and the data sources. The UrbanSim's analysis units, which are grid cells, are also described. Since there is no inventory of household level data, it is necessary to synthesize the household data based on the census. A detailed description of household synthesis is provided. Travel data from the travel demand model are also presented. Input tables for UrbanSim are prepared based on these data sets and are imported into a MySQL database. Section 3.2 lists the input tables for UrbanSim and their descriptions.

3.1 Data Preparation

The following sections describe the data collection and pre-processing procedures to extract data for this study.

3.1.1 Study Area

To test the applicability of UrbanSim to Florida's urban areas and investigate model application issues, Volusia County is selected as the study area. The selection criteria include that the study area has up-to-date geographic information system (GIS) data, including parcel-level property data, and is relatively self-contained.

Volusia County covers an area of 1,263 square miles and had a total population of 443,343 in the year 2000. It faces the Atlantic Ocean to the east. The St. John's River flows through the county to the west, and there are many lakes along its west border. Although it is surrounded by Flagler, Marion, Lake, Seminole, and Brevard counties, it is relatively isolated from the impacts of developments in these surrounding counties due to their low level of urbanization. Volusia County also has GIS data and GIS parcel data, which are essential to calibrating an UrbanSim model. Figure 3.1 shows the location of Volusia County and the surrounding counties.

Figure 3.1 Volusia County and Its Surrounding Counties

3.1.2 Data Collection

The data for this study were collected from online sources and public agencies. The on-line resources include the Florida Geographic Data Library (FGDL) and the web sites of the Volusia County Geographic Information Services, St. Johns River Water Management District, U.S. Census Bureau, Federal Highway Administration, U.S. Geology Survey, U.S. Department of Agriculture, and Florida Department of Transportation. Public agencies include the Volusia County Metropolitan Planning Organization, Volusia County Property Appraiser's Office, and the Volusia County Geographic Information Services Department. Employment data were obtained from a proprietary database purchased by the FDOT.

Grid Cells

The current release of UrbanSim uses grid cells as the data aggregation and simulation unit. A GIS grid was created for Volusia County to convert parcel data, employment data, household data, and environmental and planning layers to this grid system. Each cell in the grid is 150 meters by 150 meters (492.1245 × 492.1245 square-feet or approximately 5.56 acres) in size and is assigned a unique identification number. The total number of grid cells generated is 145,363. Figure 3.2 shows the grid cell created for a part of Volusia County, overlaid with parcel and address layers.

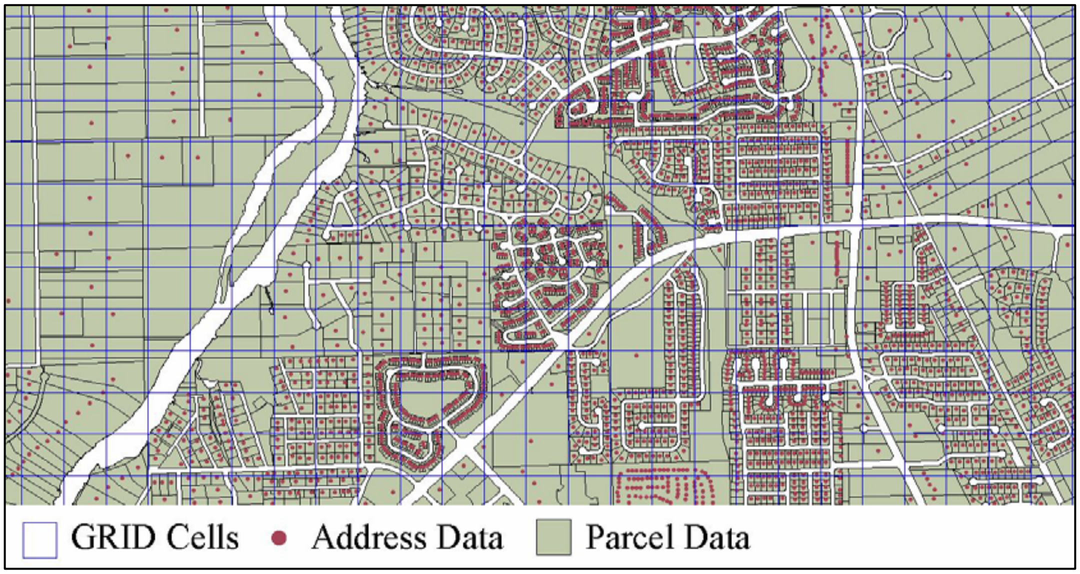


Figure 3.2 GRID Cells, Address Data, and Parcel Data for Part of Volusia County

Parcel Data

The base year for the UrbanSim model is designed to be 2000. The 2000 parcel data were obtained from the Geographic Information Services Department of Volusia County. The parcel data is a polygon coverage with 241,900 parcels. Each parcel has a unique parcel ID, designated as PID.

Address Data

The address data were downloaded from the Volusia County GIS website (http://www.volusia.org/gis/). The address database contains both address and PID information, with the former providing a link to the employment data and the latter to the parcel data, thus establishing a connection between the two datasets. Address data are provided as a point layer and include 213,264 points. There are fewer addresses than parcels. A possible reason for the difference in the number of records in the address and parcel databases may be that they were created in different years.

Property Tax Data

The 2000 property tax appraiser's database was obtained from the Volusia County Property Appraiser's Office. The data include land use code, lot size, property assessed value, taxable value, value of land, built year of the structure, total living area (or adjusted area) or usable area if non-residential, price corresponding to first or original purchase/sale of property, year and month corresponding to first or original purchase of property, name of owner, tax exemption types and amounts, etc.

The total number of records in the property tax database is 255,429. The records are in fixed-width text file format. The data were converted to dBase IV format to be joined with the parcel layer based on PID.

Employment Data

Employment data are available from state unemployment insurance records but are confidential and difficult to obtain and use. They are also available from proprietary sources such as InfoUSA, Dunn & Bradstreet, etc. This study uses the employment data purchased by the FDOT from InfoUSA. Information on the quality of the InfoUSA data may be found from an audit report that was commissioned by InfoUSA and conduced by Bass & Associates, available from the InfoUSA web site.

Multiple sources are utilized by InfoUSA to gather U.S. business data. These sources include (http://www.infousa.com/):

- 5,200 yellow page and business white page directories
- 17 Million phone calls to verify information (every business is called one to four times a year)
- County courthouse and secretary of state data
- Leading business magazines and newspapers
- Annual reports by companies
- Securities and Exchange Commission (SEC) 10-K and other filings
- New business registration and incorporations
- Postal service information including national change of address, ZIP+4 carrier route, and delivery sequence files

The Business information included in the employment data is listed below.

- Location – ZIP code, neighborhood, city, metro area, county, area code, and state
- Type of business – yellow page heading, major industry group, SIC code, or professionals (doctors, dentists, etc.)
- Business size – number of employees and sales volume
- Credit rating
- Location type – corporate headquarters, headquarters of a subsidiary, and branch
- Phone and fax numbers
- Key decision makers/executive names

According to the InfoUSA employment data, there were a total of 17,196 public entities or private establishments in Volusia County in 2000. Of all the records, 373 are missing employment size information. For such records, the average employment size for its sector is assumed. After assigning employment size, the total number of employees was found to be 178,864. Table 3.1 summarizes jobs based on industry sector. Major industries in Volusia County are retail trade (24.34%) and services (25.13%). Retail trade is subdivided into eating and drinking places and other retail trade, while service is subdivided into producer services, consumer services, and health services.

Table 3.1 Employee Composition of Sectors in Volusia County

Sector ID	Sector	Jobs	Percentage (%)
1	Resource	364	0.20
2	Construction	10,700	5.98
3	Manufacturing - Other	966	0.54
4	Manufacturing - Aviation	15,405	8.61
5	Transportation	4,213	2.36
6	Communications, Electric, Gas, and Sanitary Services	2,390	1.34
7	Wholesale Trade	10,450	5.84
8	Eating and Drinking Places	18,380	10.28
9	Other Retail Trade	25,145	14.06
10	Finance, Insurance, and Real Estate	9,670	5.41
11	Producer Services	10,135	5.67
12	Consumer Services	19,606	10.96
13	Health Services	15,211	8.50
14	Education	11,678	6.53
15	Public Administration	24,551	13.72
	Total	178,864	100.00

The locations of businesses are coded as a point layer, which was created by geocoding businesses' addresses on a street network. Since geocoding locates an address based on the centerline of a street, which may not be the actual location of the property at that address, employment data were matched with parcel data for use during model development. Since UrbanSim locates new or moving employment to a grid cell by accounting for the occupancy of properties, it is important to have the employment data associated with the parcel data. This match between businesses and the properties on which they are located also makes it possible to distinguish home-based employment from non-home-based employment. Home-based employment is located in a residential unit, while non-home-based employment is in a non-residential unit.

Environmental Layers

Environmental layers reflect the environmental characteristics in Volusia County. They include:

- Water
- Wetlands
- Floodplains
- Parks and open space
- National forests
- Steep slopes

- Stream buffers (riparian areas)

Each of the above layers is explained in more detail below.

Water

A hydrology polygon coverage was downloaded from the Volusia County GIS website. The hydrology layer contains features representing islands, drainage easement, canals, rivers, lakes, reservoirs, borrow pits, and the Atlantic Ocean. The types of hydrology features are indicated in the field *hyd_code* in the attribute table of the hydrology layer. Table 3.2 provides the hydrology code and its description. A new layer was constructed by extracting from this layer features representing drainage easement, lakes/ponds, reservoirs, borrow pits, and the ocean.

Table 3.2 Description of Hydrology Code

Hydrology Code	Description
0	Islands
514	Drainage Easement
515	Canals
519	Rivers
520	Lakes/Ponds
530	Reservoirs
540	Borrow Pit
550	Oceans

The new water layer was compared with the Major River layer downloaded from the FGDL website. Modifications were made on boundaries of Lakes George, Monroe, and Harney. The water layer is shown with stream buffers and wetland in Figure 3.3.

Stream Buffers

Stream buffers were created based on the canals and rivers in the Hydrology layer and were modified based on Major River layer from the FGDL. Modifications were made to the Tomoka River and Spruce Creek to correct their shapes. Stream buffers are also depicted in Figure 3.3.

Wetlands

The wetland layer was derived from the vegetation layer downloaded from the Volusia County website. Vegetation is a polygon layer and contains land use codes, as listed in Table 3.3. The features with land use codes ranging from 611 to 650 were selected to form a new wetland layer. The wetlands in Volusia County are illustrated in Figure 3.3.

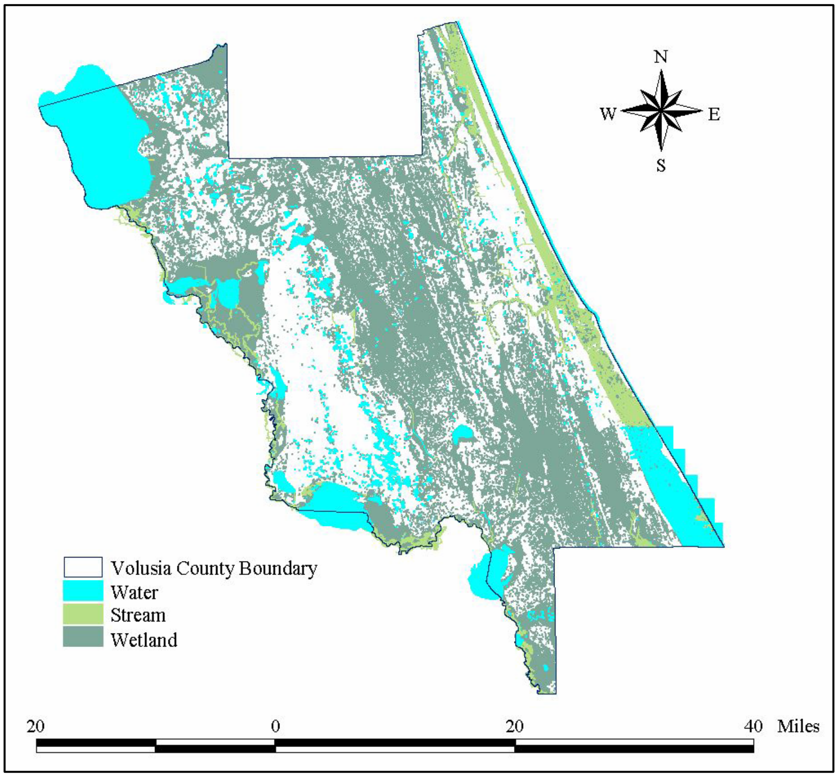


Figure 3.3 Water, Stream, and Wetland in Volusia County

Table 3.3 Description of Land Use Code

Land Use Code	Description	Land Use Code	Description
0	Incorporated or other	431	Beech-Magnolia
100	Generalized Urban	432	Scrub Oak
181	Beaches	434	Hardwood/Conifer Mix
200	Generalized Agriculture	440	Tree Plantation
270	Abandoned Fields	451	Red Cedar
311	Coastal Dune Series	452	Larch Oak Hammock
322	Coastal Scrub	500	Open Water
329	Shrub/Disturbed Wetlands	611	Bay Swamp
330	Mixed Rangeland	612	Mangrove
411	Pine Flatwood	615	Bottomland Swamp
412	Pine/Xeric Oak	616	Inland Pond
413	Sand Pine	617	Mixed Wetland Hardwoods
414	Pine/Mesic Oak	621	Cypress Swamp
415	Longleaf Pine/Sandhill	623	Atlantic White Cedar
419	Other Pine/Special	624	Cypress/Pine/Cabbage Palm
421	Xeric Oak	641	Freshwater Marsh
424	Melaleuca	642	Estuarine Marsh
425	Temperate Hammocks	643	Wet Prairie
427	Live Oak Hammock	650	Non-Vegetated Wetland
428	Cabbage Palm Hammock	740	Disturbed Lands
429	Wax Myrtle/Willow		

Floodplains

The Q3 flood data, downloaded from the St. Johns River Water Management District (http://sjrwmd.com/programs/data.html), were derived from the Flood Insurance Rate Maps (FIRMs) published by the Federal Emergency Management Agency (FEMA). The original layer was georeferenced to UTM zone 17, Datum 1983 and was re-projected to State Plane Florida East, Datum 1983. Figure 3.4 shows the floodplain in Volusia County. The FIRMs provide the basis for floodplain management, mitigation, and insurance activities for the National Flood Insurance Program (NFIP) and identify Special Flood Hazard Areas (SFHAs).

The values in the SFHA field of the attribute table indicate whether an area is inside or outside of the SFHAs. Flood hazard zones are defined in Table 3.4. An area designated as within a SFHA may be inundated by a 100-year flood, for which either Base Flood Elevations (BFEs) or velocity may have been determined. A SFHA may be classified as Zones A, AE, AO, AH, A99, AR, V, or VE. An area designated as outside a "special flood hazard area" includes Zones X or X500.

Table 3.4 Flood Hazard Zone Designation

Zone	Definition
100IC	An area where the 1% annual chance flooding is contained within the channel banks and the channel is too narrow to show to scale.
A	An area inundated by 1% annual chance flooding, for which no BFEs have been determined.
AE	An area inundated by 1% annual chance flooding, for which BFEs have been determined.
AH	An area inundated by 1% annual chance flooding (usually an area of pond), for which BFEs have been determined; flood depths range from 1 to 3 feet.
ANI	An area that is located within a community or county that is not mapped on any published FIRM.
AO	An area inundated by 1% annual chance flooding (usually sheet flow on sloping terrain), for which average depths have been determined; flood depths range from 1 to 3 feet.
D	An area of undetermined but possible flood hazards.
UNDES	A body of open water, such as a pond, lake, ocean, etc., located within a community's jurisdictional limits that have no defined flood hazard.
VE	An area inundated by 1% annual chance flooding with velocity hazard (wave action); BFEs have been determined.
X	An area that is determined to be outside the 1% and 0.2% annual chance floodplains.
X500	An area inundated by 0.2% annual chance flooding; an area inundated by 1% annual chance flooding with average depths of less than 1 foot or with drainage areas less than 1 square mile; or an area protected by levees from 1% annual chance flooding.

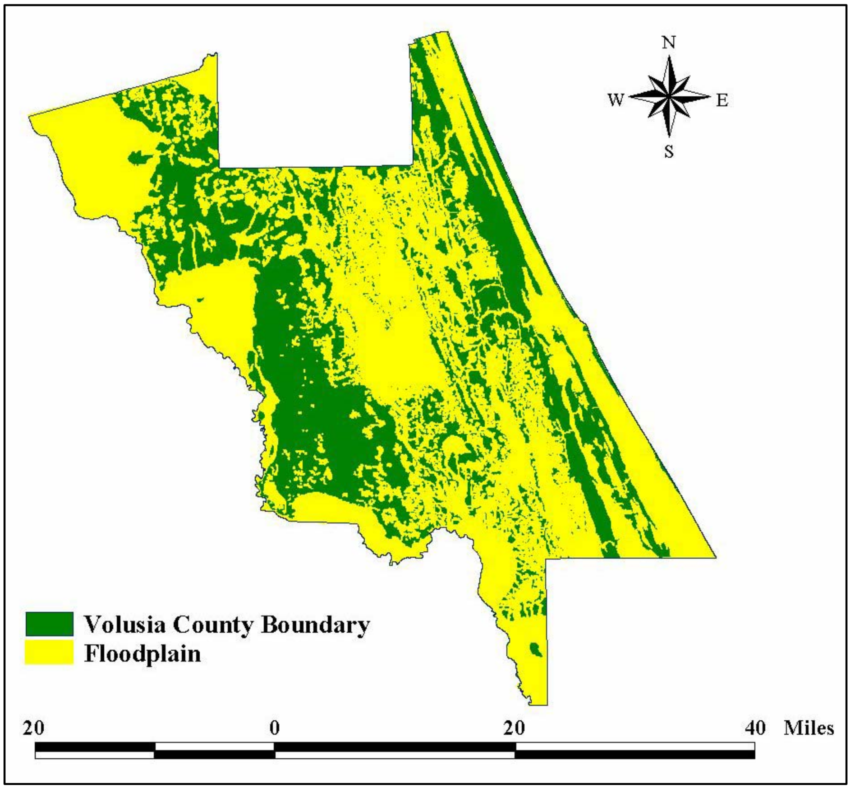


Figure 3.4 Floodplain in Volusia County

Parks and Open Space

Three GIS datasets of the Florida Natural Areas Inventory (FNAI), State Parks, and County Parks were downloaded from the FGDL and the Volusia County GIS website. The FNAI data provide information mostly on public lands identified as having a natural resource value and managed at least partially for conservation. The State Park and County Park data contain geographic boundaries and associated information. The National Forests geographic boundary was extracted from the FNAI data. However, none of the national forests were located inside Volusia County. Figure 3.5 depicts the conservation areas, natural areas, national forests and seashores, state forests and parks, local parks, wildlife refuges, and historical sites.

Slopes

The Digital Elevation Model (DEM) grids were downloaded from the St. Johns River Water Management District GIS website (http://www.sjrwmd.com/programs/plan_monitor/gis). The Spatial Analysis Tool was used to calculate the slope of the DEM. The slope of each cell was defined as the maximum rate of change from the cell to its neighboring cells in degrees. The slope was calculated by using the 'Derive Slope', one of the functions available in the Spatial Analysis Tool in ArcView, and was converted to a grid.

Cells in the slope grid were selected to generate a steep slope grid. For this study, the steep slope was defined as a slope with more than 19.3 degrees, which is equivalent to a 35 percent slope. Figure 3.6 displays the DEM in feet and the slope in percentage for Volusia County.

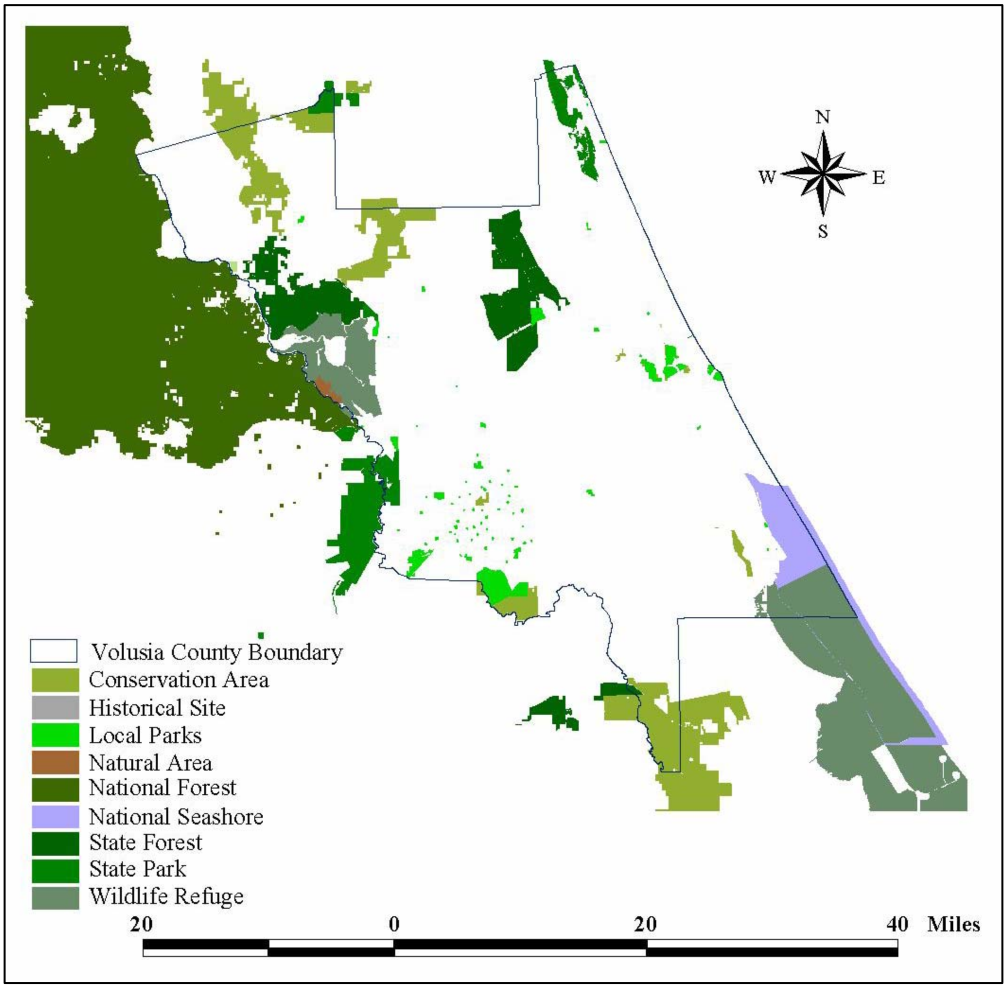
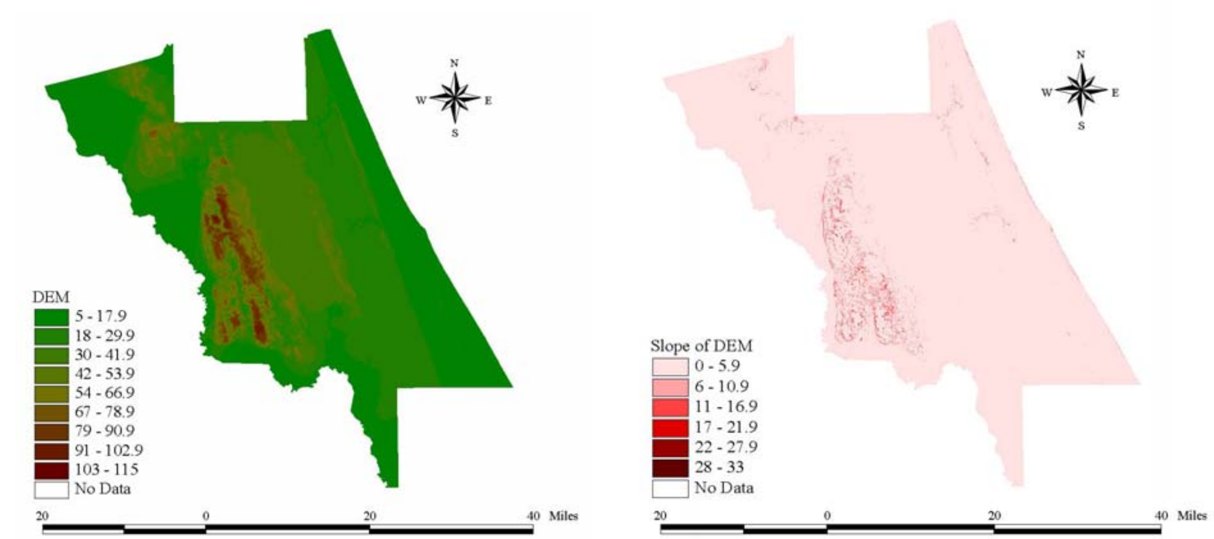


Figure 3.5 Parks and Open Space in Volusia County



(a) DEM (b) Slope

Figure 3.6 Digital Elevation Model (DEM) and Slope

Planning and Political Layers

Beside the environmental layers, the following planning and political layers are required:

- Streets
- Traffic Analysis Zones (TAZs)
- Cities
- Urban growth boundaries
- Military
- Major public lands
- Tribal lands

Tribal land data may be downloaded from the Census Bureau website for cartographic boundary files (http://www.census.gov/geo/www/cob/bdy_files.html,) and military land data from the FGDL. However, there were neither tribal lands nor military bases in Volusia County. The rest of the data sets are described below.

Streets

A GIS line layer, *allstreet*, was downloaded from the Volusia County website. The highway network and local street network were extracted from this layer. The attribute table of the *allstreet* layer contains the street code that defines road types, as shown in Table 3.5. Figure 3.7 illustrates the freeways, expressways, arterials, and local streets.

Table 3.5 Description of Street Code

STRCODE	Description
1	Interstate Highway
2	U.S. Highway
3	State Highway
4	County Road
5	Water Feature
6	Incorporated Local Road
7	Unincorporated Local Road
8	County Boundary
9	Unincorporated Major Road
10	Incorporated Major Road
11	Shoreline (Atlantic)
12	Railroads
13	Airfields

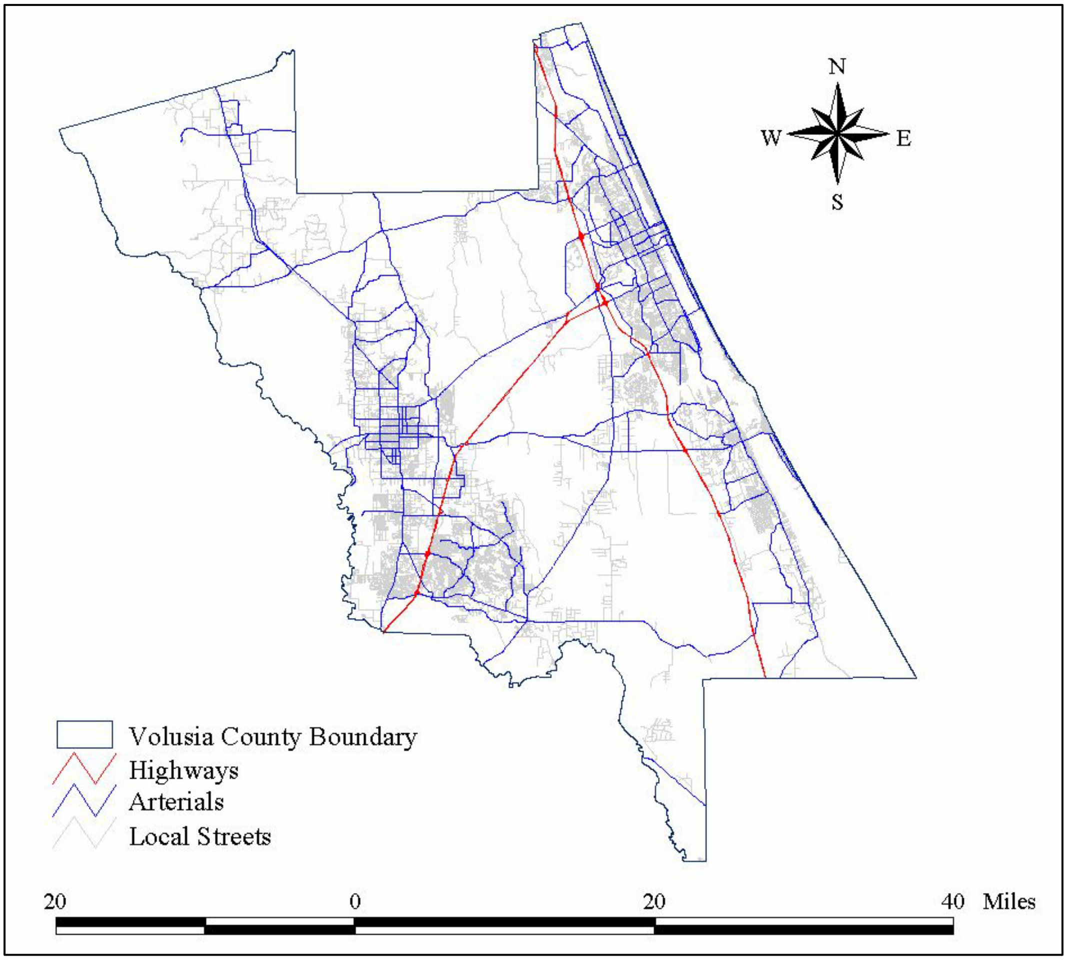


Figure 3.7 Highways, Arterials, and Local Streets

Traffic Analysis Zones Layer

The TAZs are used as the basic geographic units in analyses related to transportation, as well as in the Florida Standardized Urban Transportation Modeling Structure (FSUTMS) models. The TAZ layer of 1997, which was used in the Volusia County 2020 LRTP, was obtained from Volusia County to summarize the demographic and socioeconomic data, such as population and employment.

Since UrbanSim runs separately from the travel demand model, which is the Volusia County 1997 validated base year travel demand model in this study, consistency must be maintained between the input files of UrbanSim and the ZDATA files for the travel demand model. The ZDATA files identify 760 TAZs and 95 dummy zones. Figure 3.8 shows the 1997 TAZ boundaries of Volusia County.

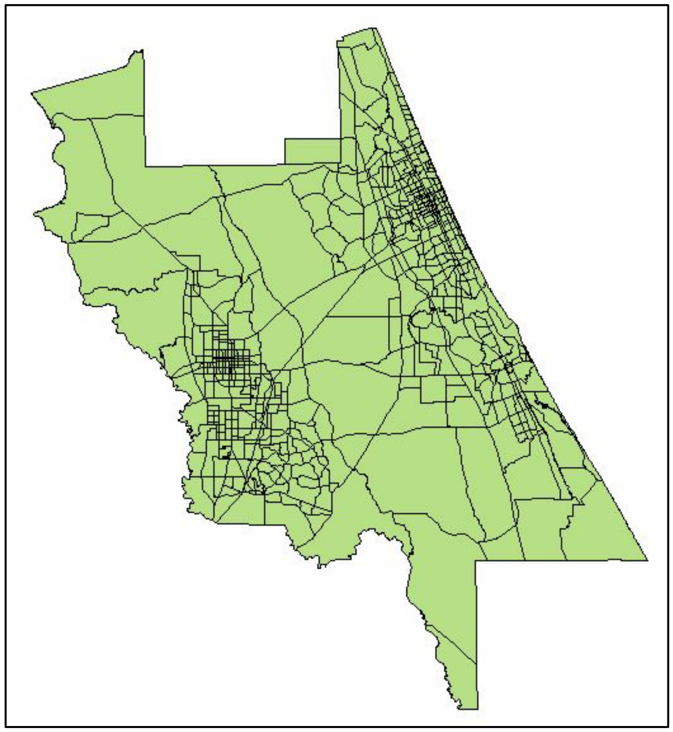


Figure 3.8 1997 Traffic Analysis Zone Boundaries

Cities and Urban Growth Boundaries

The municipal boundary layer downloaded from the Volusia County GIS website is a polygon layer depicting municipalities in Volusia County. There are seventeen municipalities in the county: Pierson, Deland, Lake Helen, Orange City, Ormond Beach, Daytona Beach, Holly Hill, South Daytona, Port Orange, Daytona Beach Shores, Ponce Inlet, New Smyrna Beach, Oak Hill, Edgewater, Debary, Deltona, and Flagler Beach. The total area of incorporated areas in Volusia County is 296 square-miles, which is approximately one quarter of the entire county area. The City of Daytona Beach has the largest area, which is 67 square-miles, while Daytona Beach Shores has the smallest area of 0.81 square-mile.

There are two urban growth boundaries in Volusia County: East Volusia and West Volusia. The West Volusia Urban Growth Boundary encompasses an area of 143 square-miles and includes five cities: Deland, Lake Helen, Orange City, Debary, and Deltona. The East Volusia Urban Growth Boundary covers an area of 247 square-miles and ten cities: Ormond Beach, Daytona Beach, Holly Hill, South Daytona, Port Orange, Daytona Beach Shores, Ponce Inlet, New Smyrna Beach, Oak Hill, and Edgewater. Figure 3.9 shows the seventeen cities within the urban growth boundaries.

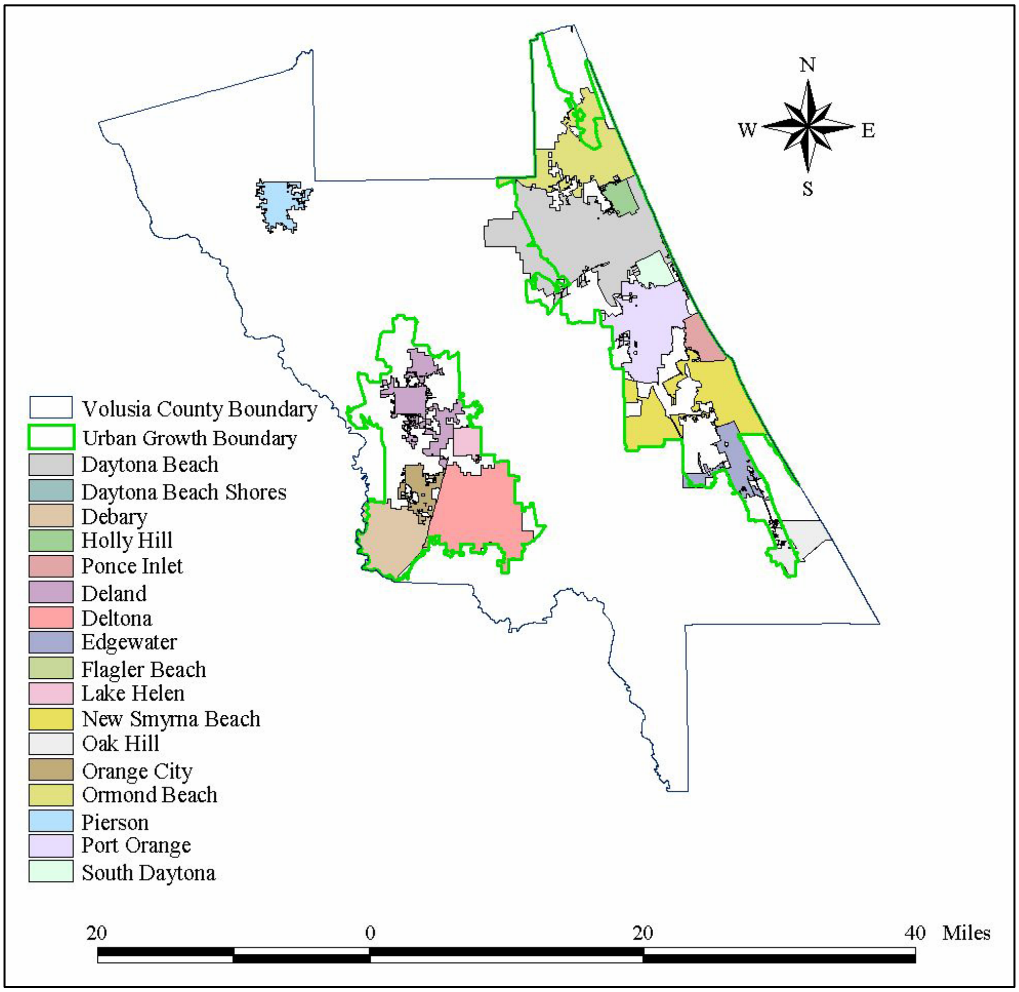


Figure 3.9 Cities and Urban Growth Boundaries

Major Public Lands

The major public land layer was derived from the 2000 property parcel layer and includes parcels identified as one of five types of public lands, including federal land, state land, county land, district land, and municipal land. Table 3.6 shows the total area of each type of public land. Figure 3.10 shows the distribution of the public lands in Volusia County.

Table 3.6 Area of Major Public Land in Volusia County

Public Land	Area (square-miles)
Federal Land	55.7
State Land	174.3
District Land	6.6
County Land	29.5
Municipal Land	25.5

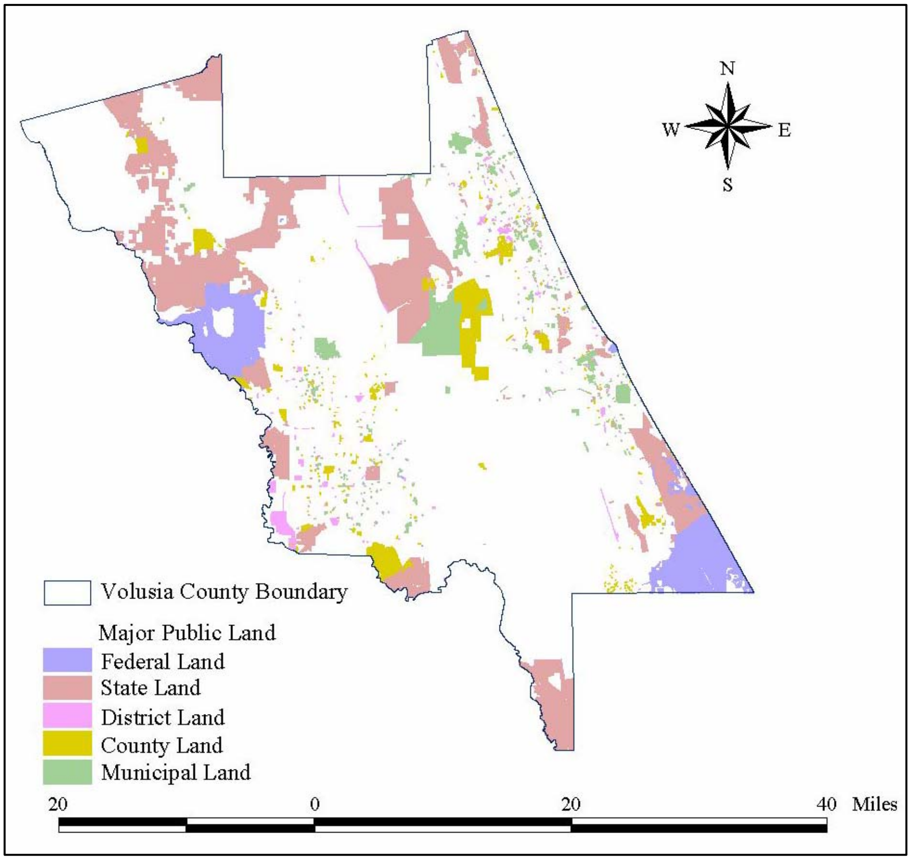


Figure 3.10 Major Public Lands in Volusia County

3.1.3 Assemble and Standardize Property Parcel Data

Property parcel data, downloaded from the Volusia County GIS website, were joined with the property tax database through a unique ID. There are two identifiers available from both datasets: parcel ID[4] (PID) and alternate key. The alternate key was used to match the property tax records with the parcel data because using the alternate key resulted in better matches.

The following information from the parcel data and the property tax database is required to prepare the input files for UrbanSim:

- Land use
- Lot size
- Housing units
- Square footage of building space
- Year built
- Assessed land value
- Assessed improvement value

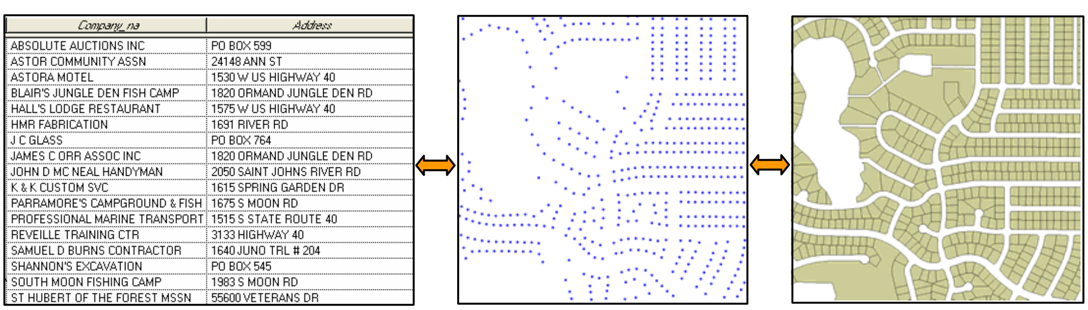
[4] Parcel ID: A complete identification number used by the Assessor of Property to locate and identify property.

A significant number of records from the parcel data and property tax data had missing values for the above attributes. By searching the Volusia County Property Appraiser's website (http://webserver.vcgov.org/vc_search.html), some of the missing values were imputed. However, since both datasets are massive in their size, it is impossible to manually impute all the missing values. An ArcView program was developed to impute the missing values of land use and year built. It examines the surrounding parcels of the same type and draws from the distribution of observed values. UrbanSim provides a script running in MySQL[5] to impute the housing unit data, which matches imputed housing units with the census block housing counts.

After the data imputation, it is necessary to standardize the names and coding of the above attributes in order to facilitate further data preparation. A look-up table was created to categorize the land use codes into standardized and more general land use categories. The look-up table contains two levels of land use category aggregation: *GENERIC_LAND_USE_1* and *GENERIC_LAND_USE_2*. *GENERIC_LAND_USE_1* aggregates land use codes into 26 categories, and *GENERIC_LAND_USE_2* into six broader categories (commercial, government, industrial, residential, non-residential, and group quarters). The look-up table is provided in Appendix II.

3.1.4 Assemble Employment Data

The employment data were geocoded using the address matching process based on street networks that contain the attribute information on house numbers, street directions, street names, and street types. Zone information such as zip code may be used as an optional input. Since the UrbanSim input data are grid cell-based, the employment data need to be geocoded to parcels to avoid introducing spatial errors. However, there are no common fields in the datasets to allow the linkage of the employment data to parcel data other than addresses. The address data downloaded from the Volusia County website (http://www.volusia.org/gis) provide both address information and parcel IDs and were used to establish the connection between employment data and parcel data, as shown in Figure 3.11. The matching process consists of two steps: matching the address data with the parcel data and matching the address data with the employment data. These two steps are described in the next two subsections.



Company_na	Address
ABSOLUTE AUCTIONS INC	PO BOX 599
ASTOR COMMUNITY ASSN	24148 ANN ST
ASTORA MOTEL	1530 W US HIGHWAY 40
BLAIR'S JUNGLE DEN FISH CAMP	1820 ORMAND JUNGLE DEN RD
HALL'S LODGE RESTAURANT	1575 W US HIGHWAY 40
HMR FABRICATION	1691 RIVER RD
J C GLASS	PO BOX 764
JAMES C ORR ASSOC INC	1820 ORMAND JUNGLE DEN RD
JOHN D MC NEAL HANDYMAN	2050 SAINT JOHNS RIVER RD
K & K CUSTOM SVC	1615 SPRING GARDEN DR
PARRAMORE'S CAMPGROUND & FISH	1675 S MOON RD
PROFESSIONAL MARINE TRANSPORT	1515 S STATE ROUTE 40
REVEILLE TRAINING CTR	3133 HIGHWAY 40
SAMUEL D BURNS CONTRACTOR	1640 JUNO TRL # 204
SHANNON'S EXCAVATION	PO BOX 545
SOUTH MOON FISHING CAMP	1983 S MOON RD
ST HUBERT OF THE FOREST MSSN	55600 VETERANS DR

Employment Data | Address Data | Parcel Data

Figure 3.11 Employment Data, Address Data, and Parcel Data

[5] MySQL, an open source software, is a multithreaded, multi-user, SQL (Structured Query Language) Database Management System (DBMS).

Matching Address Data with Parcel Data

The address data come as a GIS point layer with 213,264 points and are associated with an attribute table containing a *PID* (property ID) and address fields. The parcel data are in the form of a GIS polygon layer consisting of 241,900 parcels with *PID*s. A polygon in the parcel data corresponds to a point in address data.

Matching Address Data with Employment Data

The original employment data were matched to the address data by using the business addresses, which made it possible to assign an employment size to the parcels. The address fields from both the employment data and address data were used to join the two data sets. Table 3.7 shows different ways in which the employment data were joined with the address data. Without any changes to the original addresses from the employment data, 10,501 records out of 17,196 records in the employment dataset were matched to the address data. After the addresses in the remaining records from the employment dataset were modified by removing suite numbers, another 1,892 records were matched, still leaving 4,803 business establishments unmatched.

The third method was a manual matching process. To speed up the manual matching process, an ArcView program was developed. Figure 3.12 shows a graphical user interface of the program. The program searches for possible candidates from the address database, which are located within 1,000 feet from an unmatched record. If a match is found from the candidate list, the parcel ID is assigned to the record. If a match is not found, no changes are made and a note is made. The manual matching program records the status (verified match, incorrect but matching parcel found, and incorrect and no matching parcel found), the actions taken, reasons behind the actions, and the new *PID*. After a manual matching process, there were still some businesses that had not matched up with parcels. For these businesses, their original locations were used.

Table 3.7 Results of Address Matching between Employment Data and Address Data

Attempts	Join	Number of Records
1	Match with original address	10,501
2	Match without suite number	1,892
3	Manual match	3,004
4	PO Box	153
5	Not joined	1,646
TOTAL		17,196

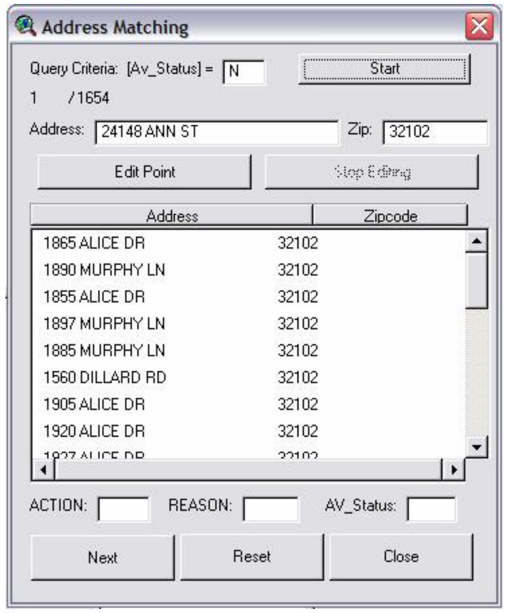


Figure 3.12 Graphical User Interface of Address Matching Program

Tables 3.8 and 3.9 summarize the coding systems used for 'ACTION' and 'REASON', respectively, when the manual matching process was conducted.

Table 3.8 Coding System for ACTION Field

Value	Type of Matching Action
1	Verified that parcel address and employment address match
2	Parcel address did NOT match employment address – a match was found and a new PIN was entered
3	Parcel address did NOT match employment address – a match was not found and no changes were made

Table 3.9 Coding System for REASON Field

Value	Reason for Address Matched or Unmatched
1	Neither employer nor parcel address was available
2	No employer address was available
3	Employer address was incomplete or insufficient to make a match
4	Parcel address was unavailable
5	One component of an address (usually street number) did not match but other components of the address matched
6	Multiple components of an address (street number, name, prefix, etc.) did not match
7	Multiple parcel address matches were found – the correct parcel to assign cannot be determined
8	Employer and parcel addresses matched perfectly
9	Employer and parcel addresses matched – only a typo or abbreviation difference existed

3.1.5 Conversion of Data to Grid

Since UrbanSim simulates urban land use based on grid cells, all data must be converted to a grid. An ArcView conversion program was developed and its graphical user interface is shown in Figure 3.13. When converting a vector layer to a grid, the grid cell values are determined according to the following rules based on the type of attribute values (such as floating, Boolean, or nominal) in the original data layer:

- Boolean or nominal values: the polygon coverage is overlaid on a point coverage that represents the grid cell centroids to assign the political, environmental, or planning geography to each cell in the grid.
- Floating/Percentage values: the polygon coverage is intersected with the grid to determine the areas of environmental or planning geography in each cell in the grid. These areas are then summed up by the grid cell and used to calculate the percentage of grid cells covered by a particular geography.

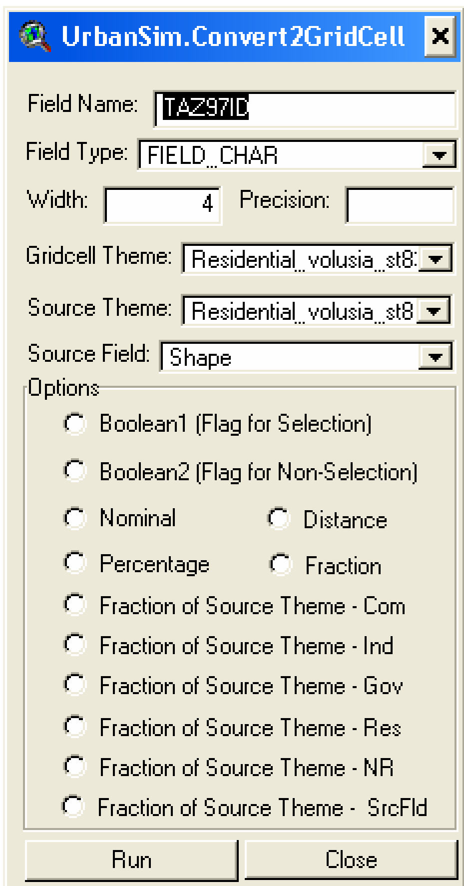


Figure 3.13 Graphical User Interface of the Conversion Program

The conversion program created the following grids:

- PERCENT_WATER – percentage of water area in a grid cell,
- PERCENT_STREAM_BUFFER – percentage of stream buffer area in a grid cell,
- PERCENT_FLOODPLAIN – percentage of flood plain area in a grid cell,

- PERCENT_WETLAND – percentage of wetland area in a grid cell,
- PERCENT_SLOPE – slope in percentage of a grid cell,
- PERCENT_OPEN_SPACE – percentage of open space in a grid cell,
- IS_OUTSIDE_URBAN_GROWTH_BOUNDARY – binary value indicating whether a grid cell is located outside of urban growth boundary,
- IS_INSIDE_NATIONAL_FOREST – binary value indicating whether a grid cell is located inside of national forest,
- IS_INSIDE_TRIBAL_LAND – binary value indicating whether a grid cell is located inside of tribal land,
- IS_INSIDE_MILITARY_BASE – binary value indicating whether a grid cell is located inside of a military base,
- ZONE_ID – TAZ ID,
- CITY_ID – City ID,
- COUNTY_ID – County ID,
- DISTANCE_TO_ARTERIAL – distance to the nearest arterial, and
- DISTANCE_TO_HIGHWAY – distance to the nearest freeway or expressway.

The slope of a grid cell identifies the maximum rate of change in the elevation value from itself to its neighbors. The slope in percentage is calculated as

$$Percent\ Slope = \frac{\Delta elevation}{dist} \times 100$$

where $\Delta elevation$ is the difference between the elevations of two adjacent grid cells and *dist* is the distance between the centroids of the two grid cells.

3.1.6 Assignment of Development Types

The parcel data were converted to grid cells to construct a composite representation of the real estate development within each cell. A development type was assigned to a cell based on its real estate composition of the number of housing units and the square footage of commercial, industrial, and government use in a particular grid cell, as shown in Table 3.10. UrbanSim defines twenty-five development types for residential, mixed use, commercial, industrial, governmental, and vacant developable and undevelopable uses. The distribution of these general uses is shown in Figure 3.14. Further classification of each use is based on density or land use intensity.

Development types are assigned using a SQL script from the UrbanSim website. The SQL script performs queries to select grid cells based on a combination of characteristics and update grid cells with appropriate development type. For instance, the percentage of a cell that is water, wetland, floodplain, steep slope, or public land is used as a criterion to determine whether or not a grid cell is considered undevelopable by UrbanSim.

Table 3.10 Description of Development Type

DEVELOPMENT_TYPE_ID	NAME	MIN_UNITS	MAX_UNITS	MIN_SQFT	MAX_SQFT	PRIMARY_USE
1	R1	1	1	0	499	Residential
2	R2	2	4	0	499	Residential
3	R3	5	9	0	499	Residential
4	R4	10	15	0	499	Residential
5	R5	16	21	0	499	Residential
6	R6	22	30	0	499	Residential
7	R7	31	75	0	4,999	Residential
8	R8	76	28,000	0	4,999	Residential
9	M1	1	9	500	4,999	Mixed Use
10	M2	1	9	5,000	24,999	Mixed Use
11	M3	1	9	25,000	4,000,000	Mixed Use
12	M4	10	30	500	4,999	Mixed Use
13	M5	10	30	5,000	24,999	Mixed Use
14	M6	10	30	25,000	4,000,000	Mixed Use
15	M7	31	2,800	5,000	99,999	Mixed Use
16	M8	31	2,800	100,000	4,000,000	Mixed Use
17	C1	0	0	1	24,999	Commercial
18	C2	0	0	25,000	99,999	Commercial
19	C3	0	0	100,000	4,000,000	Commercial
20	I1	0	0	1	24,999	Industrial
21	I2	0	0	25,000	99,999	Industrial
22	I3	0	0	100,000	4,000,000	Industrial
23	GV	0	0	1	4,000,000	Government
24	Vacant Developable	0	0	0	0	Vacant Developable
25	Undevelopable	0	0	0	0	Vacant Undevelopable

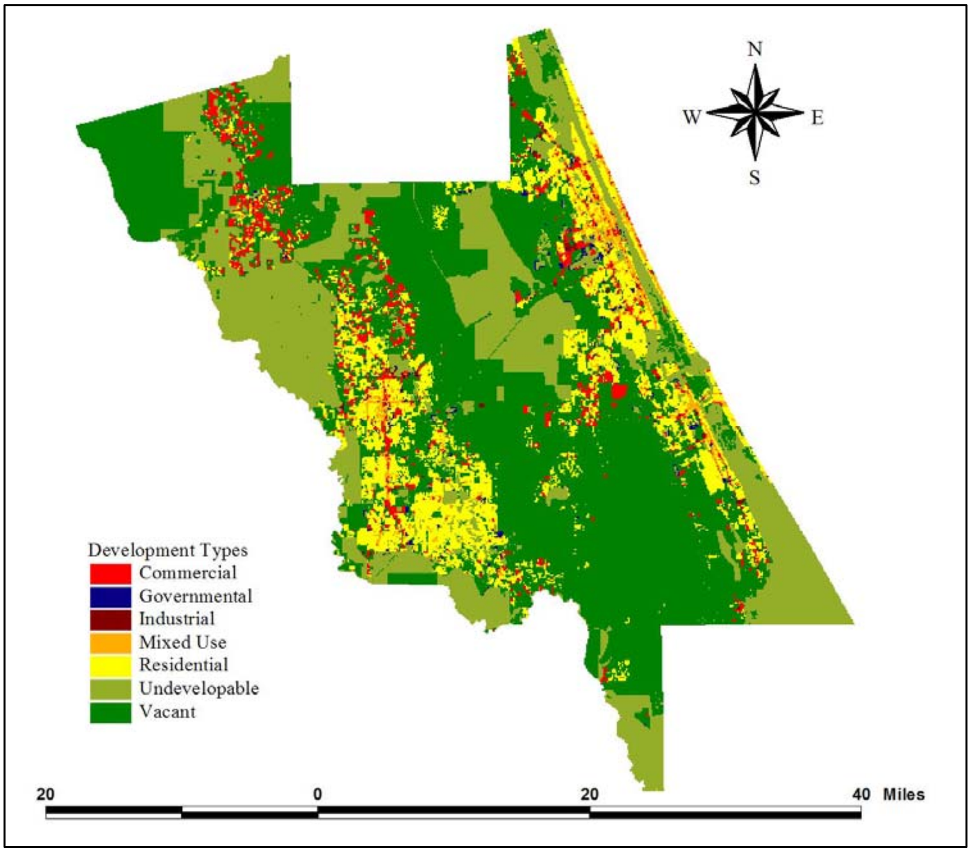


Figure 3.14 Distribution of Development Types

3.1.7 Household Synthesis

Household level data are required to run UrbanSim. Since this type of data is unavailable, it is necessary to synthesize a database of households. UrbanSim provides a Household Synthesis Utility for this purpose by utilizing the census data sources.

The U.S. Census Bureau conducts a census survey decennially to provide demographic, social, economic, and housing statistics at various geographic scales such as nation, state, county, and minor civil division. These statistics are provided in 100-percent and sample data. The 2000 census statistics are available for downloading at the Census Bureau website (http://factfinder.census.gov/servlet/DatasetMainPageServlet?_ds_name=DEC_2000_SF1_U&_program=DEC&_lang=en). The most commonly used census data are the Census 2000 Summary File 1 (SF 1) 100-Percent Data, Census 2000 Summary File 2 (SF 2) 100-Percent Data, Census 2000 Summary File 3 (SF 3) Sample Data, and Census 2000 Summary File 4 (SF 4) Sample Data.

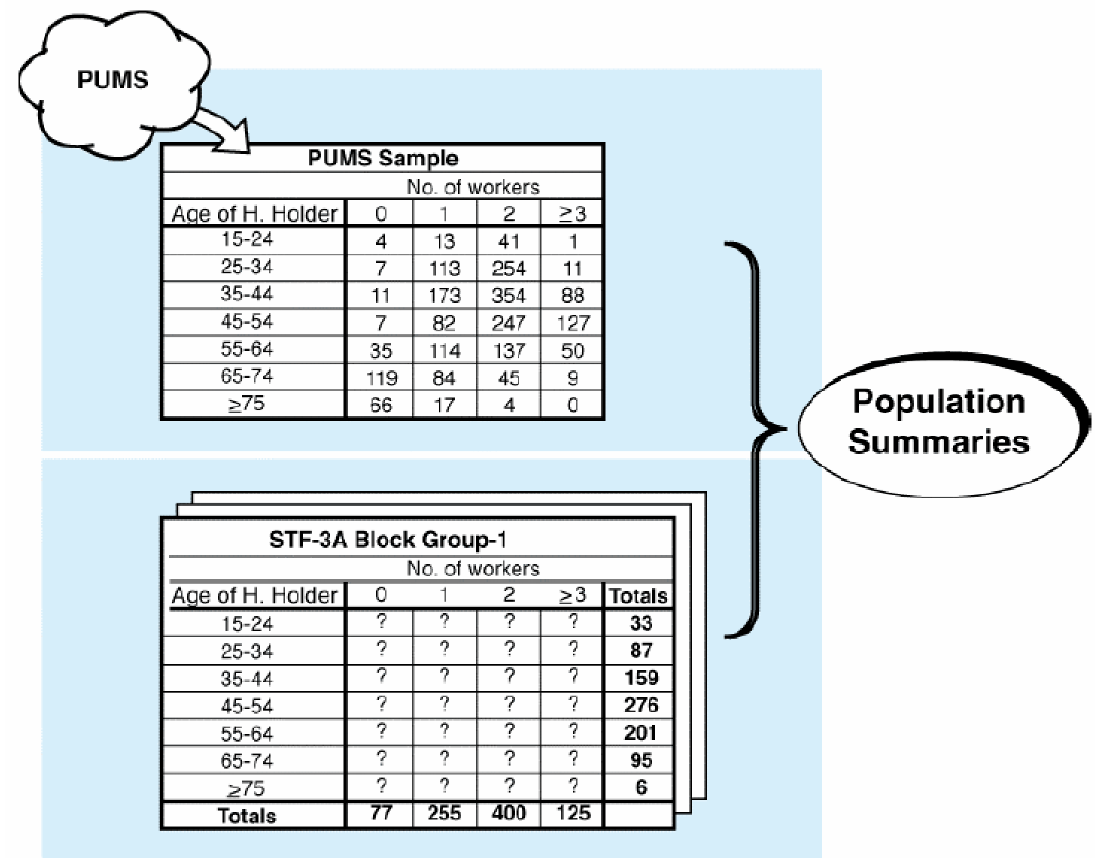
- Census 2000 Summary File 1 (SF 1) 100-Percent Data – counts and information (age, sex, race, Hispanic/Latino origin, household relationship, whether residence is owned or rented) collected from all people and housing units.
- Census 2000 Summary File 2 (SF 2) 100-Percent Data – population and housing characteristics iterated for many detailed race and Hispanic or Latino categories, and American Indian and Alaska Native tribes.

- Census 2000 Summary File 3 (SF 3) Sample Data – detailed population and housing data (such as place of birth, education, employment status, income, value of housing unit, year structure was built) were collected from a 1-in-6 sample and weighted to represent the total population.
- Census 2000 Summary File 4 (SF 4) Sample Data − tabulations by race of population and housing data collected from a sample of the population. The data are given down to the census tract level for 336 races, Hispanic or Latino, American Indian and Alaska Native, and ancestry categories.

The Public Use Microdata Sample (PUMS) files contain a one-percent or five-percent sample of individual records from the census long form and from the American Community Survey on the population and housing characteristics of the people included on those forms. Some of the housing characteristics included in the housing record are house structural characteristics, family characteristics, household characteristics, real estate characteristics, etc. The person record provides characteristics of individuals without revealing personal identity. The PUMS data are available by state at the Census Bureau website at (http://www.census.gov/Press-Release/www/2003/PUMS5.html).

The Census Transportation Planning Package provides tabulations of households, persons, and workers. The CTPP is designed for transportation planning applications and is tabulated from answers to the Census 2000 long form questionnaire, which is mailed to one in six U.S. households. The information is summarized by place of residence, place of work, and for worker-flows between home and work. The CTPP provides comprehensive and cost-effective data in a standard format, and the data are available for downloading at the CTPP website at (http://transtats.bts.gov/Tables.asp?DB_ID=630&DB_Name=Census%20Transportation%20Planning%20Package%20%28CTPP%29%202000&DB_Short_Name=CTPP%202000).

The UrbanSim Household Synthesis Utility utilizes the Census 2000 Summary File 1 (SF 1) 100-Percent Data, Census 2000 Summary File 3 (SF 3) Sample Data, the Public Use Microdata Sample (PUMS), and the Census Transportation Planning Package (CTPP) 2000. It estimates households at the block group level by iterative proportional fitting (IPF). It generates a household distribution that matches the block group marginal distributions of households while protecting the correlation structure of the joint distribution of households in the Public Use Micro Area (PUMA). The joint distribution of households is from the PUMS and the marginal distribution of households is from the Census 2000 Summary File 1 (SF 1) 100-Percent Data or the Census 2000 Summary File 3 (SF 3) Sample Data, as shown in Figure 3.15.



PUMS Sample				
	No. of workers			
Age of H. Holder	0	1	2	≥3
15-24	4	13	41	1
25-34	7	113	254	11
35-44	11	173	354	88
45-54	7	82	247	127
55-64	35	114	137	50
65-74	119	84	45	9
≥75	66	17	4	0

STF-3A Block Group-1					
	No. of workers				
Age of H. Holder	0	1	2	≥3	Totals
15-24	?	?	?	?	33
25-34	?	?	?	?	87
35-44	?	?	?	?	159
45-54	?	?	?	?	276
55-64	?	?	?	?	201
65-74	?	?	?	?	95
≥75	?	?	?	?	6
Totals	77	255	400	125	

Figure 3.15 Household Synthesis using PUMS and Census Data at Block Group Level

The household synthesis utility requires the following input tables:

- SELECTED_PUMAS: This table contains a list of PUMAs to be processed. In this study, there are three PUMAs in Volusia County, which are 01901, 01902, and 01903, as shown in Figure 3.16.
- JOINT_DISTRIBUTION_TABLES: This table provides the database table names for the joint distribution tables from the PUMS. The database table is named PUMS005H in this study.
- CLASSIFICATION_VARIABLES: This table provides the definitions of classification variables for the joint frequency distribution of households. Six variables including *age*, *workers*, *income*, *persons*, *children*, and *race* are used to describe the type of family, and four variables including *age*, *income*, *persons*, and *race* are used to describe the type of non-family household, such as dormitory.
- CLASSIFICATION_CATEGORIES: This table contains the category details for the classification variables. For example, the variable *workers* has four categories and the variable *age* has seven categories.
- JOINT_DISTRIBUTION_DETAILS: This table gives the column names in the joint distribution table for the classification variables. For example, the column name for the variable *workers* is *NWRK,* and *HDHHAGE* is used for the variable *age*.

- BLOCK_GROUP_SUMMARY: This table summarizes, by census block group, the number of households for each family type (Family and Non-Family), total population, and residential vacancy rate. The residential vacancy rate is estimated as vacant housing units divided by total housing units based on one of the Census SF 1 tables (H3. Occupancy Status).
- SELECTED_MARGINAL_DISTRIBUTION_TABLES: This table lists marginal distribution tables generated from the Census SF 1 or SF 3 data. For this study, seven tables are created: P21HHAGE (householder's age), P26HHSIZE (household size), P48WORKERS (number of workers), P76INCOME (income), P79INCOME (income), and P146RACE (race).
- MARGINAL_DISTRIBUTION_DETAILS: This table summarizes the variables, corresponding column names, categories of variables, and marginal distribution table names.
- GRIDCELLS: This table contains the TAZ ID and the number of residential units for each grid cell. This table is a copy of the gridcells from the base year database.
- GRIDCELL_BLOCK_GROUP_MAPPING: This table links grid ID with state ID, county ID, census tract ID, and census block group ID.
- HOUSEHOLD_VARIABLE_MAPPING_DIRECT: This table provides UrbanSim variable names (*persons*, *workers*, *age_of_head*, *income*, *children*, and *race_ID*) and the corresponding column names (*NPERSON*, *NWRK*, *HDHHAGE*, *HINC*, *NOC*, and *RACE*) in the joint distribution table (PUMS005F and PUMS005NF). It is used to map the data in the joint distribution tables to UrbanSim's household data.
- HOUSEHOLD_VARIABLE_MAPPING_INDIRECT: This table contains a list of UrbanSim variables that have an indirect mapping to a particular joint distribution variable in a given column. For this study, no UrbanSim variables have an indirect mapping to a joint distribution variable.
- VEHICLE_OWNERSHIP_DISTRIBUTION: This input table contains the probabilities of each household owning 0, 1, 2, or 3 and more cars based on household size (1, 2, 3, and 4+). Table 1-076 from the 2000 Census Transportation Planning Package Part 1 was used to estimate these probabilities. Table 1-076 provides tabular information based on household size, vehicle availability, and household income.
- PUMS005F and PUMS005NF: These two joint distribution tables provide household information on variables listed in the CLASSIFICATION_VARIABLES table for family and non-family, respectively. Column names used for variables are defined in the JOINT_DISTRIBUTION_DETAILS table.
- P21HHAGE, P26HHSIZE, P48WORKERS, P76INCOME, P79INCOME, P146RACE: These marginal distribution tables have one row per block group. These tables have columns for IDs of block groups, classification variables, and categories. These tables are generated from the Census SF 1 or SF 3 tables.

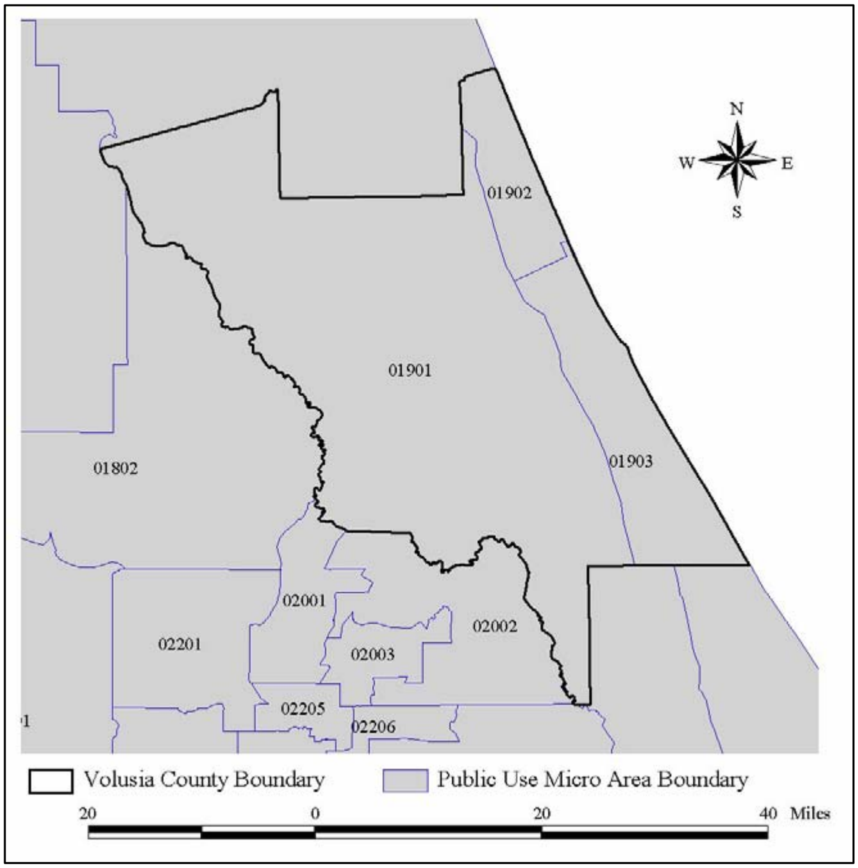


Figure 3.16 Public Use Micro Area Boundary

Table 3.11 summarizes the UrbanSim variables by family type along with the corresponding columns in PUMS and census tables from Census SF1 or SF3. For example, the variable called *persons* is used for both family and non-family family types and is summarized in the *PERSON* column in PUMS for a household. The Census tables P26 from Census SF1 and P14 from Census SF3 provide the number of persons in a household at the block group level.

Table 3.11 UrbanSim Variables with Corresponding PUMS Columns and Census Tables

UrbanSim Variables	Family Type	Column in PUMS	CENSUS Tables
PERSONS: number of persons in the household	Family, Non-Family	H.PERSONS	SF1.P26 / SF3.P14
WORKERS: number of workers in family	Family	H.WIF	SF3.P48
AGE_OF_HEAD: age of household head	Family, Non-Family	P.AGE	SF1.P21
INCOME: household income	Family, Non-Family	H.HINC	SF3.P52
CHILDREN: number of children in the household	Family	H.NOC H.NRC	SF1.P18
RACE_ID: race of householder	Family, Non-Family	P.WHITE: 1 P.BLACK: 2 P.AIAN: 3 P.ASIAN: 4 P.NHPI: 4 P.OTHER: 5	SF1.P15A SF1.P15B SF1.P15C SF1.P15D SF1.P15E SF1.P15F + SF1.P15G

3.1.8 Accessibilities

The accessibility is defined as the sum of opportunities weighted by the composite utility across all modes of travel for each zone pair. The composite utility is obtained as the logsum from the mode choice model for each origin-destination pair. Zonal opportunities may be measured by total employment or total households.

Accessibility is obtained as:

$$Accessibility_i = \sum_{j=1}^{J} (D_j \times \exp(LogSum_{aij})),$$

where D_j = quantity of activity in location j,

$LogSum_{aij}$ = logsum for households with vehicle ownership level a from TAZ i to TAZ j; and

J = number of TAZs.

UrbanSim requires logsum values from the travel model. To compute the logsum matrix, it is necessary to understand the mode choice model. The mode choice model in the Volusia Transportation Planning Model consists of a three-level nested structure, as shown in Figure 3.17. In the primary nest, the total person trips are divided into "Auto" trips and "Transit" trips. In the secondary nest, the auto trips are split into "Drive Alone" trips and "Shared Ride" trips, and the transit trips into "Walk Access" trips and "Auto Access" trips. In the third nest, shared ride trips are split into "One Passenger" and "2+ Passengers". On the transit side, the walk access trips are split into "Local Bus" trips and "Premium Modes" (light rail, fixed guideway, express bus, etc.) trips, while the auto access trips are split into "Park-and-Ride" trips and "Kiss-and-Ride" trips.

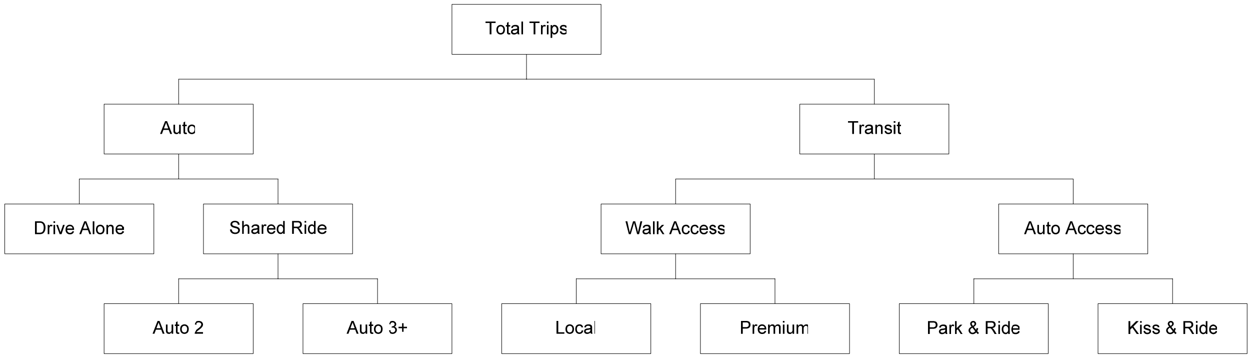


Figure 3.17 Nested Mode Choice Model

The equation of logsum is:

$$LogSum_{\text{mod}\,e} = \ln\left[\exp\left(0.8 \times LogSum_{Auto}\right) + \exp\left(0.3 \times LogSum_{Transit}\right)\right]$$

where

$$LogSum_{Auto} = \ln\left[\exp(0.2 \times LogSum_{ShareRide}) + \exp(U_{DriveAlone})\right],$$

$$LogSum_{Transit} = \ln\left[\exp(LogSum_{WalkAccess}) + \exp(0.5 \times LogSum_{AutoAccess})\right],$$

$$LogSum_{ShareRide} = \ln\left[\exp(U_{Auto2}) + \exp(U_{Auto3+})\right],$$

$$LogSum_{WalkAccess} = \ln\left[\exp(0.5 \times U_{Local}) + \exp(0.5 \times U_{\text{Pr}emium})\right], \text{ and}$$

$$LogSum_{AutoAceess} = \ln\left[\exp(0.5 \times U_{Park\&Ride}) + \exp(0.5 \times U_{Kiss\&Ride})\right].$$

Utility functions vary by the numbers of cars available in a household. The Volusia Transportation Planning Model defines four types of households: 0-car household, 1-car household, and 2+-car household. The utility functions for these household types are given below.

0 – Car Households

$U_{\text{Drive Alone}}$ = – 0.045* **Highway Terminal Time** – 0.02* **Highway Run Time** – 0.0025* **Auto Operating Costs**

$U_{\text{Auto 2}}$ = -0.4520 – 0.045* **Highway Terminal Time** – 0.02* **Highway Run Time** – 0.0025* **Auto Operating Costs**

$U_{\text{Auto 3+}}$ = -0.6280 – 0.045* **Highway Terminal Time** – 0.02* **Highway Run Time** – 0.0025* **Auto Operating Costs**

$U_{\text{Local Bus}}$ = 1.8600 – 0.045* **Transit Walk Time** – 0.02* **Transit Run Time** – 0.045* **Transit First Wait** (<7 min) – 0.023* **Transit First Wait** (>7 min) – 0.045* **Transit Second Wait** – 0.045* **Transit Number of Transfers** –0.0032* **Transit Fare**

$U_{Premium}$ = 1.8600 – 0.045* **Transit Walk Time** – 0.02* **Transit Run Time** – 0.045* **Transit First Wait** (<7 min) – 0.023* **Transit First Wait** (>7 min) – 0.045* **Transit Second Wait** – 0.045* **Transit Number of Transfers** –0.0032* **Transit Fare**

$U_{Park\ \&\ Ride}$ = 0.3600 – 0.045* **Transit Walk Time** – 0.02 * **Transit Auto Access Time** – 0.02* **Transit Run Time** – 0.045* **Transit First Wait** (<7 min) – 0.023 **Transit First Wait** (>7 min) – 0.045* **Transit Second Wait** – 0.045* **Transit Number of Transfers** –0.0032* **Transit Fare** – 0.0025* **Auto Operating Costs**

$U_{Kiss\ \&\ Ride}$ = 0.1900 – 0.045* **Transit Walk Time** – 0.02 * **Transit Auto Access Time** – 0.02* **Transit Run Time** – 0.045* **Transit First Wait** (<7 min) – 0.023* **Transit First Wait** (>7 min) – 0.045* **Transit Second Wait** – 0.045* **Transit Number of Transfers** –0.0032* **Transit Fare** – 0.0025* **Auto Operating Costs**

1 – Car Households

$U_{Drive\ Alone}$ = – 0.045* **Highway Terminal Time** – 0.02* **Highway Run Time** – 0.0025* **Auto Operating Costs**

$U_{Auto\ 2}$ = -1.1120 – 0.045* **Highway Terminal Time** – 0.02* **Highway Run Time** – 0.0025* **Auto Operating Costs**

$U_{Auto\ 3+}$ = -1.3380 – 0.045* **Highway Terminal Time** – 0.02* **Highway Run Time** – 0.0025* **Auto Operating Costs**

$U_{Local\ Bus}$ = -1.5900 – 0.045* **Transit Walk Time** – 0.02* **Transit Run Time** – 0.045* **Transit First Wait** (<7 min) – 0.023* **Transit First Wait** (>7 min) – 0.045* **Transit Second Wait** – 0.045* **Transit Number of Transfers** –0.0032* **Transit Fare**

$U_{Premium}$ = -1.5900 – 0.045* **Transit Walk Time** – 0.02* **Transit Run Time** – 0.045* **Transit First Wait** (<7 min) – 0.023* **Transit First Wait** (>7 min) – 0.045* **Transit Second Wait** – 0.045* **Transit Number of Transfers** –0.0032* **Transit Fare**

$U_{Park\ \&\ Ride}$ = -1.1090 – 0.045* **Transit Walk Time** – 0.02 * **Transit Auto Access Time** – 0.02* **Transit Run Time** – 0.045* **Transit First Wait** (<7 min) – 0.023* **Transit First Wait** (>7 min) – 0.045* **Transit Second Wait** – 0.045* **Transit Number of Transfers** –0.0032* **Transit Fare** – 0.0025* **Auto Operating Costs**

$U_{Kiss\ \&\ Ride}$ = -1.2790 – 0.045* **Transit Walk Time** – 0.02 * **Transit Auto Access Time** – 0.02* **Transit Run Time** – 0.045* **Transit First Wait** (<7 min) – 0.023* **Transit First Wait** (>7 min) – 0.045* **Transit Second Wait** – 0.045* **Transit Number of Transfers** –0.0032* **Transit Fare** – 0.0025* **Auto Operating Costs**

2+ Car Households

$U_{Drive\ Alone}$ = – 0.045* **Highway Terminal Time** – 0.02* **Highway Run Time** – 0.0025* **Auto Operating Costs**

$U_{Auto\ 2}$ = -1.8720 – 0.045* **Highway Terminal Time** – 0.02* **Highway Run Time** – 0.0025* **Auto Operating Costs**

$U_{Auto\ 3+}$ = -2.0980 – 0.045* **Highway Terminal Time** – 0.02* **Highway Run Time** – 0.0025* **Auto Operating Costs**

$U_{Local\ Bus}$ = -2.5080 – 0.045* **Transit Walk Time** – 0.02* **Transit Run Time** – 0.045* **Transit First Wait** (<7 min) – 0.023* **Transit First Wait** (>7 min) – 0.045* **Transit Second Wait** – 0.045* **Transit Number of Transfers** – 0.0032* **Transit Fare**

$U_{Premium}$ = -2.5080 – 0.045* **Transit Walk Time** – 0.02* **Transit Run Time** – 0.045* **Transit First Wait** (<7 min) – 0.023* **Transit First Wait** (>7 min) – 0.045* **Transit Second Wait** – 0.045* **Transit Number of Transfers** – 0.0032* **Transit Fare**

$U_{Park\ \&\ Ride}$ = -1.9220 – 0.045* **Transit Walk Time** – 0.02 * **Transit Auto Access Time** – 0.02* **Transit Run Time** – 0.045* **Transit First Wait** (<7 min) – 0.023* **Transit First Wait** (>7 min) – 0.045* **Transit Second Wait** – 0.045* **Transit Number of Transfers** – 0.0032* **Transit Fare** – 0.0025* **Auto Operating Costs**

$U_{Kiss\ \&\ Ride}$ = -2.0920 – 0.045* **Transit Walk Time** – 0.02 * **Transit Auto Access Time** – 0.02* **Transit Run Time** – 0.045* **Transit First Wait** (<7 min) – 0.023* **Transit First Wait** (>7 min) – 0.045* **Transit Second Wait** – 0.045* **Transit Number of Transfers** – 0.0032* **Transit Fare** – 0.0025* **Auto Operating Costs**

3.2 Data Compilation

After the data are prepared, they need to be compiled to generate input tables for UrbanSim. Detailed information on the format of input tables is provided in Appendix III. The input tables are listed below.

- ANNUAL_EMPLOYMENT_CONTROL_TOTALS – employment forecasts, by sector, by location (home-based or non-home based), and by year for each simulated year;
- ANNUAL_HOUSEHOLD_CONTROL_TOTALS – households forecasts, by race, by household size, and by year for each simulated year;
- ANNUAL_RELOCATION_RATES_FOR_HOUSEHOLDS – relocation rates of households, by household head age and by household income;
- ANNUAL_RELOCATION_RATES_FOR_JOBS – relocation rates of jobs, by sector,
- BASE_YEAR – 2000 in this study;
- CITIES – the name of cities included in the study area;
- COUNTIES – the name of counties included in the study area;
- DEVELOPER_MODEL_COEFFICIENTS – estimated coefficients of variables included in the Developer Model;
- DEVELOPER_MODEL_SPECIFICATION – variables included in the Developer Model;
- DEVELOPMENT_CONSTRAINTS – constraints that restrict the development types;
- DEVELOPMENT_CONSTRAINT_EVENTS – constraints that restrict changes of development types in the future;
- DEVELOPMENT_EVENTS – development events scheduled to take place in the future,
- DEVELOPMENT_EVENT_HISTORY – the development events that occurred prior to the base year;
- DEVELOPMENT_TYPES – definitions of all development types;
- DEVELOPMENT_TYPE_GROUPS – definitions of all development type groups;
- DEVELOPMENT_TYPE_GROUP_DEFINITIONS – the development type membership of development type groups;
- EMPLOYMENT_ADHOC_SECTOR_GROUPS – employment sector groups such as basic (resource, construction, manufacturing, transportation, communication, electric, gas, and whole sale trade), retail, and service;
- EMPLOYMENT_ADHOC_SECTOR_GROUP_DEFINITIONS – the employment sector membership of employment ad hoc sector groups;
- EMPLOYMENT_EVENTS – events that creates employment;
- EMPLOYMENT_HOME_BASED_LOCATION_CHOICE_MODEL_COEFFICIENTS – estimated coefficients of variables included in the Employment Home Based Location Choice Model;
- EMPLOYMENT_HOME_BASED_LOCATION_CHOICE_MODEL_SPECIFICATION – variables included in the Employment Home Based Location Choice Model;
- EMPLOYMENT_LOCATION_CHOICE_MODEL_COEFFICIENTS – estimated coefficients of variables included in the Employment Location Choice Model;

- EMPLOYMENT_LOCATION_CHOICE_MODEL_SPECIFICATION – variables included in the Employment Location Choice Model;
- EMPLOYMENT_NON_HOME_BASED_LOCATION_CHOICE_MODEL_COEFFICIENTS – estimated coefficients of variables included in the Employment Non-Home Based Location Choice Model;
- EMPLOYMENT_NON_HOME_BASED_LOCATION_CHOICE_MODEL_SPECIFICATION – variables included in the Employment Non-Home Based Location Choice Model;
- EMPLOYMENT_SECTORS – definitions of 15 employment sectors;
- GEOGRAPHIES – geography types such as region, city, TAZ, and grid in the study area,
- GEOGRAPHY_NAMES – names of cities and county in the study area;
- GRIDCELLS – grid cells with geographical, environmental, and political information;
- GRIDCELLS_IN_GEOGRAPHY – grid cell ID associated with corresponding geography type and ID;
- GRIDCELL_FRACTIONS_IN_ZONES – fractions of grid cells overlaying with TAZ zones;
- HOUSEHOLDS – all households in the study area;
- HOUSEHOLDS_FOR_ESTIMATION – sample set of households for estimation;
- HOUSEHOLD_CHARACTERISTICS_FOR_HLC – household characteristics for the Household Location Choice Model;
- HOUSEHOLD_CHARACTERISTICS_FOR_HT – household characteristics for the Household Transition Model;
- HOUSEHOLD_LOCATION_CHOICE_MODEL_COEFFICIENTS – estimated coefficients of variables included in the Household Location Choice Model;
- HOUSEHOLD_LOCATION_CHOICE_MODEL_SPECIFICATION – variables included in the Household Location Choice Model;
- JOBS – all jobs in the study area with corresponding sector ID, grid cell ID, and indication of home-based;
- JOBS_FOR_ESTIMATION_HOME_BASED – sample of home based jobs in the study area;
- JOBS_FOR_ESTIMATION_NON_HOME_BASED – sample of non-home based jobs in the study area;
- LAND_PRICE_MODEL_COEFFICIENTS – estimated coefficients of variables included in the Land Price Model;
- LAND_PRICE_MODEL_SPECIFICATION – variables included in the Land Price Model;
- LAND_USE_EVENTS – changes of land use scheduled for future;
- MODELS – the models to run for each year of the simulation;
- MODEL_VARIABLES – all variables provided by UrbanSim;
- PLAN_TYPES – planned land use;
- PRIMARY_USES – names for primary land uses such as residential, mixed use, commercial, industrial, governmental, vacant developable, and undevelopable;
- RACE_NAMES – the name of racial groups;
- RESIDENTIAL_LAND_SHARE_MODEL_COEFFICIENTS – estimated coefficients of variables included in the Residential Land Share Model;

- RESIDENTIAL_LAND_SHARE_MODEL_SPECIFICATION – variables included in the Residential Land Share Model;
- RESIDENTIAL_UNITS_FOR_HOME_BASED_JOBS – residential units for jobs located in residential units;
- SCENARIO_INFORMATION – description of scenario;
- SQFT_FOR_NON_HOME_BASED_JOBS – floor space for jobs that are not located in residential units;
- TARGET_VACANCIES – target vacancies used by the Developer Model;
- TRANSITION_TYPES – transitions from one development type to another development type;
- TRAVEL_DATA – outputs from the travel model that provide logsums for the Accessibility Model;
- URBANSIM_CONSTANTS – constants needed for UrbanSim; and
- ZONES – traffic analysis zones with travel time to the CBD and airport.

4. SCENARIO DESIGN

Land use and transportation scenarios are a part of the input to UrbanSim. UrbanSim is designed to test policies that deal with environmental, sociological, and economic concerns. Possible scenarios developed based on such policies may range from urban growth boundaries at the regional or metropolitan scale to street design, mixing of uses, and development at the neighborhood or site-specific scale. Scenarios that may be tested by UrbanSim are:

- Macroeconomic assumptions;
- Household and employment control totals;
- Development constraints determined by any combinations of political and planning overlays, environmental overlays, and land use plan designation;
- Development constraints based on development types that cannot occur;
- Transportation infrastructure; and
- User-specified events.

The scenarios that will be applied to UrbanSim, in conjunction with the FSUTMS, are developed based on the LRTP and socioeconomic projections provided by the BEBR. Section 4.1 describes three alternatives based on the transportation improvements that are defined in the Volusia LRTP. Section 4.2 discusses the socio-demographic forecasts from the BEBR. Section 4.3 summarizes scenarios to be tested in this study.

4.1 Long Range Transportation Plan

The Long Range Transportation Plan (LRTP) is periodically updated by the Volusia County Metropolitan Planning Organization (MPO) to guide the expenditure of federal and state transportation funding based on countywide transportation planning. In the 2020 Long Range Transportation Plan, three different alternatives of transportation improvements and the final plan are documented. In this study, alternative 2, alternative 3, and the final plan are selected for the scenario testing. They are summarized in Tables 4.1, 4.2, and 4.3, respectively.

Table 4.1 Alternative 2 Transportation Improvement Plan

Road Name	Limits (From-To)	Improvement
Florida Intrastate Highway System (FIHS) Roads		
I-95	US 92 to Brevard County	Widen to 6 Lanes
I-4	Seminole County to SR 472	Widen to 8 Lanes (6+2)
I-4	SR 472 to I-95	Widen to 6 Lanes
I-4	Connector Between I-4 & US 92	Remove Connector Ramps
SR 40	Lake County to Cone Rd	Widen to 4 Lanes
Non-FIHS State Roads		
US 17	SR 40 to Ponce DeLeon Blvd	Widen to 4 Lanes
US 17/92	SR 15A to SR 472	Widen to 6 Lanes
SR A1A	Sandra Dr to Neptune Av	Widen to 3 Lanes
SR 40	Tymber Creek Rd to Nova Rd	Leave as 4 Lane Road
SR 44	Blue Lake Av to Summit Av	Widen to 4 Lanes
SR 400 (Beville Rd)	Clyde Morris Blvd to Nova Rd	Widen to 6 Lanes
SR 415	SR 44 to Seminole County	Widen to 4 Lanes

SR 421 (Dunlawton Av)	Nova Rd to Spruce Creek Rd	Widen to 6 Lanes
SR 430 (Mason Av)	Clyde Morris Blvd to Beach St	Widen to 6 Lanes
SR 442 (Indian River Blvd)	SR 415 to I-95	Extend as 2 Lane Road
SR 472	Kentucky Av to I-4	Widen to 6 Lanes
Local Roads		
Airport Rd	Taylor Rd to Pioneer Trail	Widen to 4 Lanes
Airport Rd	Pioneer Trail to SR 44	Extend as 2 Lane Road
Airport Rd	SR 44 to Park Av Extension	Extend as 2 Lane Road
Bellevue Av Extension	US 92 to Williamson Blvd	Widen to 4 Lanes
Beresford Av	SR 15A to US 17/92	Widen to 4 Lanes
Beresford Av	Blue Lake Av to Summit Av	Extend as 2 Lane Road
Blue Lake Av	Orange Camp Rd to SR 472	Extend as 2 Lane Road
Clyde Morris Blvd	Falls Way to LPGA Blvd	Widen to 4 Lanes
Clyde Morris Blvd	LPGA Blvd to Jimmy Ann Rd	Widen to 4 Lanes
Clyde Morris Blvd	Jimmy Ann Dr to US 92	Leave as 4 Lane Road
Clyde Morris Blvd	US 92 to Beville Rd	Widen to 6 Lanes
Clyde Morris Blvd	Beville Rd to Dunlawton Av	Leave as 4 Lane Road
CR 92	SR 15A to US 17/92	Widen to 4 Lanes
Deltona Blvd	Enterprise Rd to DeBary Av	Widen to 4 Lanes
Dirksen/DeBary/Doyle	US 17/92 to I-4	Widen to 4 Lanes
Dirksen/DeBary/Doyle	I-4 to Deltona Blvd	Widen to 6 Lanes
Dirksen/DeBary/Doyle	Providence Blvd to SR 415	Widen to 4 Lanes
Dunn Av	LPGA Blvd to Williamson Blvd	Extend as 2 Lane Road
Dunn Av	Williamson Blvd to Bill France Blvd	Widen to 4 Lanes
Elkcam Blvd Extension	Riverhead Dr to SR 415	Extend as 2 Lane Road
Enterprise Rd	Deltona Blvd to Main St	Widen to 4 Lanes
Enterprise Rd	US 17/92 to Saxon Blvd	Widen to 6 Lanes
Frontage Rd (along I-4)	Summit Av to Orange Camp Rd	Extend as 2 Lane Road
Frontage Rd (along I-4)	Orange Camp Rd to SR 472	Widen to 4 Lanes
Garfield Av	SR 44 to Taylor Rd	Extend as 2 Lane Road
Hand Av	SR 40 to Tymber Creek Rd	Extend as 2 Lane Road
Hand Av	Tymber Creek Rd to Nova Rd	Extend and Widen as 4 Lane Road
Howland Blvd	I-4 to Providence Blvd	Widen to 6 Lanes
Howland Blvd	Providence Blvd to SR 415	Widen to 4 Lanes
Kentucky Av	SR 472 to Graves Av	Widen to 4 Lanes
LPGA Blvd	US 1 to Nova Rd	Widen to 4 Lanes
LPGA Blvd	Nova Rd to Clyde Morris Blvd	Leave as 4 Lane Road
LPGA Blvd	Clyde Morris Blvd to I-95	Widen to 6 Lanes
LPGA Blvd	I-95 to Tymber Creek Rd Extension	Widen to 4 Lanes
LPGA Blvd	US 92 to I-4	Extend as 2 Lane Road + 2 Interchanges
LPGA Blvd	I-4 to Tomoka Farms RD	Extend as 2 Lane Road
Mason Av	Williamson Blvd to Bill France Blvd	Widen to 4 Lanes
Old Mission Rd	Josephine St to Eslinger Rd	Widen to 4 Lanes
Orange Av/Silver Beach Bridge	End of 2 lane segment to Peninsula Dr	Widen to 4 Lanes

Orange Camp Rd	US 17/92 to I-4	Widen to 4 Lanes
Pioneer Trail	Tomoka Farms Rd to I-95 (plus Interchange)	Widen to 4 Lanes
Pioneer Trail	I-95 to Turnbull Bay rd	Widen to 4 Lanes
Pioneer Trail	Turnbull Bay Rd to Sugar Mill Rd	Widen to 4 Lanes
Providence Blvd/ Sixma Rd	Lake Helen Osteen Rd to Howland Blvd	Extend as 2 Lane Road
Providence Blvd	Howland Blvd to Ft Smith Blvd	Widen to 4 Lanes
Providence Blvd	Tivoli Dr to Saxon Blvd	Widen to 4 Lanes
Providence Blvd	Saxon Blvd to Doyle Rd	Widen to 4 Lanes
Rhode Island Av	Westside Connector to US 17/92	Extend as 2 Lane Road
Rhode Island Av	Veterans Mem Pkwy to Normandy Blvd	Extend as 2 Lane Road
Saxon Blvd	Westside Connector to US 17/92	Extend as 2 Lane Road
Saxon Blvd	US 17/92 to Normandy Blvd	Widen to 6 Lanes
Saxon Blvd	Tivoli Dr to Providence Blvd	Widen to 4 Lanes
Sugar Mill Rd	I-4 to Williamson Blvd	Widen to 4 Lanes
Taylor Rd	Tomoka Farms Rd to Williamson Blvd/Airport Rd	Widen to 4 Lanes
Taylor Rd	Williamson Blvd/Airport Rd to I-95 (at Dunlawton Av)	Widen to 6 Lanes
Tomoka Farms Rd Extension	LPGA Blvd (north end) to Dunn Av Extension	Extend as 2 Lane Road
Tomoka Farms Rd Extension	Dunn Av Extension to US92	Extend as 2 Lane Road
Tomoka Farms Rd	US 92 to SR 44	Widen to 4 Lanes
Tymber Creek Rd Extension	Riverbend Rd to LPGA Blvd	Extend as 2 Lane Road
Westside Connector	SR 44 to Highbanks Rd	Connect as 2 Lane Corridor
W. Volusia Bltwy/ Veteran's Mem Pkwy	SR 44 to Saxon Blvd	Widen to 4 Lanes
Williamson Blvd	Hand Av to Indigo/Dunn Av	Widen to 4 Lanes
Williamson Blvd	US 92 to Beville Rd	Widen to 6 Lanes
Williamson Blvd	Beville Rd to Turnbull Bay Rd/Pioneer Trail	Widen to 4 Lanes

Table 4.2 Alternative 3 Transportation Improvement Plan

Road Name	Limits (From-To)	Improvement
Florida Intrastate Highway System (FIHS) Roads		
I-95	US 92 to Brevard County	Widen to 6 Lanes
I-95	between US 1 and SR 40	New Interchange
I-95	At Pioneer Trail	New Interchange
I-4	SR 472 to I-95	Widen to 6 Lanes
I-4	At Shuntz Rd	New Interchange
SR 40	Lake County to Cone Rd	Widen to 4 Lanes
Non-FIHS State Roads		
US 1	SR 40 to Park Av	Intersection Improvements
US 17	SR 40 to Ponce DeLeon Blvd	Widen to 4 Lanes
US 92	Nova Rd to US 1	Widen to 6 Lanes

US 17/92	SR 15A to SR 472	Widen to 6 Lanes
SR 400 (Beville Rd)	I-95 to Clyde Morris Blvd	Widen to 6 Lanes
SR 400 (Beville Rd)	Clyde Morris Blvd to Nova Rd	Widen to 6 Lanes
SR 400 (Beville Rd)	Nova Rd to US 1	Widen to 6 Lanes
SR 415	SR 44 to Seminole County	Widen to 4 Lanes
SR 421 (Taylor Rd)	Williamson Blvd/Airport Rd to I-95 (at Dunlawton Av)	Widen to 6 Lanes
SR 483 (Clyde Morris Blvd)	US 92 to Beville Rd	Widen to 6 Lanes
Local Roads		
Airport Rd	Taylor Rd to Pioneer Trail	Widen to 4 Lanes
Beresford Av	Blue Lake Av to Summit Av	Extend as 2 Lane Road
Blue Lake Av	Orange Camp Rd to SR 472	Extend as 2 Lane Road
Clyde Morris Blvd	Falls Way to LPGA Blvd	Widen to 4 Lanes
CR 92	SR 15A to US 17/92	Widen to 4 Lanes
Deltona Blvd	Enterprise Rd to DeBary Av	Widen to 4 Lanes
Dirksen/DeBary/Doyle	US 17/92 to I-4	Widen to 4 Lanes
Dirksen/DeBary/Doyle	I-4 to Deltona Blvd	Widen to 6 Lanes
Dunn Av	LPGA Blvd to Williamson Blvd	Extend as 2 Lane Road
Dunn Av	Williamson Blvd to Bill France Blvd	Widen to 4 Lanes
Enterprise Rd	Deltona Blvd to Main St	Widen to 4 Lanes
Enterprise Rd	US 17/92 to Saxon Blvd	Widen to 6 Lanes
Frontage Rd (along I-4)	Summit Av to Orange Camp Rd	Extend as 2 Lane Road
Frontage Rd (along I-4)	Orange Camp Rd to SR 472	Widen to 4 Lanes
Hand Av	Tymber Creek Rd to Williamson Blvd	Extend as 2 Lane Road
Howland Blvd	I-4 to Providence Blvd	Widen to 6 Lanes
Howland Blvd	Providence Blvd to Courtland Blvd	Widen to 4 Lanes
LPGA Blvd	Clyde Morris Blvd to I-95	Widen to 6 Lanes
LPGA Blvd	I-95 to Tymber Creek Rd Extension	Widen to 4 Lanes
LPGA Blvd	US 92 to Tomoka Farms Rd	Extend as 2 Lane Road
Main Street Bridge	Beach St to Halifax Av	Widen to 4 Lanes
Mangoe-Matanzas	Cassadaga Rd to Rhode Island Av	Extend as 2 Lane Road
Mason Av	Williamson Blvd to Bill France Blvd	Widen to 4 Lanes
Orange Camp Rd	US 17/92 to I-4	Widen to 4 Lanes
Pioneer Trail	Tomoka Farms Rd to Turnbull Bay Rd	Widen to 4 Lanes
Providence Blvd/ Sixma Rd	Lake Helen Osteen Rd to Howland Blvd	Extend as 2 Lane Road
Providence Blvd	Howland Blvd to Ft Smith Blvd	Widen to 4 Lanes
Providence Blvd	Tivoli Dr to Saxon Blvd	Widen to 4 Lanes
Rhode Island Av	Westside Connector to US 17/92	Extend as 2 Lane Road
Rhode Island Av	Veterans Mem Pkwy to Normandy Blvd	Extend as 2 Lane Road
Saxon Blvd	Westside Connector to US 17/92	Extend as 2 Lane Road
Saxon Blvd	US 17/92 to 4 lane portion west of Enterprise Rd	Widen to 4 Lanes
Saxon Blvd	Enterprise Rd to I-4	Widen to 6 Lanes
Saxon Blvd	Tivoli Dr to Providence Blvd	Widen to 4 Lanes
Shuntz Rd/Madeline Av	I-4 to Williamson Blvd	Extend as 2 Lane Road
Spruce Creek Rd	Herbert St to Central Park Blvd	Extend as 2 Lane Road &

		Widen to 4 Lanes
Taylor Rd	Tomoka Farms Rd to Williamson Blvd/Airport Rd	Widen to 4 Lanes
Tymber Creek Rd Extension	Riverbend Rd to LPGA Blvd	Extend as 2 Lane Road
Westside Connector	SR 44 to Saxon Blvd	Connect as New 2 Lane Corridor
W. Volusia Bltwy/ Veteran's Mem Pkwy	US 92 to Saxon Blvd	Widen to 4 Lanes
Williamson Blvd	Hand Av to Indigo/Dunn Av	Widen to 4 Lanes
Williamson Blvd	Beville Rd to Taylor Rd	Widen to 4 Lanes
Williamson Blvd	Current terminus to Pioneer Trail/Turnbull Bay Rd	Extend as 2 Lane Road
Yorktowne Blvd	Dunlawton Av to Taylor Rd	Extend as 2 Lane Road

Table 4.3 Final Transportation Improvement Plan

Road Name	Limits (From-To)	Improvement
Florida Intrastate Highway System (FIHS) Roads		
I-4	St Johns River Bridge to Saxon Blvd	Widen to 6 Lanes
I-4	Saxon Blvd to SR 472	Widen to 6 Lanes
I-4	At LPGA Blvd	New Interchange
I-4	SR 472 to I-95	Widen to 6 Lanes
I-4	At Taylor Rd (extension)	New Interchange
I-95	Flagler County Line to SR 40	Widen to 6 Lanes
I-95	SR 40 to LPGA Blvd	Widen to 8 Lanes
I-95	LPGA Blvd to US 92	Widen to 6 Lanes
I-95	US 92 to Brevard County	Widen to 6 Lanes
I-95	At Pioneer Trail	New Interchange
SR 40	Cone Rd to Tymber Creek Rd	Widen to 4 Lanes
SR 40	SR 11 to Cone Rd	Widen to 4 Lanes
SR 40	Lake County to SR 11	Widen to 4 Lanes
Non-FIHS State Roads		
US 17/92	Enterprise Rd to Highbanks Rd	Widen to 4 Lanes
US 17/92	Plantation Rd to Seminole County	Widen to 4 Lanes
SR 5A (Nova Rd)	US 1 to Wilmette Av	Widen to 4 Lanes
SR 5A (Nova Rd)	Wilmette Av to Flomich Av	Widen to 6 Lanes
SR 5A (Nova Rd)	Herbert St to Village Trail	Widen to 4 Lanes
SR 5A (Nova Rd)	Village Trail to US 1	Widen to 4 Lanes
SR 15A	US 17 to Greens Dairy Rd	Widen to 4 Lanes
SR 15A	Greens Dairy Rd to Plymouth Av	Widen to 4 Lanes
SR 15A	Beresford Av to US 17/92	Widen to 4 Lanes
SR 44	Summit Av to I-4	Widen to 4 Lanes
SR 44	I-4 to Pioneer Trail	Widen to 4 Lanes
SR 44	Pioneer Trail to SR 415	Widen to 4 Lanes
SR 442 (Indian River Blvd)	I-95 to Air Park Rd	Widen to 4 Lanes
SR 442 (Indian River Blvd)	Air Park Rd to US 1	Widen to 4 Lanes
US 1	SR 40 to Park Av	Intersection Improvements

US 17	SR 40 to Ponce DeLeon Blvd	Widen to 4 Lanes
US 17/92	SR 15A (Taylor Rd) to SR 472	Widen to 6 Lanes
SR 415	Howland Blvd to Seminole County	Widen to 4 Lanes
SR 483 (Clyde Morris Blvd)	US 92 to Beville Rd	Widen to 6 Lanes
US 92	Nova Rd to US 1	Widen to 6 Lanes
SR 400 (Beville Rd)	SR 483 (Clyde Morris Blvd) to Nova Rd	Widen to 6 Lanes
SR 400 (Beville Rd)	Nova Rd to US 1	Widen to 6 Lanes
SR 415	SR 44 to Howland Blvd	Widen to 4 Lanes
SR 421 (Dunlawton Av)	Nova Rd to Spruce Creek Rd	Widen to 6 Lanes
SR 430 (Mason Av)	SR 483 (Clyde Morris Blvd) to Seabreeze Bridge	Widen to 6 Lanes
SR 442 (Indian River Blvd)	) Airport Rd to I-95	Extend as 2 Ln Rd
Local Roads		
Airport Rd	Taylor Rd/Williamson Blvd to Summertrees Rd	Widen to 4 Lanes
Atlantic Av	Flagler Av to 6th St	Widen to 3 Lanes
Big Tree Rd	Nova Rd to Kenilworth Av	Widen to 3 Lanes
CR 92 Widen to 4 Lanes	SR 15A to US 17/92	Widen to 4 Lanes
Clyde Morris Blvd	Falls Way to LPGA Blvd	Widen to 4 Lanes
Enterprise Rd	Saxon Blvd to Highbanks Rd	Widen to 4 Lanes
Enterprise Rd	Highbanks Rd to Deltona Blvd	Widen to 4 Lanes
Dunn Ave	Williamson Blvd to Bill France Blvd	Extend as 2 Ln Rd
Dirksen/BeBary (realign)	I-4 to Providence Blvd	Widen to 4 Lanes
Howland Blvd	Extension SR472/I-4 to Deltona High School	Extend as 4 Ln Rd
Howland Blvd	Deltona High School to Providence Blvd	Widen to 4 Lanes
Howland Blvd	Elkcam Blvd to Newmark Dr	Widen to 4 Lanes
LPGA Blvd	Jimmy Ann Dr to Nova Rd (SR 5A)	Widen to 4 Lanes
Madeline Av	Sauls Rd to US 1	Extend as 3 Ln Rd
Providence/Idlewise/Sixma	Lake Helen Osteen Rd to Catalina Blvd	Extend as 2 Ln Rd
Providence Blvd	Elkcam Blvd to Ft. Smith Blvd	Widen to 4 Lanes
Providence Blvd	Ft. Smith Blvd to Tivoli Dr	Widen to 4 Lanes
Saxon Blvd	US 17/92 to W.of Enterprise Rd (4 lane portion)	Widen to 4 Lanes
Saxon Blvd	Normandy Blvd to Sumatra Av	Widen to 4 Lanes
Saxon Blvd	Sumatra Av to Tivoli Dr	Widen to 4 Lanes
Westside Connector (Fatio Rd)	SR 44 to Beresford Av	Extend as 2 Ln Rd
Westside Connector (Hamilton Av)	20th St to French Av	Extend as 2 Ln Rd
W. Volusia Bltwy/Veteran's Memorial Pkwy	SR 472 to Graves Av	Widen to 4 Lanes
Williamson Blvd.	Indigo Dr to US 92	Widen to 4 Lanes
Beresford Av	Blue Lake Av to Summit Av	Extend as 2 Ln Rd
Dunn Av	LPGA Blvd to Williamson Blvd	Extend as 2 Ln Rd
Dunn Av	Williamson Blvd to Clyde Morris Blvd	Widen to 4 Lanes
Elkcam Blvd	Riverhead Dr to SR 415	Extend as 2 Ln Rd
Enterprise Rd	US 17/92 to Saxon Blvd	Widen to 6 Lanes
Frontage Road (along I-4)	Summit Av to Orange Camp Rd	Extend as 2 Ln Rd

and Realignment		
Frontage Road (along I-4)	Orange Camp Rd to SR 472	Extend as 2 Ln Rd
Howland Blvd	Providence Blvd to Elkcam Blvd	Widen to 4 Lanes
Howland Blvd	Newmark Dr to Courtland Blvd	Widen to 4 Lanes
Howland Blvd	Courtland Blvd to SR 415	Widen to 4 Lanes
LPGA Blvd	US 1 to Nova Rd	Widen to 4 Lanes
LPGA Blvd	US 92 to Tomoka Farms Rd	Extend as 2 Ln Rd
Providence Blvd	Tivoli Dr to Doyle Rd	Widen to 4 Lanes
Rhode Island Av	Veteran's Memorial Pkwy to Normandy Blvd	Extend as 2 Ln Rd
Saxon Blvd	Westside Connector to US 17/92	Extend as 2 Ln Rd
Saxon Blvd	Enterprise Rd. to I-4	Widen to 6 Lanes
Spruce Creek Rd	Herbert St to Dunlawton Av	Extend as 2 Ln Rd
Tymber Creek Rd	Riverbend Rd to LPGA Blvd	Extend as 2 Ln Rd
W. Volusia Bltwy/Veteran's Mem Pkwy	Graves Av to Harley Strickland Blvd	Widen to 4 Lanes
Williamson Blvd	Current terminus to Pioneer Trail/Turnbull Bay	Extend as 2 Ln Rd
Airport Rd	Summer Tree to Pioneer Trail	Widen to 4 Lanes
Airport Rd	Pioneer Trail to SR 44	Extend as 2 Ln Rd
Airport Rd	SR 44 to SR 442	Extend as 2 Ln Rd
Blue Lake Av	Orange Camp Rd to SR 472	Extend as 2 Ln Rd
Deltona Blvd	Enterprise Rd to DeBary Av	Widen to 4 Lanes
Dirksen Dr	US 17/92 to I-4	Widen to 4 Lanes
Enterprise Rd	Deltona Blvd to Main St/Lexington Av	Widen to 4 Lanes
Hand Av	Tymber Creek Rd to Williamson Blvd	Extend as 2 Ln Rd
Hand Av	Williamson Rd to Nova Rd	Widen to 4 Lanes
Knox Bridge	At Highbridge Rd	Reconstruct Bridge
LPGA Blvd	Clyde Morris Blvd to I-95	Widen to 6 Lanes
LPGA Blvd	I-95 to Tymber Creek Rd extension	Widen to 4 Lanes
Madeline Av	LPGA Blvd extension to Williamson Blvd	Extend as 2 Ln Rd
Main Street Bridge	Beach St to Halifax Av	4 Lane High Rise Bridge
Mason Av	Williamson Blvd to Bill France Blvd	Widen to 4 Lanes
Memorial Bridge (Orange Av)	City Island to Peninsula Dr	4 Lane High Rise Bridge
Orange Camp Rd	US 17/92 to I-4	Widen to 4 Lanes
Pioneer Trail	Tomoka Farms Rd to Turnbull Bay Rd	Widen to 4 Lanes
Providence/Idlewise/Sixma	Catalina Blvd to Howland Blvd	Extend as 2 Ln Rd
Providence Blvd	Howland Blvd to Elkcam Blvd	Widen to 4 Lanes
Rhode Island Av	Westside Connector to US 17/92	Extend as 2 Ln Rd
Saxon Blvd	Tivoli Dr to Providence Blvd	Widen to 4 Lanes
Taylor Rd (CR 421)	I-4 to Tomoka Farms Rd (see I-4 for Interchange)	Extend as 2 Ln Rd
Taylor Rd (CR 421)	Tomoka Farms Rd to Williamson Blvd/Airport Rd	Widen to 4 Lanes
Tomoka Farms Rd (CR 415)	Taylor Rd to SR 44	Widen to 4 Lanes
Westside Connector v	Beresford Av to 20th/Hamilton A	New 2 Lane Corridor
Westside Connector	French Av to Saxon Blvd	New 2 Lane Corridor

W. Volusia Bltwy/ Veteran's Mem Pkwy	SR 44 to SR 472	Widen to 4 Lanes
Williamson Blvd	Hand Av to Indigo/Dunn Av	Widen to 4 Lanes
Williamson Blvd	Beville Rd to Taylor Rd	Widen to 4 Lanes
Yorktowne Blvd	Dunlawton Av to Taylor Rd	Extend as 2 Ln Rd

4.2 Employment and Demographic Forecast

The BEBR collects economic and demographic data for Florida and its local areas. It also conducts economic, demographic, and public policy research and distributes data and research findings throughout the state and the nation. The BEBR prepares countywide population projections annually for all counties in Florida. These projections include a low-, medium-, and high-range projection for each county. In the Volusia County 2020 Long Range Transportation Plan, the mid-range population projections were adopted. Figure 4.1 and 4.2 shows low-, medium-, and high-range projections of households and employment, respectively, for Volusia County. Tables 4.4 and 4.5 summarize low-, medium-, and high-range projections of households and employment by year.

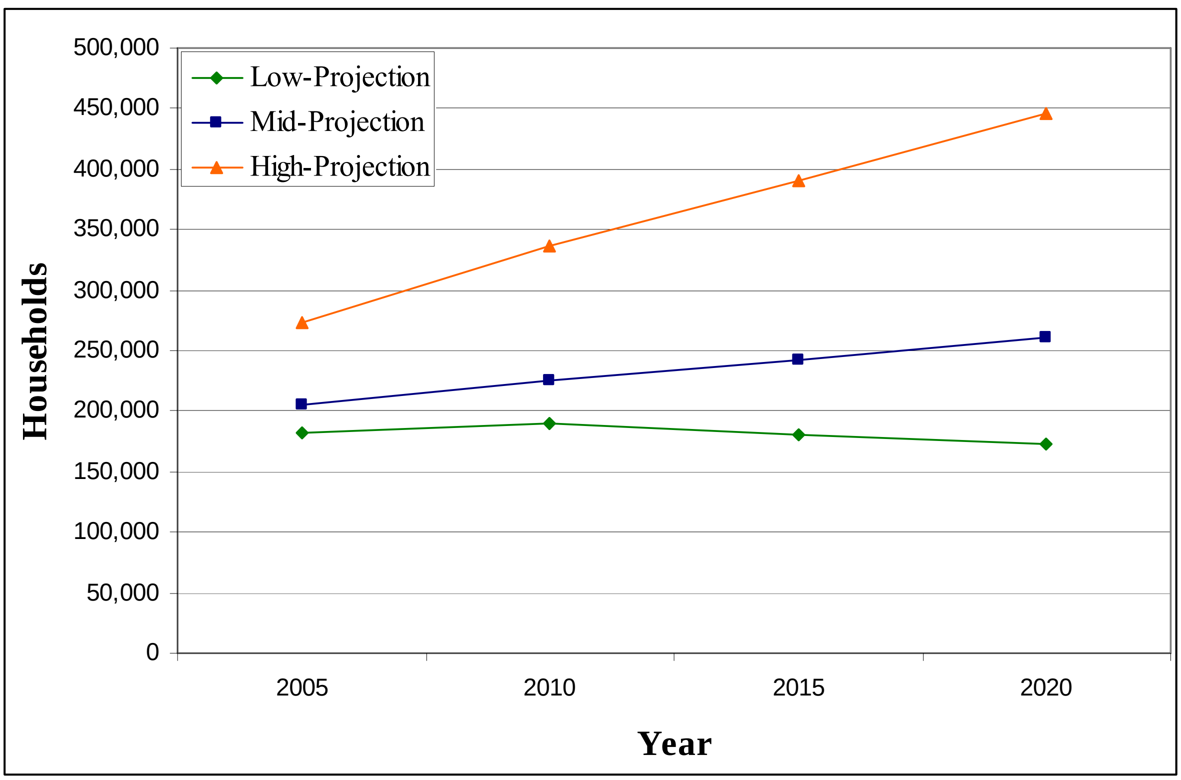


Figure 4.1 Household Forecast by Projection Range

Table 4.4 Household Forecast by Projection Range

Projection	2005	2010	2015	2020
Low	182,322	189,130	181,175	173,220
Midium	204,746	225,568	243,040	260,513
High	272,862	336,257	390,814	445,369

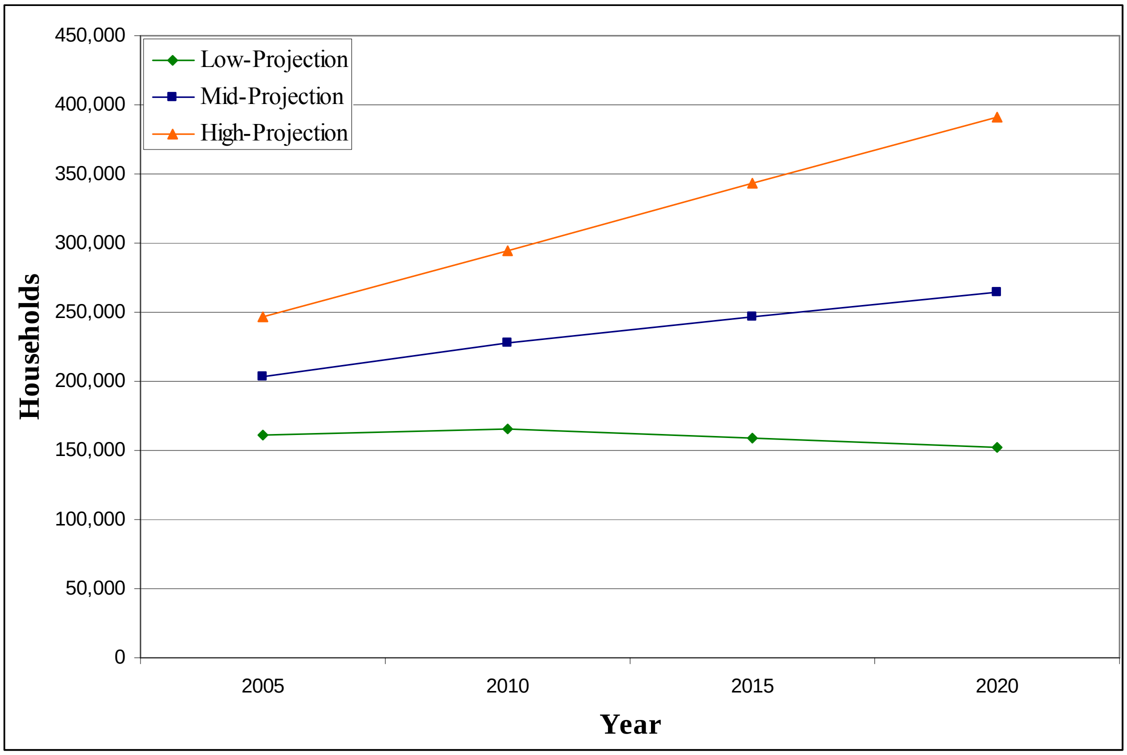


Figure 4.2 Employment Forecast by Projection Range

Table 4.5 Employment Forecast by Projection Range

Projection	2005	2010	2015	2020
Low	161,024	165,680	158,907	152,135
Midium	203,511	228,159	246,291	264,423
High	246,774	294,565	342,860	391,155

4.3 Test Scenarios

In this study, five scenarios are developed based on three transportation improvement plans and three ranges of socioeconomic forecasts. These five scenarios are shown in Table 4.6.

Table 4.6 Test Scenarios for Volusia County

Scenarios	Transportation Improvement Plans	Socioeconomic Projections
1	Final	Medium
2	Alternative 2	Medium
3	Alternative 3	Medium
4	Final	Low
5	Final	High

In addition to the base year data described in Chapter 3, scenario data are required for these five scenarios. Scenario data include FSUTMS networks representing different alternatives and annual control totals to UrbanSim for different socioeconomic projections. The FSUTMS network had been coded based on the base year transportation network with transportation improvements listed in the LRTP.

5. MODEL SPECIFICATION AND ESTIMATION

Estimation is an important process in implementing UrbanSim. UrbanSim consists of submodels that need to be estimated from historical data collected from the study area. This allows UrbanSim to simulate future developments based on the past development pattern.

The submodels of UrbanSim include the Accessibility Model, Economic Transition Model, Demographic Transition Model, Employment Mobility Model, Household Mobility Model, Employment Location Choice Model, Household Location Choice Model, Developer Model, and Land Price Model. UrbanSim generates model coefficients for some of these submodels, but model parameters need to be estimated by users for the Employment Location Choice Model, Household Location Choice Model, Developer Model, and Land Price Model. All the models except the Land Price Model use a multinomial logit model; the Land Price Model uses a hedonic regression model.

Model coefficient estimations in this study were accomplished by using the Limdep econometric software. In the following sections, descriptions of UrbanSim structure and each model will be presented, along with detailed results from model estimations as applicable.

5.1 UrbanSim Structure

As previously mentioned, UrbanSim consists of nine submodels. The Accessibility Model is responsible for providing accessibility indices, which are used in the Employment Location Choice Model and Household Location Choice Model. The Economic Transition Model simulates job creation and loss by computing the sectoral employment growth or decline from the preceding year, based on annual employment control totals. The Demographic Transition Model simulates births and deaths in households. The Employment Mobility Model (or Employment Relocation Choice Model) determines which jobs will move from their current locations during a particular year. The Household Mobility Model simulates household movement. The Employment Location Choice Model is responsible for determining the location of jobs. The Household Location Choice Model chooses a location for a new or moving household. The Developer Model simulates new development and redevelopment. The Land Price Model simulates land prices.

Figure 5.1 depicts the execution sequence of these submodels during each simulation period, as well as two external inputs from a travel demand model and two other user inputs: scenario assumptions and user specified events. These submodels are discussed in the following subsections.

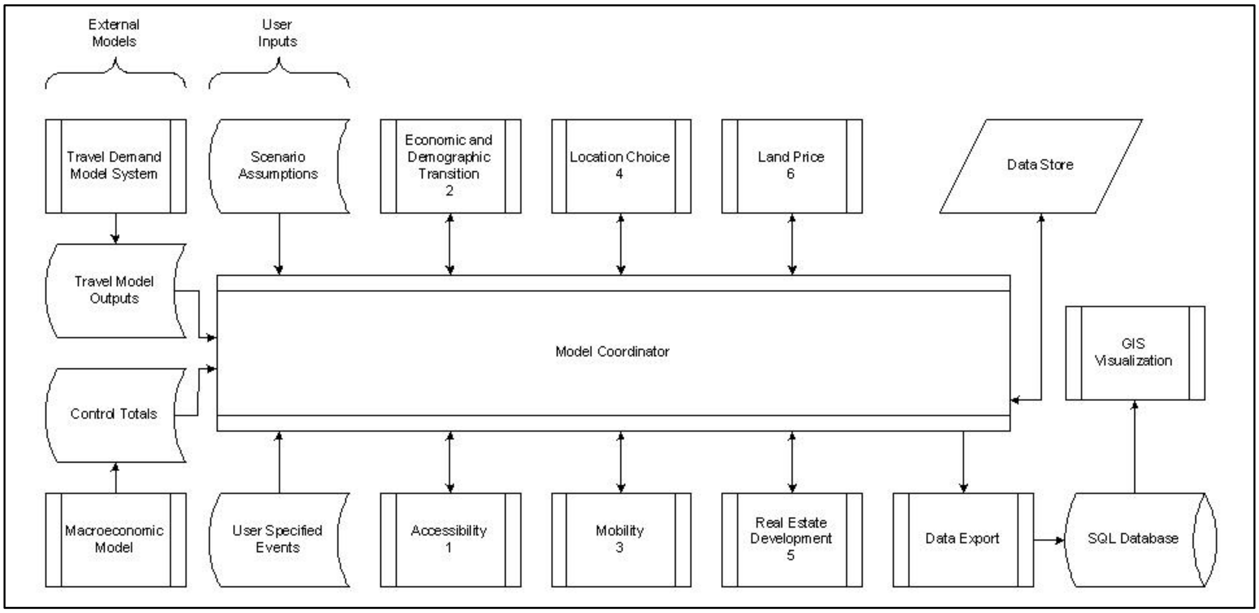


Figure 5.1 UrbanSim Model Structure

5.1.1 Accessibility Model

The Accessibility Model requires four input tables, including travel data that contains the logsum matrix from the output of the travel model, jobs, households, and zones, as shown in Figure 5.2.

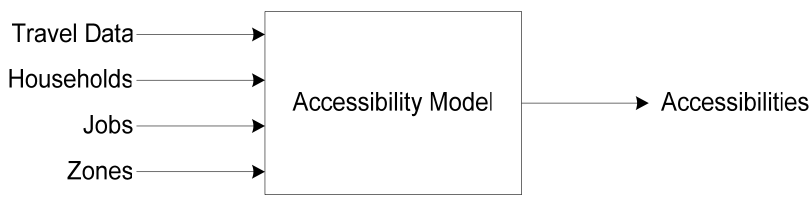


Figure 5.2 Accessibility Model

Since the input data to the FSUTMS are updated for the years of 2002, 2005, 2007, 2010, 2013, 2017, and 2020 in this study, the FSUTMS model runs only for those years. Therefore, the travel impedance remains constant until it is updated with the next travel model output. The activity distribution and the accessibility indices, however, are updated annually because UrbanSim simulates land use change annually.

In this study, several accessibility measures are used, which are listed below.

- Logsum-weighted accessibility
- Travel time to the central business district (CBD) and the airport
- Trip-weighted travel time by single-occupancy vehicle for home-based work trips
- Trip-weighted composite utility by single-occupancy vehicle for home-based work trips

5.1.2 Economic Transition Model

The input and output of the economic transition model is shown in Figure 5.3. Employment is classified into 15 employment sectors, based on the SIC (Standard Industrial Classification) code. The sectoral employment forecasts are exogenous input to UrbanSim and are created from total employment forecasts obtained from the BEBR by assuming that the sectoral distribution is the same as the current distribution.

Figure 5.3 Employment Transition Model

During each simulation period, the Economic Transition Model updates the UrbanSim database by either queuing jobs from sectors that experience growth and are to be placed in the Employment Location Choice Model, or removing jobs from sectors that are declining from the database. In cases where jobs are being removed, the space (land) occupied by these jobs becomes available for other jobs.

5.1.3 Demographic Transition Model

The simulation of births and deaths in households is based on household control totals provided exogenously by the BEBR. Figure 5.4 shows the inputs and outputs of the model. The annual household control totals are given by household type, characterized household size and race of the household head. Since only the total household forecast is available from the BEBR, it is assumed that the distribution of households by type remains unchanged.

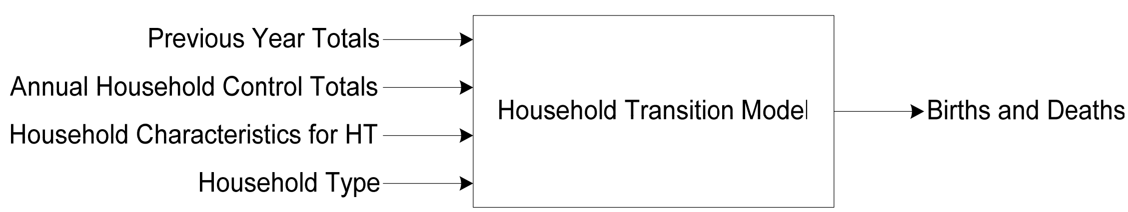


Figure 5.4 Demographic Transition Model

Household births are added to the list of households that will be spatially distributed by the Household Location Choice Model. In the case of household deaths, households are randomly selected and removed from the set of existing households. The housing vacancies created upon their removal become available to other households.

5.1.4 Employment Mobility Model

This model determines the movement of jobs based on annual relocation rates by sector. Jobs are extracted from the InfoUSA employment data. In the employment model, the creation, loss, and movement of individual jobs are simulated. Jobs that are chosen to be moved are subtracted from the current allocated jobs and added to the unallocated new jobs by sector and calculated in the Economic Transition Model. The Employment Location Choice Model will locate both the new jobs and those that will be moved. As jobs are subtracted from the collection of currently allocated jobs, the data on vacant nonresidential space are updated and the space occupied by these jobs becomes available for allocating new jobs in the Employment Location Choice Model. Figure 5.5 illustrates the data flow of the Employment Mobility Model.

Figure 5.5 Employment Mobility Model

5.1.5 Household Mobility Model

The Household Mobility Model (or Household Relocation Choice Model) simulates decisions by households on whether or not to relocate. The mobility probabilities are estimated from the Census Current Population Survey, which provides data on the annual rate of moving and the characteristics of movers and non-movers by the type of move. Figure 5.6 shows the inputs and output of this model.

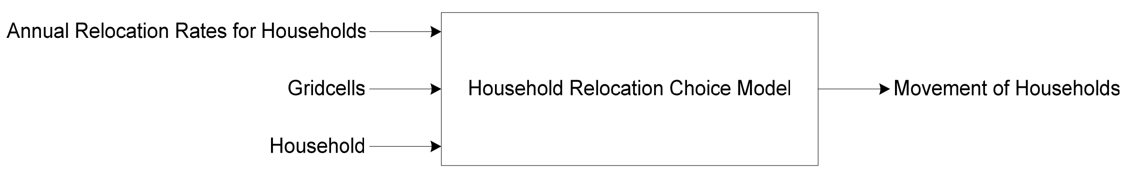


Figure 5.6 Household Mobility Model

Once a household has chosen to move, it is subtracted from the housing stock by cell and is added to the new household database. The Household Location Choice Model will locate both the new and moving households. As movers are subtracted, housing vacancy is updated by making the housing available for occupation in the Household Location Choice Model.

5.1.6 Employment Location Choice Model

The Employment Location Choice Model chooses a location for each job that is either new (from the Economic Transition Model) or is moved within the region (from the Employment Mobility Model). Figure 5.7 depicts the required input files to the model and output from the model. In this study, the jobs are further divided, based on their locations, into home-based jobs and non-home-based jobs. Home-based jobs are defined as jobs located in residential units whereas non-home-based jobs are located on nonresidential properties. Therefore, residential units of a cell must be considered when locating home-based jobs, while non-home-based jobs depend on the total square footage of nonresidential floor space in a cell. Total home-based employment is defined by a maximum rate of home-based employment, which is one of the user's inputs and generally estimated from local employment data.

For each job, a sample of locations with unoccupied space measured in square feet is randomly selected from the set of all possible job locations. Since grid cells are used as the basic geographic unit of analysis, job locations are represented by cells that have nonresidential floor space and residential units. The probability that a job will be located to a particular cell will be estimated by using a multinomial logit model with separate equations estimated for each employment sector.

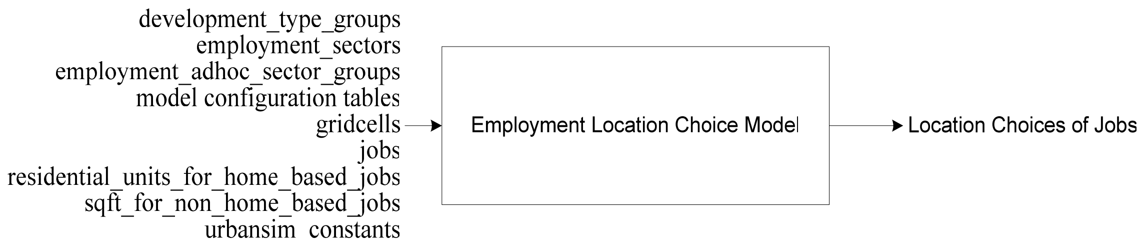


Figure 5.7 Employment Location Choice Model

5.1.7 Household Location Choice Model

The Household Location Choice Model determines a location for each household that is either new (from the Household Transition Model) or is moving (from the Household Mobility Model). For each household, a sample of locations with vacant housing units is randomly selected from the set of all vacant housing units, which also includes those units vacated by movers in the current year. The probability of each alternative in the sample is calculated by using a multinomial logit model calibrated to historical data. The model chooses the most desirable location for the household based on the calculated probability. Figure 5.8 shows the required input files and output from the model.

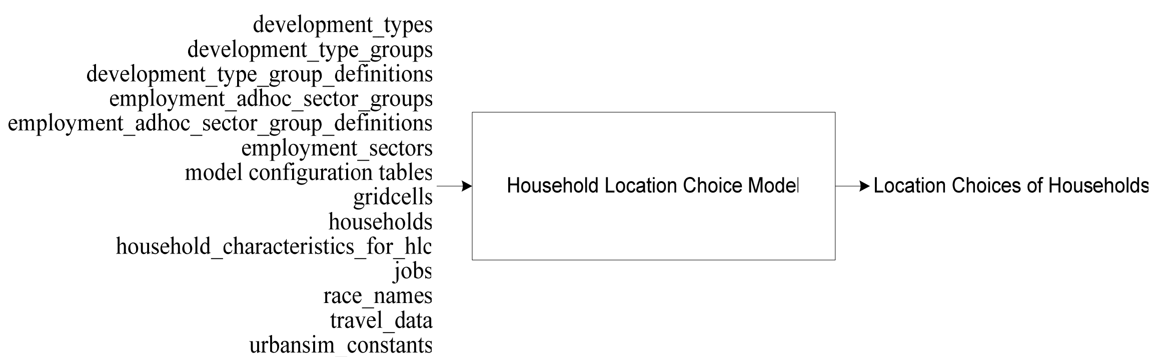


Figure 5.8 Household Location Choice Model

5.1.8 Developer Model

The Developer Model (or the Real Estate Development Model) simulates the construction of new development and redevelopment of existing structures in each grid cell. For each simulation year, the model determines the development types of grid cells and creates a list of possible transition alternatives, as shown in Figure 5.9. There are 25 development types available based on a grid cell's real estate composition. A multinomial logit model is used to calculate the probability of each alternative being chosen and a Monte Carlo sampling process is used to simulate the commitment of development with the estimated probabilities. These commitments are added to the 'development event' queue in the UrbanSim database, to be built as scheduled.

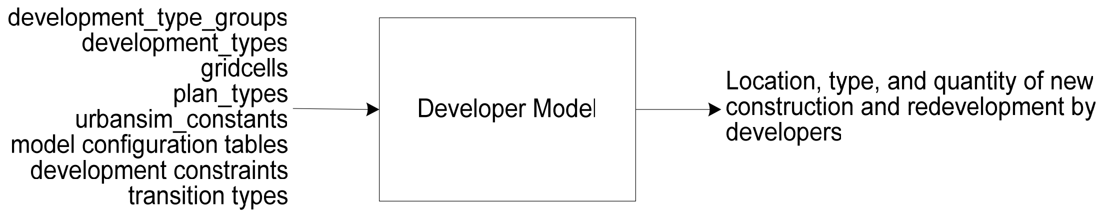


Figure 5.9 Developer Model

The following constraints on development alternatives are considered for this study, and they were assigned to grid cells through spatial processing by using GIS overlay techniques.

- Water
- Wetlands
- Floodplains
- Parks and open space
- National forests
- Steep slopes
- Stream buffers (riparian areas)
- Cities
- Urban growth boundaries
- Major public lands
- Land use plan designation

5.1.9 Land Price Model

The Land Price Model simulates land prices in each grid cell during a simulation period, based on the characteristics of the grid cell and its neighboring cells, including geographical attributes such as percentage of the area that is water or public space, land use mix, density of development, proximity to highways and arterials, accessibilities, land use plan, and demographic attributes. The model is developed using a hedonic linear regression in which land price is modeled as a dependent variable and the characteristics of a grid cell and its neighboring cells are considered to be explanatory variables. To estimate the model, land values are obtained from tax assessor records, which are coded to grid cells in the data preparation process. Figure 5.10 illustrates the input files to and the output from the model.

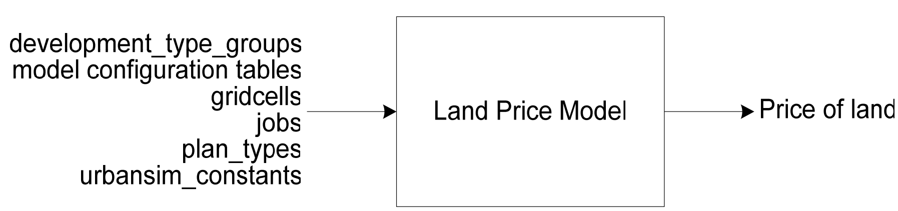


Figure 5.10 Land Price Model

5.2 Model Estimation

Among the submodels described in the previous chapter, the coefficients of four models must be estimated based on the prepared databases. They are the Land Price Model, Household Location

Choice Model, Employment Location Choice Model, and Developer Model. The estimation processes generally have four steps:

1. Specifying models – determining variables to be included in the models and creates specification tables, which are input tables required by UrbanSim.

2. Running estimation data writers – computing variables for the models and generating datasets for model estimation.

3. Estimating coefficients – estimating coefficients of variables in models using Limdep.

4. Updating coefficients in the database – creating coefficient tables corresponding to specification tables.

UrbanSim provides a list of variables that may be implemented in submodels, which are presented in Appendix IV. The user may specify the submodels by selecting variables from the list and must provide specification tables for the submodels to run the estimation data writers. User-specified variables that are not included in the list, such as disutilities from the travel model, may be included in the specification table.

Five estimation data writers, which are stand-alone Java programs for computing variables to be included in the submodels, are listed below:

- Land Price Model Estimation Data Writer
- Housing Location Choice Model Estimation Data Writer
- Employment Location Choice Model Estimation Data Writer
- Employment Home Based Location Choice Model Estimation Data Writer
- Developer Model Estimation Data Writer

These writers generate estimation data that are further analyzed by LIMDEP[6] to estimate coefficients for the UrbanSim submodels. The results of the estimation are presented in the following sections.

5.2.1 Estimation of the Land Price Model

The Land Price Model Estimation Data Writer requires the most input tables. Once the data writer generates estimation data for land prices, they are used by Limdep to estimate the coefficients of the model. The estimated model has an adjusted R^2 of 0.62, indicating that the model explains approximately 62 percent of the variation in land values in the base year. Table 5.1 summarizes the Land Price Model coefficients, along with descriptive statistics.

[6] An econometric software package

Table 5.1 Land Price Model Coefficients

Submodel	Coefficient Name	Estimate	Standard Error	T-statistics	P-Value
-2	constant	0.45423	0.02746	16.54	0.0000
-2	is_near_arterial	0.72954	0.01904	38.32	0.0000
-2	is_near_highway	-0.80253	0.04120	-19.48	0.0000
-2	ln_total_nonresidential_sqft_within_walking_distance	0.31855	0.00341	93.33	0.0000
-2	percent_high_income_households_within_walking_distance	0.02679	0.00042798	62.60	0.0000
-2	percent_mid_income_households_within_walking_distance	0.02407	0.00045309	53.12	0.0000
-2	plantype_1	5.97969	0.03307	180.80	0.0000
-2	plantype_2	6.07226	0.03512	172.91	0.0000
-2	plantype_3	6.13894	0.03562	172.35	0.0000
-2	plantype_4	5.48307	0.13649	40.17	0.0000
-2	plantype_6	9.43481	0.04167	226.42	0.0000
-2	plantype_7	8.28565	0.03783	219.02	0.0000
-2	plantype_8	10.21146	0.06929	147.38	0.0000
-2	plantype_9	7.50680	0.11757	63.85	0.0000
-2	plantype_10	7.47475	0.13459	55.54	0.0000
-2	plantype_11	8.67486	0.30073	28.85	0.0000
-2	plantype_12	7.22658	0.20681	34.94	0.0000
-2	plantype_13	8.37794	0.29997	27.93	0.0000
-2	plantype_15	7.67431	0.07097	108.14	0.0000
-2	plantype_16	5.94542	0.10499	56.63	0.0000
-2	plantype_17	8.99969	0.03267	275.48	0.0000
-2	plantype_18	8.17553	0.46469	17.59	0.0000
-2	plantype_19	8.80478	0.30372	28.99	0.0000
-2	plantype_22	6.21219	0.23013	26.99	0.0000
-2	percent_water	0.02967	0.00127	23.30	0.0000
-2	utility_for_SOV	-0.04127	0.01298	-3.18	0.0015

Notes: is_near_arterial – Indicator for cells near an arterial defined by 1 mile
is_near_highway – Indicator for cells near a highway defined by 1 mile
ln_total_nonresidential_sqft_within_walking_distance – Log of non-residential sq. ft. within walking distance (0.25 mile)
percent_high_income_households_within_walking_distance – Percent of households within walking distance (0.25 mile) that are designated as high-income
percent_mid_income_households_within_walking_distance – Percent of households within walking distance (0.25 mile) that are designated as mid-income
plantype_? – Indicator for plantype "?". There is exactly one variable corresponding to each defined plan type, where "?" is the plan type (e.g. plantype_1, plantype_2)
percent_water – Percent of cell covered by water
utility for SOV – Values from utility function for single occupancy vehicle

5.2.2 Estimation of the Household Location Choice Model

The Housing Location Choice Model Estimation Data Writer requires the following tables from the base year database as input data to generate estimation data, which will be used to calculate the coefficients for the model:

- base_year – 2000 in this study;
- cities – the name of the cities included in the study area;
- counties – the name of the counties included in the study area;
- annual_employment_control_totals – employment forecasts by sector;
- annual_household_control_totals – household forecasts by household type;
- annual_relocation_rates_for_jobs – relocation rates for jobs by sector;
- annual_relocation_rates_for_households – relocation rates for households by household type;
- development_types – definitions of all development types;
- development_type_groups – definitions of all development type groups;
- development_type_group_defintions – the memberships of development types for development type groups;
- employment_sectors – definitions of 15 employment sectors;
- employment_adhoc_sector_groups – employment sector groups such as basic, retail, and service;
- employment_adhoc_sector_group_definitions – the employment sector membership of employment ad hoc sector groups;
- gridcells – gird cells with geographical, environmental, and political information;
- households – all households in the study area;
- households_for_estimation – sample set of households for estimation;
- household_location_choice_model_specification – variables included in the model;
- jobs – all jobs in the study area with corresponding sector ID, grid cell ID, and indication of home-based;
- models – the models to run for each year of the simulation;
- model_variables – all variables provided by UrbanSim;
- plan_types – planned land use;
- primary_uses – names for primary land uses such as residential, mixed use, commercial, industrial, governmental, vacant developable, and undevelopable;
- race_names – the name of racial groups;
- residential_units_for_home_based_jobs – residential units for jobs located in residential units;
- sqft_for_non_home_based_jobs – floor space for jobs that are not located in residential units;
- transition_types – transitions from one development type to another development type;
- travel_data – outputs from the travel model, which has logsums for the Accessibility Model;
- urbansim_constants – constants needed by UrbanSim; and
- zones – traffic analysis zones with travel time to CBD and airport.

The model coefficients and descriptive statistics estimated by Limdep are listed in Table 5.2. In the estimation process, three criteria are applied to optimize the model by screening out insignificant variables. The criteria include the levels of correlation, level of significance, and goodness-of-fit. The variables selected for the model are examined for their correlations to the dependent variable and other independent variables for possible exclusion. After correlated variables are screened, a trial-and-error approach is used to ensure that all variables are significant and that the R-squared value is high.

Two criteria, significance and goodness-of-fit, however, sometimes cannot be satisfied simultaneously. For instance, when all variables are significant, the R-squared value is too low to be accepted. In Table 5.2, variables *ln_commercial_sqft_within_walking_distance* and *cost_to_income_ratio* have higher p-values, compared with 0.01 at the 99-percent confidence level, which means they are not significant in the model. Nonetheless, they are included in the model because removing these two variables from the model causes other variables to become insignificant and eventually results in a low R-squared value of the model. To ensure goodness-of-fit of the model, some of the variables with higher p-values are retained (see the Puget Sound Region study[7] conducted by the Center for Urban Simulation and Policy Analysis, University of Washington).

The (pseudo) R-squared value is defined as:

$$R^2 = 1 - (LogL / LogL^*)$$

where

LogL: log likelihood function and
LogL*: constant only log likelihood function.

[7] Waddell, P. and et al. (2004), "UrbanSim: Model Estimation for the Puget Sound Region," Technical Report CUSPA-04-02, Center for Urban Simulation and Policy Analysis, University of Washington, Available from http://www.urbansim.org/estimation_archive/estimation_main_2004_09_17.pdf.

Table 5.2 Household Location Choice Model Coefficients

Coefficient Name	Estimate	Standard Error	T-Statistics	P-Value
young_household_in_high_density_residential	0.19159	0.06284	3.049	0.0023
percent_minority_households_within_walking_distance	0.01671	0.00070	23.958	0.0000
residential_units_when_household_has_children	-0.00503	0.00059	-8.459	0.0000
ln_commercial_sqft_within_walking_distance	-0.00654	0.00346	-1.891	0.0586
is_near_arterial	-0.24939	0.02167	-11.508	0.0000
cost_to_income_ratio	-0.01717	0.00991	-1.733	0.0831
ln_retail_within_walking_distance_fewer_cars_than_workers	-0.20372	0.05465	-3.728	0.0002
income_and_ln_residential_units	0.0000015	0.00000026	5.609	0.0000
building_age	-0.000067	0.00001111	-6.067	0.0000
percent_low_income_households_within_walking_distance_if_low_income	0.22803	0.00425	53.597	0.0000
percent_mid_income_households_within_walking_distance_if_mid_income	0.08292	0.00061	-7.731	0.0000
percent_high_income_households_within_walking_distance_if_high_income	0.09591	0.00084	114.534	0.0000
ln_residential_units_within_walking_distance	-0.13021	0.01684	136.846	0.0000
ln_total_improvement_value	-0.02599	0.00478	-5.435	0.0000

5.2.3 Estimation of the Employment Location Choice Model

Based on job locations, the Employment Location Choice Model is separated into two submodels: the Employment Non-Home-Based Location Choice Model and the Employment Home-Based Location Choice Model. As discussed in the previous chapter, the Employment Non-Home Based Location Choice Model models job locations on nonresidential properties, whereas the Employment Home-Based Location Choice Model locates jobs to residential units.

The Employment Non-Home Based Location Choice Model consists of 15 submodels, one for each employment sector, based on the SIC codes shown in Table 5.3. In the table, the number of employees is the number of persons whose job belongs to that particular sector, which was obtained from the InfoUSA database.

Table 5.3 Employment Sector Definitions

Sector	SIC	Number of Employees
Resource	01 ~ 14 (except 07)	364
Construction	15 ~ 17	10,700
Manufacturing - Other	37	966
Manufacturing - Aviation	20 ~ 39 (except 37)	15,405
Transportation	40 ~ 47	4,213
Communications, Electric, Gas, and Sanitary Services	48 ~ 49	2,390
Wholesale Trade	50 ~ 51	10,450
Eating and Drinking Places	58	18,380
Other Retail Trade	52 ~ 59 (except 58)	25,145
Finance, Insurance and Real Estate	60 ~ 67	9,670
Producer Services	07, 73, 81, 87	10,135
Consumer Services	70, 72, 75 ~ 79, 83, 84, 86, 88, 89	19,606
Health Services	80	15,211
Education	82	11,678
Public Administration	91 ~ 99	24,551
Total		178,864

The Employment Location Choice Model Estimation Data Writers require the following tables from the base year database:

- base_year – 2000 in this study
- cities – the name of the cities included in the study area
- county – the name of the counties included in the study area
- jobs_for_estimation_non_home_based (or jobs_for_estimation_home_based) – sample of non-home based jobs in the study area

Table 5.4 summarizes the coefficients for the 15 submodels in the Employment Non-Home-Based Location Choice Model. The coefficients of the Employment Home-Based Location Choice Model are presented in Table 5.5. Some of the variables are insignificant, but they are kept in the model to improve the model R-squared value.

Table 5.4 Employment Non-Home Based Location Choice Model Coefficients

Submodel	Coefficient Name	Estimate	Standard Error	T-Statistic	P-Value
1	building_age	-0.00086	0.00015	-5.770	0.0000
1	is_near_arterial	2.7173	0.44212	6.146	0.0000
1	ln_commercial_sqft	0.0435	0.03024	1.439	0.1502
1	ln_distance_to_highway	0.7643	0.13669	5.591	0.0000
1	ln_industrial_sqft_within_walking_distance	0.2005	0.02655	7.550	0.0000
1	ln_total_improvement_value	-1.2560	0.09050	-13.879	0.0000
1	ln_work_access_to_employment	15.6167	6.85738	2.277	0.0228
1	ln_work_access_to_population	-15.6339	6.85737	-2.280	0.0226
1	percent_mid_income_households_within_walking_distance	0.0180	0.00295	6.114	0.0000
1	service_sector_employment_within_walking_distance	0.0385	0.00365	10.557	0.0000
2	building_age	0.000097	0.00003	2.980	0.0029
2	ln_commercial_sqft_within_walking_distance	-0.2172	0.01280	-16.963	0.0000
2	ln_distance_to_highway	-0.1171	0.03273	-3.579	0.0003
2	ln_industrial_sqft_within_walking_distance	0.0777	0.00833	9.324	0.0000
2	ln_residential_units	0.1399	0.02648	5.284	0.0000
2	ln_total_employment_within_walking_distance	2.3084	0.04368	52.844	0.0000
2	ln_total_improvement_value	-0.9980	0.02683	-37.203	0.0000
2	ln_work_access_to_employment	-15.2754	1.89963	-8.041	0.0000
2	ln_work_access_to_population	15.2504	1.89958	8.028	0.0000
2	trip_weighted_travel_time_from_zone_for_SOV	0.2128	0.01457	14.603	0.0000
3	building_age	0.00071	0.00006	12.196	0.0000
3	is_near_arterial	1.2463	0.18397	6.775	0.0000
3	ln_commercial_sqft_within_walking_distance	-0.1101	0.01249	-8.813	0.0000
3	ln_distance_to_highway	-0.2041	0.04005	-5.097	0.0000
3	ln_industrial_sqft_within_walking_distance	0.3465	0.01365	25.379	0.0000
3	ln_residential_units	0.4979	0.04407	11.298	0.0000
3	ln_total_improvement_value	-0.3207	0.04126	-7.774	0.0000
3	percent_low_income_households_within_walking_distance	-0.0683	0.01347	-5.071	0.0000
3	service_sector_employment_within_walking_distance	0.0560	0.00206	27.128	0.0000
3	trip_weighted_travel_time_from_zone_for_SOV	-0.1107	0.02183	-5.073	0.0000

4	building_age	-0.00006	0.00002	-2.963	0.0030
4	is_in_floodplain	0.20087	0.05138	3.909	0.0001
4	is_in_wetland	-0.41662	0.12194	-3.417	0.0006
4	ln_commercial_sqft	0.01314	0.00469	2.799	0.0051
4	ln_industrial_sqft_within_walking_distance	0.29611	0.45476	65.113	0.0000
4	ln_residential_units_within_walking_distance	-0.13952	0.02009	-6.945	0.0000
4	ln_total_improvement_value	-0.26002	0.01142	-22.763	0.0000
4	ln_work_access_to_population	-1.26429	0.62261	-2.031	0.0423
4	percent_low_income_households_within_walking_distance	-0.03173	0.00447	-7.104	0.0000
4	retail_sector_employment_within_walking_distance	0.02450	0.15704	15.600	0.0000
4	service_sector_employment_within_walking_distance	0.02651	0.00079	33.759	0.0000
4	trip_weighted_travel_time_from_zone_for_SOV	0.02991	0.00735	4.067	0.0000
5	building_age	-0.00028	0.00003	-10.431	0.0000
5	is_in_wetland	-0.68934	0.13744	-5.016	0.0000
5	is_near_arterial	-0.27828	0.05611	-4.960	0.0000
5	ln_commercial_sqft_within_walking_distance	-0.77111	0.00542	-14.222	0.0000
5	ln_distance_to_highway	-0.07336	0.01459	-5.027	0.0000
5	ln_industrial_sqft_within_walking_distance	0.01769	0.00606	2.921	0.0035
5	ln_residential_units	-0.30537	0.02457	-12.429	0.0000
5	ln_total_improvement_value	-0.37083	0.01245	-29.780	0.0000
5	percent_high_income_households_within_walking_distance	-0.02700	0.00352	-7.675	0.0000
5	percent_low_income_households_within_walking_distance	-0.01671	0.00420	-3.975	0.0001
5	percent_mid_income_households_within_walking_distance	0.01252	0.00087	14.355	0.0000
5	retail_sector_employment_within_walking_distance	0.03565	0.00158	22.601	0.0000
5	service_sector_employment_within_walking_distance	0.02954	0.00088	33.671	0.0000
5	trip_weighted_travel_time_from_zone_for_SOV	0.02436	0.00718	3.396	0.0007
6	building_age	-0.00060	0.00004	-14.026	0.0000
6	is_in_floodplain	1.34819	0.07651	17.621	0.0000
6	is_in_wetland	-1.51474	0.25184	-6.015	0.0000
6	is_near_arterial	0.75225	0.09268	8.117	0.0000
6	is_near_highway	-5.15717	0.44625	-11.557	0.0000
6	ln_distance_to_highway	-0.69961	0.04251	-16.456	0.0000

6	ln_industrial_sqft_within_walking_distance	0.31031	0.00884	35.093	0.0000
6	ln_residential_units	-0.59928	0.04357	-13.755	0.0000
6	ln_total_improvement_value	0.10217	0.04211	2.426	0.0153
6	ln_total_nonresidential_sqft_within_walking_distance	-0.86265	0.04610	-18.712	0.0000
6	ln_work_access_to_population	-1.37223	0.16234	-8.453	0.0000
6	percent_high_income_households_within_walking_distance	-0.00780	0.00289	-2.700	0.0069
6	percent_low_income_households_within_walking_distance	0.22666	0.00282	8.039	0.0000
6	retail_sector_employment_within_walking_distance	0.08774	0.00304	28.904	0.0000
6	service_sector_employment_within_walking_distance	0.43371	0.00140	31.002	0.0000
6	trip_weighted_travel_time_from_zone_for_SOV	-0.03002	0.01193	-2.516	0.0119
7	building_age	-0.00036	0.21369	-16.814	0.0000
7	is_near_arterial	0.25577	0.44853	5.702	0.0000
7	ln_commercial_sqft	-0.14058	0.36593	-38.417	0.0000
7	ln_distance_to_highway	0.23980	0.15913	15.070	0.0000
7	ln_industrial_sqft_within_walking_distance	-0.03033	0.47309	-6.410	0.0000
7	ln_residential_units	-0.48604	0.19344	-25.126	0.0000
7	ln_total_improvement_value	-0.22852	0.99022	-23.078	0.0000
7	ln_work_access_to_population	-0.01682	0.89309	-1.884	0.0596
7	percent_high_income_households_within_walking_distance	-0.00581	0.13873	-4.188	0.0000
7	percent_low_income_households_within_walking_distance	0.00895	0.18910	4.733	0.0000
7	percent_mid_income_households_within_walking_distance	-0.00312	0.91882	-3.390	0.0007
7	retail_sector_employment_within_walking_distance	0.03081	0.13129	23.470	0.0000
7	trip_weighted_travel_time_from_zone_for_SOV	-0.04586	0.55174	-8.312	0.0000
8	basic_sector_employment_within_walking_distance	0.05005	0.20944	23.897	0.0000
8	building_age	0.00013	0.17157	7.318	0.0000
8	ln_commercial_sqft	0.08309	0.59144	14.049	0.0000
8	ln_distance_to_highway	-0.11424	0.91662	-12.463	0.0000
8	ln_industrial_sqft_within_walking_distance	-0.23960	0.12046	-19.890	0.0000
8	ln_total_improvement_value	-0.11608	0.10704	-10.845	0.0000
8	ln_work_access_to_population	-0.00146	0.59914	-2.434	0.0149
8	percent_high_income_households_within_walking_distance	-0.01013	0.19314	-5.243	0.0000
8	percent_mid_income_households_within_walking_distance	-0.01475	0.12946	-11.394	0.0000

8	service_sector_employment_within_walking_distance	0.01877	0.68540	27.390	0.0000
8	trip_weighted_travel_time_from_zone_for_SOV	0.02144	0.62912	3.408	0.0007
9	building_age	0.00009	0.00004	2.408	0.0161
9	is_near_arterial	1.15948	0.15578	7.443	0.0000
9	ln_commercial_sqft	-0.17623	0.01451	-12.148	0.0000
9	ln_distance_to_highway	-0.09283	0.02962	-3.134	0.0017
9	ln_industrial_sqft_within_walking_distance	-0.07593	0.01237	-6.139	0.0000
9	ln_total_employment_within_walking_distance	1.72079	0.05330	32.286	0.0000
9	ln_total_improvement_value	-0.80012	0.02723	-29.385	0.0000
9	retail_sector_employment_within_walking_distance	0.03817	0.00229	16.688	0.0000
9	service_sector_employment_within_walking_distance	-0.01210	0.00114	-10.661	0.0000
9	trip_weighted_travel_time_from_zone_for_SOV	0.15714	0.01277	12.306	0.0000
10	basic_sector_employment_within_walking_distance	0.81920	0.00228	35.990	0.0000
10	building_age	-0.00008	0.00002	-3.947	0.0001
10	is_in_floodplain	-0.61138	0.07022	-8.707	0.0000
10	is_in_wetland	-0.52306	0.18562	-2.818	0.0048
10	is_near_arterial	1.19371	0.07722	15.459	0.0000
10	ln_commercial_sqft	0.08481	0.00711	11.925	0.0000
10	ln_distance_to_highway	-0.10897	0.01118	-9.748	0.0000
10	ln_industrial_sqft_within_walking_distance	-0.19026	0.11976	-15.887	0.0000
10	ln_residential_units	-0.16271	0.01645	-9.889	0.0000
10	ln_total_improvement_value	0.32815	0.01939	16.925	0.0000
10	ln_total_nonresidential_sqft_within_walking_distance	-0.87375	0.02305	-37.903	0.0000
10	ln_work_access_to_population	-0.24165	0.00073	-3.331	0.0009
10	percent_high_income_households_within_walking_distance	-0.42985	0.00136	-3.149	0.0016
10	retail_sector_employment_within_walking_distance	0.03898	0.00127	30.724	0.0000
10	trip_weighted_travel_time_from_zone_for_SOV	-0.07484	0.00852	-8.781	0.0000
11	basic_sector_employment_within_walking_distance	0.00680	0.00214	3.171	0.0015
11	building_age	0.00020	0.00003	6.591	0.0000
11	is_in_floodplain	0.21766	0.09137	2.382	0.0172
11	is_near_arterial	0.50682	0.11477	4.416	0.0000
11	ln_commercial_sqft	-0.22457	0.01144	-19.622	0.0000

11	ln_distance_to_highway	-0.17422	0.02518	-6.919	0.0000
11	ln_industrial_sqft_within_walking_distance	-0.06848	0.00976	-7.017	0.0000
11	ln_residential_units	0.32486	0.02329	13.951	0.0000
11	ln_total_employment_within_walking_distance	2.15133	0.03905	55.096	0.0000
11	ln_total_improvement_value	-0.89136	0.02525	-35.307	0.0000
11	ln_work_access_to_population_1	-0.00386	0.00161	-2.393	0.0167
11	percent_mid_income_households_within_walking_distance	-0.00465	0.00123	-3.769	0.0002
11	retail_sector_employment_within_walking_distance	-0.00482	0.00089	-5.392	0.0000
11	trip_weighted_travel_time_from_zone_for_SOV	0.22213	0.01295	17.153	0.0000
12	basic_sector_employment_within_walking_distance	-0.00199	0.00053	-3.716	0.0002
12	building_age	0.00012	0.00003	4.652	0.0000
12	ln_commercial_sqft	-0.26471	0.00830	-31.892	0.0000
12	ln_distance_to_highway	-0.13587	0.01996	-6.808	0.0000
12	ln_industrial_sqft_within_walking_distance	-0.12762	0.00971	-13.146	0.0000
12	ln_residential_units_within_walking_distance	0.17259	0.01978	8.724	0.0000
12	ln_total_employment_within_walking_distance	1.78549	0.02939	60.754	0.0000
12	ln_total_improvement_value	-0.67176	0.01886	-35.623	0.0000
12	ln_work_access_to_population	0.43267	0.00054	8.024	0.0000
12	retail_sector_employment_within_walking_distance	-0.00129	0.00062	-2.083	0.0372
12	trip_weighted_travel_time_from_zone_for_SOV	0.12130	0.00910	13.331	0.0000
13	basic_sector_employment_within_walking_distance	0.06201	0.00234	26.488	0.0000
13	building_age	-0.00008	0.00002	-3.910	0.0001
13	is_in_floodplain	-0.25554	0.05720	-4.468	0.0000
13	is_near_arterial	0.95022	0.05567	17.070	0.0000
13	ln_commercial_sqft	-0.19454	0.00399	-48.796	0.0000
13	ln_distance_to_highway	-0.10363	0.01259	-8.232	0.0000
13	ln_industrial_sqft_within_walking_distance	-0.17096	0.00794	-21.530	0.0000
13	ln_total_improvement_value	-0.24441	0.00871	-28.049	0.0000
13	percent_mid_income_households_within_walking_distance	0.00196	0.00070	2.805	0.0050
13	retail_sector_employment_within_walking_distance	0.32194	0.00153	21.039	0.0000
13	trip_weighted_travel_time_from_zone_for_SOV	-0.09118	0.00619	-14.735	0.0000
14	basic_sector_employment_within_walking_distance	0.06816	0.00291	23.390	0.0000

14	building_age	-0.00046	0.00003	-17.032	0.0000
14	is_in_floodplain	-1.03247	0.07859	-13.137	0.0000
14	is_in_wetland	-1.13226	0.19205	-5.896	0.0000
14	is_near_arterial	0.09605	0.04754	2.020	0.0434
14	is_near_highway	0.47106	0.15830	2.976	0.0029
14	ln_commercial_sqft	-0.24478	0.00486	-50.416	0.0000
14	ln_distance_to_highway	0.15284	0.02399	6.372	0.0000
14	ln_industrial_sqft_within_walking_distance	-0.39828	0.02327	-17.113	0.0000
14	ln_residential_units	-0.64626	0.02372	-27.250	0.0000
14	ln_total_improvement_value	0.14044	0.01961	7.160	0.0000
14	ln_total_nonresidential_sqft_within_walking_distance	-0.54639	0.02334	-23.410	0.0000
14	percent_high_income_households_within_walking_distance	-0.00255	0.00127	-2.012	0.0442
14	percent_low_income_households_within_walking_distance	0.01730	0.00216	8.005	0.0000
14	percent_mid_income_households_within_walking_distance	-0.00526	0.00108	-4.889	0.0000
14	retail_sector_employment_within_walking_distance	0.05901	0.00187	31.575	0.0000
14	trip_weighted_travel_time_from_zone_for_SOV	-0.08179	0.00677	-12.086	0.0000
15	basic_sector_employment_within_walking_distance	0.12450	0.00244	50.959	0.0000
15	building_age	0.00007	0.00002	3.542	0.0004
15	is_in_floodplain	-0.32810	0.05715	-5.741	0.0000
15	is_near_arterial	0.30364	0.05125	5.925	0.0000
15	ln_commercial_sqft	-0.19288	0.00448	-43.036	0.0000
15	ln_distance_to_highway	-0.28834	0.01416	-2.036	0.0418
15	ln_industrial_sqft_within_walking_distance	-0.15875	0.00697	-22.770	0.0000
15	ln_residential_units	-0.16922	0.01774	-9.540	0.0000
15	ln_total_improvement_value	-0.14421	0.01002	-14.394	0.0000
15	percent_high_income_households_within_walking_distance	-0.01077	0.00173	-6.210	0.0000
15	percent_mid_income_households_within_walking_distance	-0.00908	0.00115	-7.878	0.0000
15	retail_sector_employment_within_walking_distance	0.04036	0.00154	26.172	0.0000
15	trip_weighted_travel_time_from_zone_for_SOV	0.04443	.00565	7.863	0.0000

Table 5.5 Employment Home-Based Location Choice Model

Coefficient Name	Estimate	Standard Error	T-Statistic	P-Value
basic_sector_employment_within_walking_distance	0.04489	0.00251	17.915	0.0000
building_age	0.00008	0.000018	4.590	0.0000
is_in_floodplain	-0.42541	0.04662	-9.125	0.0000
is_in_wetland	-0.63773	0.10241	-6.227	0.0000
is_near_arterial	0.39919	0.03556	11.226	0.0000
ln_commercial_sqft_within_walking_distance	-0.10709	0.00595	-18.000	0.0000
ln_distance_to_highway	-0.10171	0.01781	-5.709	0.0000
ln_industrial_sqft_within_walking_distance	-0.09836	0.01938	-5.076	0.0000
ln_residential_units_within_walking_distance	-0.09631	0.01614	-5.968	0.0000
ln_total_improvement_value	-36.93741	19.71882	-1.873	0.0610
ln_total_value	37.05174	19.71949	1.879	0.0603
ln_work_access_to_employment	11.98435	0.95984	12.486	0.0000
ln_work_access_to_population	-11.97930	0.95984	-12.481	0.0000
percent_high_income_households_within_walking_distance	-0.05445	0.00086	-63.648	0.0000
percent_low_income_households_within_walking_distance	-0.09504	0.23042	-41.247	0.0000
retail_sector_employment_within_walking_distance	0.04827	0.00188	25.634	0.0000
service_sector_employment_within_walking_distance	0.03751	0.00100	37.368	0.0000

5.2.4 Estimation of the Developer Model

Input tables required by the Developer Model Estimation Data Writer to estimate the coefficients of the Developer Model are listed below:

- base_year – 2000 in this study;
- cities – the name of the cities included in the study area;
- counties – the name of the counties included in the study area;
- development_constraints – constraints that restrict the development types;
- development_event_history– the development events that occurred prior to the base year;
- developer_model_specification – variables included in the Developer Model;
- development_types – definitions of all development types;
- development_type_groups – definitions of all development type groups;
- development_type_group_defintions – the group membership of development types;
- employment_sectors – definitions of the15 employment sectors;
- employment_adhoc_sector_groups – employment sector groups such as basic, retail, and service;
- employment_adhoc_sector_group_definitions – the employment sector membership of employment ad hoc sector groups;
- gridcells – grid cells with geographical, environmental, and political information;
- households – all households in the study area;
- jobs – all jobs in the study area with corresponding sector ID, grid cell ID, and indication of home-based;
- model_variables – all variables provided by UrbanSim;
- models – the models to run for each year of the simulation;
- plan_types – planned land use;
- primary_uses – names for primary land uses such as residential, mixed use, commercial, industrial, governmental, vacant developable, and undevelopable;
- race_names – the name of racial groups;
- residential_units_for_home_based_jobs – residential units for jobs located in residential units;
- sqft_for_non_home_based_jobs – floor space for jobs that are not located in residential units;
- transition_types – transitions from one development type to another development type;
- urbansim_constants – constants needed for UrbanSim; and
- zones – traffic analysis zones with travel time to CBD and airport.

Since the transition from one development type (the starting development type) to another development type (the ending development type) is expected to depend on different factors, the Developer Model is subdivided into multiple multinomial logit models based on a starting development type. Each submodel consists of several equations, each of which corresponds to an ending development type. Table 5.6 summarizes the coefficients of the Developer Model with a submodel ID and equation ID. The submodel ID is the starting development type of each transition, and the equation ID indicates an ending development type of the transition. Some of the variables are insignificant, but they are included in the model to improve the model's R-squared value.

Table 5.6 Developer Model Coefficients

Sub-model ID	Equation ID	Coefficient Name	Estimate	Standard Error	T-Statistic	P-Value
1	-1	alternative specific constant	-1.79436			
1	1	alternative specific constant	-4.05251			
1	1	is_in_floodplain	-0.47792	0.30581	-1.563	0.1181
1	1	is_near_highway	-1.84421	0.64258	-2.870	0.0041
1	1	n_recent_transitions_to_developed_within_walking_distance	4.71179	0.30480	15.459	0.0000
1	1	service_sector_employment_within_walking_distance	0.01223	0.00776	1.576	0.1150
1	1	ln_distance_to_highway	-0.19392	0.07822	-2.479	0.0132
1	1	ln_total_nonresidential_sqft_within_walking_distance	0.46746	0.04165	11.225	0.0000
1	1	ln_work_access_to_employment	-0.15362	0.11586	-1.326	0.1849
1	2	alternative specific constant	-2.34384			
1	2	commercial_sqft_recently_added_within_walking_distance	-0.01503	0.00299	-5.019	0.0000
1	2	is_in_wetland	-0.83173	0.30635	-2.715	0.0066
1	2	is_near_highway	-1.84421	0.64258	-2.870	0.0041
1	2	n_recent_transitions_to_developed_within_walking_distance	6.41565	0.28243	22.716	0.0000
1	2	percent_same_type_cells_within_walking_distance	-0.08032	0.00480	-16.737	0.0000
1	2	percent_water	0.02858	0.01400	2.041	0.0413
1	2	proximity_to_development	14.51020	2.56402	5.659	0.0000
1	2	percent_mid_income_households_within_walking_distance	-0.00633	0.00205	-3.089	0.0020
1	9	alternative specific constant	-1.28519			
1	9	commercial_sqft_recently_added_within_walking_distance	0.00159	0.00068	2.353	0.0186
1	9	is_in_floodplain	-1.98295	1.03124	-1.923	0.0545
1	9	is_near_arterial	1.12671	0.37006	3.045	0.0023
1	9	is_near_highway	-1.84421	0.64258	-2.870	0.0041
1	9	n_recent_transitions_to_developed_within_walking_distance	4.67796	0.47382	9.873	0.0000
1	9	percent_same_type_cells_within_walking_distance	-0.06940	0.00714	-9.719	0.0000
1	9	proximity_to_development	171.24461	47.23895	3.625	0.0003
1	9	service_sector_employment_within_walking_distance	0.01197	0.00612	1.957	0.0503
1	9	ln_total_nonresidential_sqft_within_walking_distance	0.80052	0.09803	8.166	0.0000

1	9	ln_work_access_to_employment_1	-7.51658	2.07211	-3.627	0.0003
1	10	alternative specific constant	13.16852			
1	10	commercial_sqft_recently_added_within_walking_distance	0.00621	0.00067	9.213	0.0000
1	10	percent_developed_within_walking_distance	-0.44863	0.05383	-8.334	0.0000
1	10	is_near_highway	-1.84421	0.64258	-2.870	0.0041
1	10	n_recent_transitions_to_developed_within_walking_distance	7.23772	1.49017	4.857	0.0000
1	10	service_sector_employment_within_walking_distance	0.05694	0.01585	3.592	0.0003
1	10	ln_total_nonresidential_sqft_within_walking_distance	4.73033	0.60424	7.829	0.0000
2	-1	alternative specific constant	-0.07886			
2	2	alternative specific constant	0.36589			
2	2	commercial_sqft_recently_added_within_walking_distance	0.01279	0.00215	5.948	0.0000
2	2	is_in_floodplain	-0.20885	0.09446	-2.211	0.0270
2	2	is_near_arterial	-0.22683	0.06999	-3.241	0.0012
2	2	ln_distance_to_highway	-0.08508	0.02738	-3.107	0.0019
2	2	ln_residential_units	4.36800	0.12364	35.329	0.0000
2	2	ln_total_job_locations	1.32444	0.32385	4.090	0.0000
2	2	percent_developed_within_walking_distance	-0.01540	0.00417	-3.695	0.0002
2	2	percent_low_income_households_within_walking_distance	-0.01201	0.00453	-2.652	0.0080
2	2	percent_mid_income_households_within_walking_distance	-0.00482	0.00137	-3.525	0.0004
2	2	percent_same_type_cells_within_walking_distance	-0.00748	0.00175	-4.273	0.0000
2	2	proximity_to_development	-4.39303	0.90789	-4.839	0.0000
2	3	alternative specific constant	2.49441			
2	3	commercial_sqft_recently_added_within_walking_distance	0.00986	0.00215	4.582	0.0000
2	3	is_in_floodplain	-0.20885	0.09446	-2.211	0.0270
2	3	is_near_arterial	-0.22683	0.06999	-3.241	0.0012
2	3	ln_distance_to_highway	-0.08508	0.02738	-3.107	0.0019
2	3	ln_residential_units	6.10609	0.15921	38.352	0.0000
2	3	ln_total_job_locations	0.93842	0.36176	2.594	0.0095
2	3	ln_work_access_to_employment	-0.15544	0.01718	-9.050	0.0000
2	3	percent_developed_within_walking_distance	-0.01540	0.00417	-3.695	0.0002
2	3	percent_mid_income_households_within_walking_distance	-0.01017	0.00241	-4.228	0.0000
2	3	percent_same_type_cells_within_walking_distance	-0.03356	0.00180	-18.678	0.0000

2	3	percent_water	0.01421	0.00566	2.512	0.0120
2	3	proximity_to_development	-4.39303	0.90789	-4.839	0.0000
2	9	alternative specific constant	1.72279			
2	9	basic_sector_employment_within_walking_distance	-0.42165	0.07662	-5.503	0.0000
2	9	commercial_sqft_recently_added_within_walking_distance	0.01657	0.00215	7.695	0.0000
2	9	is_near_arterial	2.04022	0.49425	4.128	0.0000
2	9	ln_commercial_sqft	0.33942	0.07578	4.479	0.0000
2	9	ln_total_job_locations	2.90623	0.46429	6.260	0.0000
2	9	ln_work_access_to_employment	-0.15544	0.01718	-9.050	0.0000
2	9	percent_developed_within_walking_distance	-0.01540	0.00417	-3.695	0.0002
2	9	percent_high_income_households_within_walking_distance	-0.07463	0.02323	-3.213	0.0013
2	9	percent_same_type_cells_within_walking_distance	-0.04080	0.00486	-8.394	0.0000
2	10	alternative specific constant	26.07880			
2	10	basic_sector_employment_within_walking_distance	5.33206	0.07495	71.143	0.0000
2	10	commercial_sqft_recently_added_within_walking_distance	0.03536	0.00216	16.360	0.0000
2	10	is_in_floodplain	-11.38287	2.64272	-4.307	0.0000
2	10	is_near_arterial	-41.26462	1.60426	-25.722	0.0000
2	10	ln_commercial_sqft	0.76067	0.22681	3.354	0.0008
2	10	ln_distance_to_highway	7.67870	0.88933	8.634	0.0000
2	10	ln_residential_units	-97.40498	3.72514	-26.148	0.0000
2	10	ln_total_job_locations	46.04368	1.26773	36.320	0.0000
2	10	ln_work_access_to_employment	-0.15544	0.01718	-9.050	0.0000
2	10	percent_developed_within_walking_distance	-0.01540	0.00417	-3.695	0.0002
2	10	percent_high_income_households_within_walking_distance	-0.07463	0.02323	-3.213	0.0013
2	10	percent_mid_income_households_within_walking_distance	-0.26807	0.03138	-8.542	0.0000
2	10	percent_same_type_cells_within_walking_distance	-0.21993	0.03369	-6.528	0.0000
2	10	percent_water	1.05280	0.12003	8.771	0.0000
2	10	proximity_to_development	-74.00106	19.39013	-3.816	0.0001
3	-1	alternative specific constant	-0.22534			
3	3	alternative specific constant	1.90152			
3	3	commercial_sqft_recently_added_within_walking_distance	0.00400	0.00194	2.054	0.0400
3	3	is_in_floodplain	-0.42063	0.09984	-4.213	0.0000

3	3	is_near_arterial	-0.16424	0.05659	-2.902	0.0037
3	3	ln_industrial_sqft_within_walking_distance	-0.55326	0.12828	-4.313	0.0000
3	3	ln_residential_units	4.83371	0.19396	24.921	0.0000
3	3	ln_residential_units_within_walking_distance	-1.97456	0.15058	-13.113	0.0000
3	3	ln_total_improvement_value	0.15646	0.00763	20.498	0.0000
3	3	n_recent_transitions_to_developed_within_walking_distance	0.81442	0.04435	18.364	0.0000
3	3	percent_low_income_households_within_walking_distance	-0.01469	0.00631	-2.330	0.0198
3	3	percent_same_type_cells_within_walking_distance	0.00824	0.00103	7.980	0.0000
3	3	proximity_to_development	-19.83438	0.62418	-31.777	0.0000
3	4	alternative specific constant	6.53210			
3	4	commercial_sqft_recently_added_within_walking_distance	0.00400	0.00194	2.054	0.0400
3	4	is_in_floodplain	-0.42063	0.09984	-4.213	0.0000
3	4	is_near_arterial	-0.16424	0.05659	-2.902	0.0037
3	4	ln_commercial_sqft	-0.13257	0.03377	-3.926	0.0001
3	4	ln_industrial_sqft_within_walking_distance	-0.55326	0.12828	-4.313	0.0000
3	4	ln_residential_units	4.83371	0.19396	24.921	0.0000
3	4	ln_residential_units_within_walking_distance	-1.97456	0.15058	-13.113	0.0000
3	4	ln_total_improvement_value	0.15646	0.00763	20.498	0.0000
3	4	n_recent_transitions_to_developed_within_walking_distance	0.81442	0.04435	18.364	0.0000
3	4	percent_developed_within_walking_distance	0.00330	0.00058	5.689	0.0000
3	4	percent_low_income_households_within_walking_distance	-0.01469	0.00631	-2.330	0.0198
3	4	percent_same_type_cells_within_walking_distance	-0.02062	0.00116	-17.760	0.0000
3	4	proximity_to_development	-19.83438	0.62418	-31.777	0.0000
3	9	alternative specific constant	3.44137			
3	9	commercial_sqft_recently_added_within_walking_distance	0.00536	0.00200	2.685	0.0073
3	9	is_in_floodplain	-0.42063	0.09984	-4.213	0.0000
3	9	is_near_arterial	-0.65723	0.30444	-2.159	0.0309
3	9	ln_commercial_sqft	0.61762	0.08661	7.131	0.0000
3	9	ln_industrial_sqft_within_walking_distance	-0.55326	0.12828	-4.313	0.0000
3	9	ln_residential_units	4.83371	0.19396	24.921	0.0000
3	9	ln_residential_units_within_walking_distance	-1.97456	0.15058	-13.113	0.0000
3	9	ln_total_improvement_value	0.15646	0.00763	20.498	0.0000

3	9	n_recent_transitions_to_developed_within_walking_distance	0.81442	0.04435	18.364	0.0000
3	9	percent_developed_within_walking_distance	-0.04808	0.00444	-10.832	0.0000
3	9	percent_low_income_households_within_walking_distance	-0.01469	0.00631	-2.330	0.0198
3	9	percent_same_type_cells_within_walking_distance	-0.01658	0.00408	-4.059	0.0000
3	9	proximity_to_development	-19.83438	0.62418	-31.777	0.0000
4	-1	alternative specific constant	-2.57544			
4	4	alternative specific constant	32.59300			
4	4	commercial_sqft_recently_added_within_walking_distance	0.00456	0.00243	1.874	0.0610
4	4	ln_total_improvement_value	0.26916	0.01077	24.987	0.0000
4	4	ln_work_access_to_employment	0.14687	0.00462	31.760	0.0000
4	4	n_recent_transitions_to_residential_within_walking_distance	0.86512	0.05217	16.583	0.0000
4	4	percent_developed_within_walking_distance	-0.02295	0.00477	-4.816	0.0000
4	4	percent_low_income_households_within_walking_distance	-0.01832	0.00759	-2.414	0.0158
4	4	percent_same_type_cells_within_walking_distance	-0.00307	0.00095	-3.225	0.0013
4	4	proximity_to_development	-9.43830	1.03706	-9.101	0.0000
4	4	retail_sector_employment_within_walking_distance	-0.02799	0.01037	-2.698	0.0070
4	5	alternative specific constant	32.65053			
4	5	ln_total_improvement_value	0.26916	0.01077	24.987	0.0000
4	5	n_recent_transitions_to_residential_within_walking_distance	0.86512	0.05217	16.583	0.0000
4	5	percent_developed_within_walking_distance	-0.02295	0.00477	-4.816	0.0000
4	5	percent_high_income_households_within_walking_distance	-0.01131	0.00326	-3.474	0.0005
4	5	percent_low_income_households_within_walking_distance	-0.01832	0.00759	-2.414	0.0158
4	5	percent_same_type_cells_within_walking_distance	-0.00307	0.00095	-3.225	0.0013
4	5	percent_water	-0.03970	0.02126	-1.867	0.0619
4	5	proximity_to_development	-9.43830	1.03706	-9.101	0.0000
5	-1	alternative specific constant	-0.13695			
5	5	alternative specific constant	1.73225			
5	5	ln_commercial_sqft	0.17330	0.05129	3.379	0.0007
5	5	ln_residential_units	4.02830	0.64413	6.254	0.0000
5	5	n_recent_transitions_to_developed_within_walking_distance	2.10193	0.12440	16.897	0.0000
5	5	percent_developed_within_walking_distance	-0.29542	0.02745	-10.762	0.0000
5	5	percent_same_type_cells_within_walking_distance	-0.01097	0.00234	-4.689	0.0000

5	5	proximity_to_development	34.20841	4.28229	7.988	0.0000
5	6	alternative specific constant	50.00000			
5	6	ln_residential_units	9.75501	0.88724	10.995	0.0000
5	6	n_recent_transitions_to_developed_within_walking_distance	2.10193	0.12440	16.897	0.0000
5	6	percent_developed_within_walking_distance	-0.29542	0.02745	-10.762	0.0000
5	6	percent_same_type_cells_within_walking_distance	-0.01097	0.00234	-4.689	0.0000
5	6	proximity_to_development	-3.87734	1.51068	-2.567	0.0103
6	-1	alternative specific constant	-2.10162			
6	6	alternative specific constant	50.00000			
6	6	is_near_arterial	-0.77306	0.22205	-3.481	0.0005
6	6	ln_commercial_sqft_within_walking_distance	0.27976	0.09230	3.031	0.0024
6	6	ln_distance_to_highway	0.20386	0.07919	2.574	0.0100
6	6	n_recent_transitions_to_developed_within_walking_distance	2.81991	0.20860	13.518	0.0000
6	6	percent_developed_within_walking_distance	-0.03211	0.00787	-4.078	0.0000
6	6	percent_same_type_cells_within_walking_distance	-0.02054	0.00510	-4.030	0.0001
6	6	proximity_to_development	4.02540	0.27638	14.565	0.0000
6	6	retail_sector_employment_within_walking_distance	-0.16158	0.02153	-7.505	0.0000
6	7	alternative specific constant	50.00000			
6	7	is_near_arterial	-0.77306	0.22205	-3.481	0.0005
6	7	ln_commercial_sqft_within_walking_distance	0.27976	0.09230	3.031	0.0024
6	7	ln_distance_to_highway	0.20386	0.07919	2.574	0.0100
6	7	n_recent_transitions_to_developed_within_walking_distance	2.81991	0.20860	13.518	0.0000
6	7	percent_developed_within_walking_distance	-0.03211	0.00787	-4.078	0.0000
6	7	percent_same_type_cells_within_walking_distance	-0.02054	0.00510	-4.030	0.0001
6	7	retail_sector_employment_within_walking_distance	-0.16158	0.02153	-7.505	0.0000
7	-1	alternative specific constant	-1.72693			
7	7	alternative specific constant	50.00000			
7	7	ln_distance_to_highway	-0.68312	0.22918	-2.981	0.0029
7	7	ln_residential_units	1.82195	0.72466	2.514	0.0119
7	7	ln_total_improvement_value	0.53286	0.06850	7.779	0.0000
7	7	ln_work_access_to_employment	-14.13325	5.58543	-2.530	0.0114
7	7	n_recent_transitions_to_developed_within_walking_distance	0.91613	0.23056	3.974	0.0001

7	7	proximity_to_development	303.05087	129.49546	2.340	0.0193
8	-1	alternative specific constant	-0.59214			
8	8	alternative specific constant	20.59800			
8	8	ln_work_access_to_employment	-0.09848	0.03915	-2.515	0.0119
8	8	n_recent_transitions_to_developed_within_walking_distance	1.48495	0.42032	3.533	0.0004
9	-1	alternative specific constant	-0.00687			
9	9	alternative specific constant	0.11983			
9	9	commercial_sqft_recently_added_within_walking_distance	0.00063	0.00021	3.079	0.0021
9	9	ln_residential_units	0.82730	0.11936	6.931	0.0000
9	9	ln_total_improvement_value	0.28198	0.02118	13.311	0.0000
9	9	ln_total_nonresidential_sqft_within_walking_distance	-0.66249	0.11071	-5.984	0.0000
9	9	percent_same_type_cells_within_walking_distance	0.04603	0.00983	4.683	0.0000
9	9	proximity_to_development	-10.67496	1.29467	-8.245	0.0000
9	10	alternative specific constant	0.49901			
9	10	commercial_sqft_recently_added_within_walking_distance	0.00063	0.00021	3.079	0.0021
9	10	ln_commercial_sqft	0.13773	0.05144	2.677	0.0074
9	10	ln_distance_to_highway	-0.13219	0.05006	-2.640	0.0083
9	10	ln_residential_units	0.82730	0.11936	6.931	0.0000
9	10	ln_total_improvement_value	0.28198	0.02118	13.311	0.0000
9	10	percent_same_type_cells_within_walking_distance	-0.02429	0.00598	-4.064	0.0000
9	10	proximity_to_development	-10.67496	1.29467	-8.245	0.0000
9	12	alternative specific constant	1.87904			
9	12	commercial_sqft_recently_added_within_walking_distance	0.00063	0.00021	3.079	0.0021
9	12	ln_residential_units	0.82730	0.11936	6.931	0.0000
9	12	ln_total_improvement_value	0.28198	0.02118	13.311	0.0000
9	12	percent_same_type_cells_within_walking_distance	-0.03838	0.00607	-6.320	0.0000
9	12	proximity_to_development	-10.67496	1.29467	-8.245	0.0000
10	-1	alternative specific constant	-0.04605			
10	10	alternative specific constant	2.92044			
10	10	commercial_sqft_recently_added_within_walking_distance	-0.00018	0.00007	-2.716	0.0066
10	10	ln_residential_units	7.06998	2.10406	3.360	0.0008
10	10	ln_residential_units_within_walking_distance	-7.19248	2.12694	-3.382	0.0007

10	10	ln_total_improvement_value	0.20721	0.03719	5.572	0.0000
10	10	n_recent_transitions_to_developed_within_walking_distance	3.08124	0.30854	9.986	0.0000
10	10	proximity_to_development	-10.60643	1.32241	-8.021	0.0000
10	13	alternative specific constant	3.78652			
10	13	commercial_sqft_recently_added_within_walking_distance	-0.00018	0.00007	-2.716	0.0066
10	13	ln_residential_units	7.06998	2.10406	3.360	0.0008
10	13	ln_residential_units_within_walking_distance	-7.19248	2.12694	-3.382	0.0007
10	13	ln_total_improvement_value	0.20721	0.03719	5.572	0.0000
10	13	n_recent_transitions_to_developed_within_walking_distance	3.08124	0.30854	9.986	0.0000
10	13	percent_same_type_cells_within_walking_distance	-0.03705	0.00383	-9.685	0.0000
10	13	proximity_to_development	-10.60643	1.32241	-8.021	0.0000
11	-1	alternative specific constant	-0.83660			
11	11	alternative specific constant	50.00000			
11	11	ln_distance_to_highway	-1.10359	0.19280	-5.724	0.0000
11	11	ln_total_improvement_value	0.60392	0.09786	6.171	0.0000
11	11	percent_high_income_households_within_walking_distance	-0.06365	0.02126	-2.994	0.0028
12	-1	alternative specific constant	-0.07756			
12	12	alternative specific constant	2.73927			
12	12	is_near_arterial	-1.12589	0.36837	-3.056	0.0022
12	12	ln_commercial_sqft	-0.11911	0.04346	-2.741	0.0061
12	12	ln_total_improvement_value	0.48210	0.04006	12.035	0.0000
12	12	ln_work_access_to_employment	-0.74948	0.08794	-8.523	0.0000
12	12	percent_same_type_cells_within_walking_distance	0.01641	0.00672	2.442	0.0146
12	13	alternative specific constant	11.73149			
12	13	ln_commercial_sqft	7.68104	1.98659	3.866	0.0001
12	13	ln_work_access_to_employment	-6.59929	1.46202	-4.514	0.0000
12	13	percent_same_type_cells_within_walking_distance	0.01641	0.00672	2.442	0.0146
13	-1	alternative specific constant	-0.03694			
13	13	alternative specific constant	3.88896			
13	13	commercial_sqft_recently_added_within_walking_distance	-0.00019	0.00010	-1.910	0.0561
13	13	ln_total_employment_within_walking_distance	0.50376	0.18485	2.725	0.0064
13	13	ln_total_value	0.22199	0.07465	2.974	0.0029

13	13	ln_work_access_to_employment	-0.62340	0.11882	-5.247	0.0000
13	13	n_recent_transitions_to_developed_within_walking_distance	4.05832	0.70388	5.766	0.0000
14	-1	alternative specific constant	-0.89665			
14	14	alternative specific constant	46.61823			
14	14	ln_total_improvement_value	0.92824	0.37107	2.502	0.0124
14	14	ln_work_access_to_employment	-1.54838	0.66993	-2.311	0.0208
15	-1	alternative specific constant	-0.49167			
15	15	alternative specific constant	50.00000			
15	15	ln_residential_units	-0.63542	0.19126	-3.322	0.0009
15	15	n_recent_transitions_to_developed_within_walking_distance	4.66633	1.06141	4.396	0.0000
16	-1	alternative specific constant	50.00000			
16	16	alternative specific constant	50.00000			
17	-1	alternative specific constant	-0.00736			
17	9	alternative specific constant	-1.35383			
17	9	ln_residential_units_within_walking_distance	2.40786	0.31919	7.544	0.0000
17	9	ln_total_employment_within_walking_distance	-0.41679	0.19886	-2.096	0.0361
17	9	ln_total_value	0.13206	0.03278	4.029	0.0001
17	9	percent_developed_within_walking_distance	-0.04018	0.00479	-8.387	0.0000
17	9	percent_water	0.04912	0.01678	2.928	0.0034
17	17	alternative specific constant	1.81885			
17	17	ln_total_employment_within_walking_distance	0.25796	0.05983	4.311	0.0000
17	17	ln_total_value	0.47703	0.02128	22.415	0.0000
17	17	percent_developed_within_walking_distance	-0.07635	0.00396	-19.277	0.0000
17	18	alternative specific constant	4.52462			
17	18	ln_total_employment_within_walking_distance	0.62553	0.09893	6.323	0.0000
17	18	ln_total_value	0.79003	0.22694	3.481	0.0005
17	18	ln_work_access_to_employment	-0.97143	0.50718	-1.915	0.0554
17	18	percent_developed_within_walking_distance	-0.06514	0.03475	-1.875	0.0608
18	-1	alternative specific constant	-0.05026			
18	18	alternative specific constant	4.96867			
18	18	ln_total_improvement_value	0.33795	0.07686	4.397	0.0000
18	18	ln_work_access_to_employment	11.37335	5.70207	1.995	0.0461

18	18	n_recent_transitions_to_developed_within_walking_distance	2.22418	0.40688	5.466	0.0000
18	18	proximity_to_development	-276.39811	131.55326	-2.101	0.0356
19	-1	alternative specific constant	-1.80821			
19	19	alternative specific constant	50.00000			
20	-1	alternative specific constant	-0.02015			
20	20	alternative specific constant	1.18669			
20	20	ln_residential_units	-10.70933	1.90499	-5.622	0.0000
20	20	n_recent_transitions_to_developed_within_walking_distance	5.24548	0.84071	6.239	0.0000
20	20	proximity_to_development	-5.73469	1.31221	-4.370	0.0000
21	-1	alternative specific constant	-0.00409			
21	21	alternative specific constant	0.53228			
21	21	ln_commercial_sqft	0.43615	0.13731	3.176	0.0015
21	21	ln_distance_to_highway	-2.01638	0.74644	-2.701	0.0069
21	21	ln_work_access_to_employment_1	1.38895	0.53793	2.582	0.0098
22	-1	alternative specific constant	0.00000			
22	22	alternative specific constant	50.00000			
23	-1	alternative specific constant	-4.49631			
23	23	alternative specific constant	50.00000			
23	23	ln_commercial_sqft	0.32654	0.06252	5.223	0.0000
23	23	ln_distance_to_highway	-0.33606	0.16921	-1.986	0.0470
23	23	n_recent_transitions_to_developed_within_walking_distance	0.18634	0.05458	3.414	0.0006
23	23	percent_developed_within_walking_distance	0.02730	0.01503	1.816	0.0693
24	-1	alternative specific constant	0.01041			
24	1	alternative specific constant	-2.15085			
24	1	basic_sector_employment_within_walking_distance	0.86712	0.33406	2.596	0.0094
24	1	is_near_arterial	0.93684	0.11565	8.101	0.0000
24	1	ln_distance_to_highway	-0.06137	0.02367	-2.593	0.0095
24	1	ln_residential_units	8.32621	0.79547	10.467	0.0000
24	1	n_recent_transitions_to_governmental_within_walking_distance	37.68332	3.06149	12.309	0.0000
24	1	percent_high_income_households_within_walking_distance	0.05992	0.00369	16.252	0.0000

24	1	percent_low_income_households_within_walking_distance	0.09507	0.02812	3.381	0.0007
24	1	percent_mid_income_households_within_walking_distance	0.05573	0.00382	14.606	0.0000
24	1	percent_same_type_cells_within_walking_distance	-0.01705	0.00235	-7.251	0.0000
24	1	percent_water	0.02894	0.00647	4.475	0.0000
24	1	service_sector_employment_within_walking_distance	0.04658	0.01891	2.463	0.0138
24	2	alternative specific constant	0.65799			
24	2	basic_sector_employment_within_walking_distance	0.72067	0.33415	2.157	0.0310
24	2	is_near_arterial	1.56713	0.28899	5.423	0.0000
24	2	ln_distance_to_highway	-0.06137	0.02367	-2.593	0.0095
24	2	ln_residential_units	9.02859	0.80222	11.254	0.0000
24	2	ln_work_access_to_employment	-0.34462	0.01749	-19.704	0.0000
24	2	n_recent_transitions_to_governmental_within_walking_dista nce	37.68332	3.06149	12.309	0.0000
24	2	percent_high_income_households_within_walking_distance	0.05992	0.00369	16.252	0.0000
24	2	percent_low_income_households_within_walking_distance	0.09507	0.02812	3.381	0.0007
24	2	percent_mid_income_households_within_walking_distance	0.05573	0.00382	14.606	0.0000
24	2	percent_same_type_cells_within_walking_distance	-0.01705	0.00235	-7.251	0.0000
24	2	percent_water	0.02894	0.00647	4.475	0.0000
24	2	retail_sector_employment_within_walking_distance	-2.93730	0.87391	-3.361	0.0008
24	2	service_sector_employment_within_walking_distance	0.04658	0.01891	2.463	0.0138
24	17	alternative specific constant	-0.05739			
24	17	basic_sector_employment_within_walking_distance	1.13335	0.33396	3.394	0.0007
24	17	is_near_arterial	1.39304	0.08342	16.700	0.0000
24	17	ln_distance_to_highway	0.07012	0.02158	3.249	0.0012
24	17	ln_work_access_to_employment	-0.01662	0.00591	-2.812	0.0049
24	17	n_recent_transitions_to_industrial_within_walking_distance	183.30211	0.77387	236.865	0.0000
24	17	n_recent_transitions_to_governmental_within_walking_dista nce	32.56349	3.06299	10.631	0.0000
24	17	percent_mid_income_households_within_walking_distance	-0.01689	0.00972	-1.737	0.0823
24	17	percent_same_type_cells_within_walking_distance	-0.01801	0.00223	-8.091	0.0000
24	17	retail_sector_employment_within_walking_distance	0.15626	0.05156	3.031	0.0024
24	17	service_sector_employment_within_walking_distance	0.04960	0.01846	2.687	0.0072

24	18	alternative specific constant	-2.16944			
24	18	basic_sector_employment_within_walking_distance	1.13335	0.33396	3.394	0.0007
24	18	is_near_arterial	3.19669	0.51770	6.175	0.0000
24	18	ln_distance_to_highway	-0.29280	0.06150	-4.761	0.0000
24	18	ln_work_access_to_employment	-0.01662	0.00591	-2.812	0.0049
24	18	n_recent_transitions_to_industrial_within_walking_distance	183.30211	0.77387	236.865	0.0000
24	18	n_recent_transitions_to_governmental_within_walking_dista nce	35.86679	3.06457	11.704	0.0000
24	18	percent_mid_income_households_within_walking_distance	-0.01689	0.00972	-1.737	0.0823
24	18	percent_same_type_cells_within_walking_distance	-0.04356	0.00593	-7.347	0.0000
24	18	retail_sector_employment_within_walking_distance	0.17924	0.05160	3.474	0.0005
24	18	service_sector_employment_within_walking_distance	0.04960	0.01846	2.687	0.0072
24	20	alternative specific constant	21.54908			
24	20	basic_sector_employment_within_walking_distance	1.06057	0.33415	3.174	0.0015
24	20	ln_distance_to_highway	-0.12388	0.03701	-3.347	0.0008
24	20	ln_residential_units	9.90515	0.79795	12.413	0.0000
24	20	ln_work_access_to_employment	-0.60960	0.06123	-9.956	0.0000
24	20	n_recent_transitions_to_governmental_within_walking_dista nce	115.90603	3.06463	37.821	0.0000
24	20	percent_high_income_households_within_walking_distance	0.53954	0.00404	13.346	0.0000
24	20	percent_low_income_households_within_walking_distance	0.07384	0.02964	2.491	0.0127
24	20	percent_mid_income_households_within_walking_distance	0.03826	0.00458	8.349	0.0000
24	20	percent_same_type_cells_within_walking_distance	-0.94858	0.00831	-114.12	0.0000
24	20	percent_water	0.02045	0.01191	1.718	0.0858
24	20	service_sector_employment_within_walking_distance	0.06137	0.01910	3.214	0.0013
24	23	alternative specific constant	50.00000			
24	23	basic_sector_employment_within_walking_distance	2.57255	0.33411	7.700	0.0000
24	23	is_near_arterial	-33.98544	0.34378	-98.858	0.0000
24	23	ln_distance_to_highway	-0.12388	0.03701	-3.347	0.0008
24	23	ln_residential_units	9.90515	0.79795	12.413	0.0000
24	23	ln_work_access_to_employment	-0.60960	0.06123	-9.956	0.0000
24	23	n_recent_transitions_to_governmental_within_walking_dista	115.90603	3.06463	37.821	0.0000

		nce				
24	23	percent_high_income_households_within_walking_distance	0.53954	0.00404	13.346	0.0000
24	23	percent_low_income_households_within_walking_distance	0.07384	0.02964	2.491	0.0127
24	23	percent_mid_income_households_within_walking_distance	0.03826	0.00458	8.349	0.0000
24	23	percent_same_type_cells_within_walking_distance	-0.01656	0.00772	-2.144	0.0320
24	23	percent_water	0.02045	0.01191	1.718	0.0858
24	23	retail_sector_employment_within_walking_distance	-3.43144	0.12536	-27.374	0.0000
24	23	service_sector_employment_within_walking_distance	0.06137	0.01910	3.214	0.0013
24	24	alternative specific constant	50.00000			
24	24	basic_sector_employment_within_walking_distance	0.79959	0.33431	2.392	0.0168
24	24	is_near_arterial	0.64404	0.17056	3.776	0.0002
24	24	ln_distance_to_highway	-0.12388	0.03701	-3.347	0.0008
24	24	ln_residential_units	9.90515	0.79795	12.413	0.0000
24	24	ln_work_access_to_employment	-0.60960	0.06123	-9.956	0.0000
24	24	n_recent_transitions_to_governmental_within_walking_dista nce	115.90603	3.06463	37.821	0.0000
24	24	percent_high_income_households_within_walking_distance	0.53954	0.00404	13.346	0.0000
24	24	percent_low_income_households_within_walking_distance	0.07384	0.02964	2.491	0.0127
24	24	percent_mid_income_households_within_walking_distance	0.03826	0.00458	8.349	0.0000
24	24	percent_same_type_cells_within_walking_distance	0.04372	0.00698	6.266	0.0000
24	24	percent_water	0.02045	0.01191	1.718	0.0858
24	24	service_sector_employment_within_walking_distance	0.06137	0.01910	3.214	0.0013
25	-1	alternative specific constant	0			
25	25	alternative specific constant	50.00000			

5.2.5 Estimation of the Residential Land Share Model

The Residential Land Share Model is a linear regression model that predicts the fraction of land in a grid cell that will be residential. The dependent variable is the natural log of the ratio of residential land to nonresidential land in a grid cell, and the independent variables are residential units, nonresidential floor space, and development types (e.g., devtype1, devtype2, devtype3, and so on) in the grid cell. The variables for the development types are binary. A detailed description of variables is presented in Appendix IV. The developed model has an adjusted R^2 of 0.66. Table 5.7 lists the estimated coefficients with their descriptive statistics.

Table 5.7 Residential Land Share Model Coefficients

Coefficient Name	Estimate	Standard Error	T-Statistic	P-Value
constant	0.01324	0.00069471	19.06	0.0000
residential_units	0.00224	0.00016178	13.87	0.0000
non_residential_sqft	-0.00000117	0.0000001424592	-8.21	0.0000
devtype1	0.54729	0.00207	263.87	0.0000
devtype2	0.54333	0.00230	236.59	0.0000
devtype3	0.54035	0.00273	198.00	0.0000
devtype4	0.60787	0.00334	181.95	0.0000
devtype5	0.65035	0.00486	133.77	0.0000
devtype6	0.64904	0.00654	99.17	0.0000
devtype7	0.57193	0.00915	62.51	0.0000
devtype8	0.24169	0.02548	9.49	0.0000
devtype9	0.31842	0.00417	76.34	0.0000
devtype10	0.22827	0.00533	42.86	0.0000
devtype11	0.14952	0.01126	13.28	0.0000
devtype12	0.49731	0.00701	70.94	0.0000
devtype13	0.37061	0.00743	49.90	0.0000
devtype14	0.18605	0.01810	10.28	0.0000
devtype15	0.30228	0.01477	20.46	0.0000
devtype17	0.04395	0.00205	21.48	0.0000
devtype18	0.04439	0.00940	4.72	0.0000
devtype19	0.14745	0.03865	3.82	0.0001
devtype20	0.04414	0.00845	5.22	0.0000
devtype21	0.04179	0.01600	2.61	0.0090
devtype22	0.10945	0.10193	1.07	0.2830
devtype23	0.03870	0.00499	7.76	0.0000
devtype24	0.01764	0.00087176	20.23	0.0000

6. INTEGRATION OF URBANSIM WITH TRAVEL MODEL

UrbanSim and the FSUTMS model are executed alternately with data exchanges between the two programs during each iteration. UrbanSim provides socioeconomic data for the FSUTMS input files (ZDATA1 and ZDATA2), and results from the FSUTMS are fed back to UrbanSim in the form of composite travel utilities. Significant changes to the transportation supply, such as new or modified facilities or cumulative congestion effects caused by the growth and spatial distribution of jobs and population are reflected in the accessibility measures obtained based on the output from the FSUTMS model. Figure 6.1 shows the data flow between UrbanSim and the FSUTMS model.

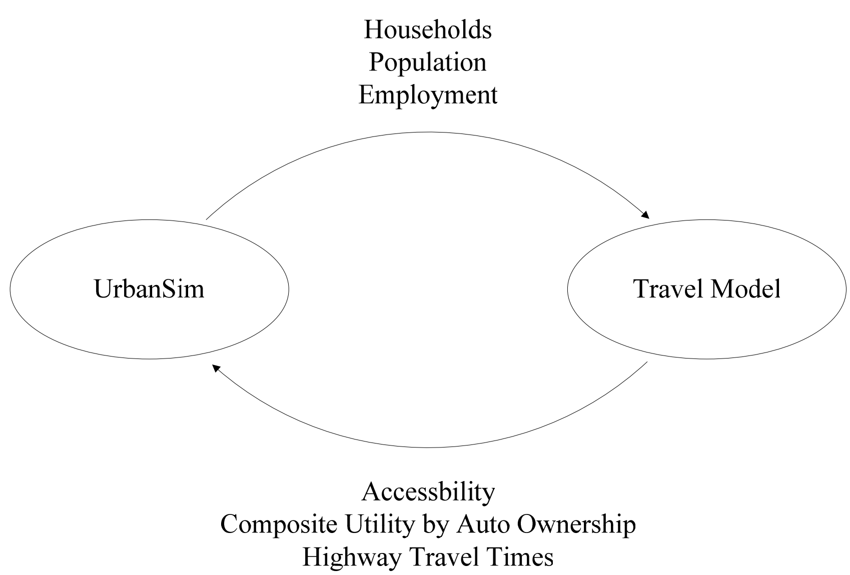


Figure 6.1 Data Flow between UrbanSim and FSUTMS

The simulation steps are listed below:

1. Run UrbanSim for a chosen time interval
   a. Create a base year database (e.g. “Volusia_Baseyear”).
   b. Create a scenario database for a given scenario (e.g., “Volusia_Scenario1_1”).
   c. Run UrbanSim.
2. Transfer data from UrbanSim to FSUTMS by creating ZDATA1 and ZDATA2 from the output database (e.g. “Volusia_Output1”).
3. Run the FSUTMS model
4. Transfer data from the FSUTMS model to UrbanSim
   a. Create a new scenario database (e.g., “Volusia_Scenario1_2”).
   b. Create tables of households, jobs, and grid cells.
5. Repeat steps 1 through 4 until the entire simulation is complete with the forecast year reached.

The time interval for simulation is based on the consideration of the need to (1) allow adequate interactions between the transportation and land use systems, (2) simulate the gradual changes in land use, and (3) limit the total simulation time required because the UrbanSim simulation is computationally intensive. UrbanSim simulates land use changes on an annual basis. It is interacted with the travel model based on intervals ranging between two to four years and defined by time points of the years 2000 (the base year), 2002, 2005, 2007, 2010, 2013, 2017, and 2020. The FSUTMS model is run at these time points, as shown in Figure 6.2.

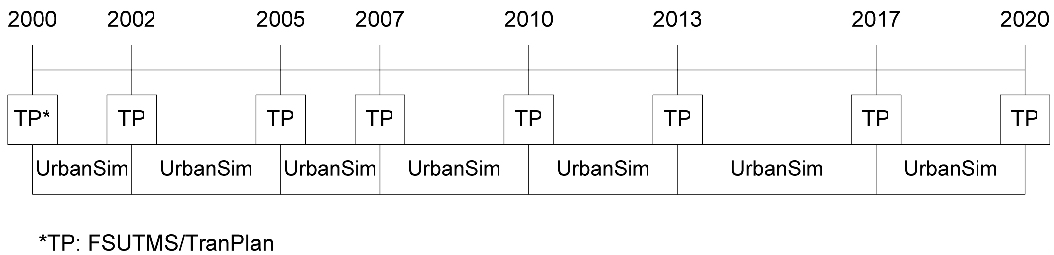


Figure 6.2 Interface Design of UrbanSim and FSUTMS

7. MODEL VALIDATION

Model validation is conducted for scenario 1, which assumes the med-range projection of countywide socioeconomic and demographic changes. For the validation purpose, UrbanSim and FSUTMS are run until the forecast year of 2020 using the first scenario, which is based on the adopted LRTP. Results are compared with two data sources: socioeconomic and demographic data adopted in the 2020 LRTP and the 2005 InfoUSA employment data. A comparison with socioeconomic and demographic data is made for the years 2000 (Base Year), 2010, and 2020, while the 2005 InfoUSA employment data are compared with the simulation results from UrbanSim for 2005.

7.1 Comparison with the Adopted Data in the LRTP

The household and employment forecasts for the years 2000, 2010, and 2020 are extracted from the UrbanSim base year database and its output databases, and are compared with the socioeconomic and demographic forecasts adopted in the 2020 LRTP. Since the control totals for Scenario 1 are the same as those adopted in the 2020 LRTP, the total number of households and employment from UrbanSim are very close to those from the socioeconomic and demographic data for the 2020 LRTP.

The cumulative percentage of TAZs is graphed against the household and population differences between UrbanSim forecasts and the 2020 LRTP (= UrbanSim – the 2020 LRTP) in Figure 7.1. The cumulative percentage for the year 2000's households (in orange color) shows that errors at the zonal level range from -100 to 100 percent for over 80 percent of the TAZs. The TAZ distributions that are based on the errors for 2010 and 2020 are similar to each other, with the 2010 distribution indicating slightly better results than the 2020 distribution. As the simulation time moves further from the base year, the difference in households between the UrbanSim output and the 2020 LRTP grows larger. This trend is also found in the zonal population obtained from UrbanSim and the LRTP. Compared with the 2010 and 2020 zonal populations, the 2000 zonal populations from UrbanSim and the 2020 LRTP have the least discrepancies.

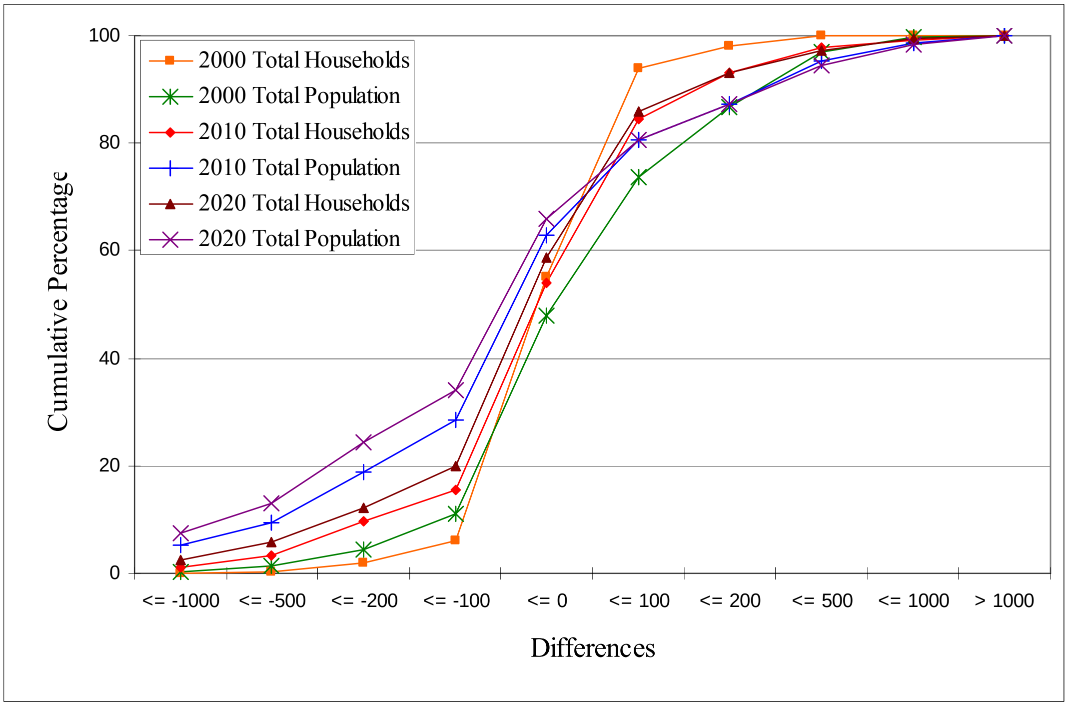


Figure 7.1 Cumulative Percentage of TAZs versus Differences in Zonal Households and Population

Figure 7.2 through 7.4 plot the distributions of TAZs as the cumulative percentage versus employment differences between the UrbanSim outputs for the base year, 2010 and 2020, respectively, and those from the 2020 LRTP (= UrbanSim – the 2020 LRTP). Zonal employment is shown in terms of the total, as well as industrial, commercial, and service employments. For the base year, industrial employment shows the least discrepancies, and over 60 percent of the TAZs show a difference between -100 and 100. The cumulative percentage of the TAZs whose differences in the employment area range between -100 and 100 is approximately 60 percent in 2010 and 50 percent in 2020. As the simulation year progresses, the percentage of TAZs with differences in this range (–100 to +100) decreases.

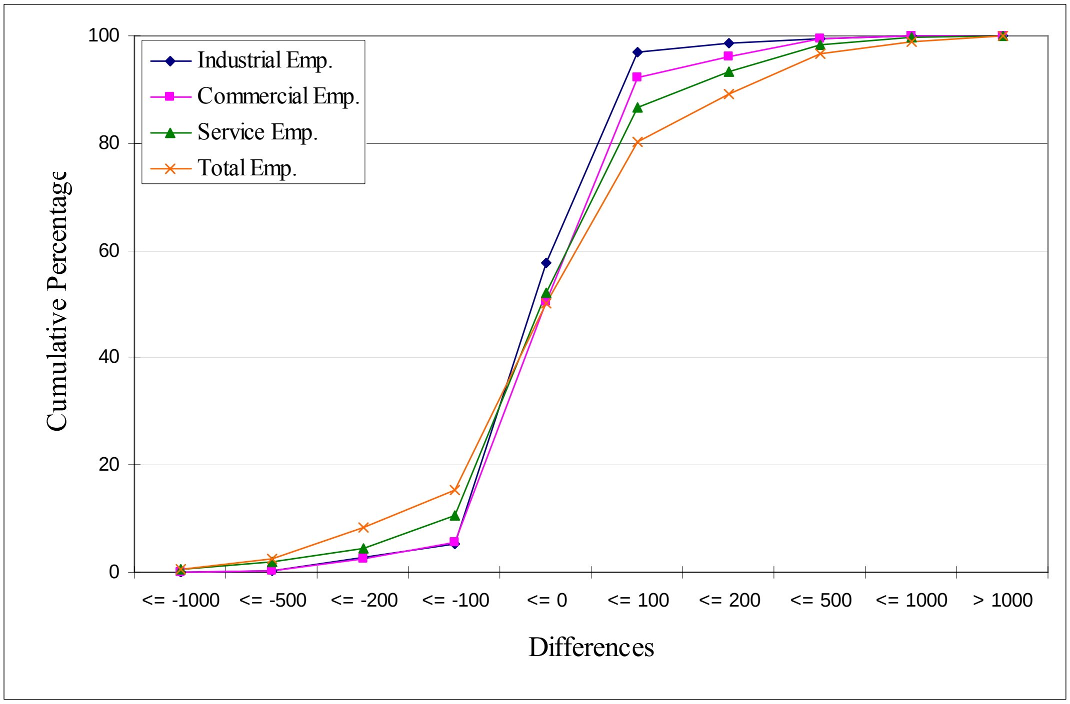


Figure 7.2 Cumulative Percentage of TAZs versus Differences in Employment in Base Year

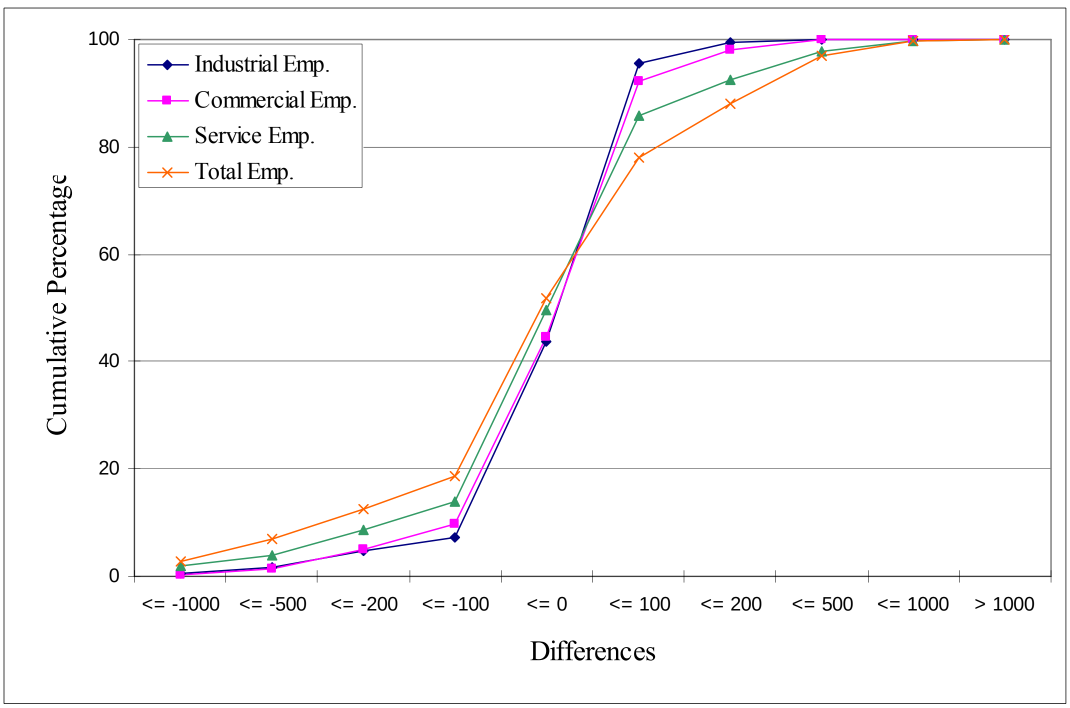


Figure 7.3 Cumulative Percentage of TAZs versus 2010 Employment Difference

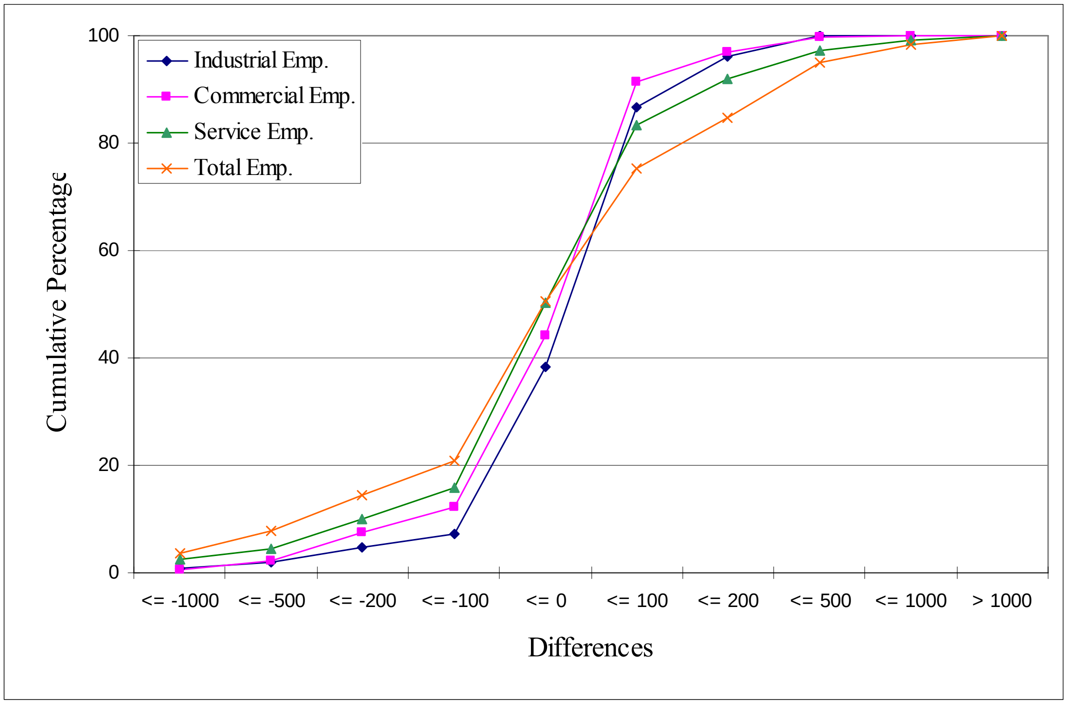


Figure 7.4 Cumulative Percentage of TAZs versus 2020 Employment Difference

7.2 Comparison to 2005 InfoUSA Employment Data

In addition to comparing the results from UrbanSim with the socioeconomic and demographic data adopted in the 2020 LRTP, the 2005 employment data from the UrbanSim output is compared with the 2005 InfoUSA employment data.

The employment is summarized at the TAZ level and grouped into three types of employment based on the SIC: industrial employment, commercial employment, and service employment. The total employment is the sum of the employment of the three types. Figure 7.5 shows the cumulative percentage of differences (= UrbanSim output – the 2005 InfoUSA employment) in the zonal employment for each type. The negative differences indicate that employment was under-estimated by UrbanSim, whereas the positive differences point to over-estimation. Among the three types of employment, industrial employment shows the least discrepancies between the UrbanSim output and the 2005 InfoUSA employment. Over 90 percent of all the TAZs have differences in industrial employment between -100 and 100. About 70-80 percent of all the TAZs have a difference in the same range for commercial and service employment, respectively. The total employment shows that 60 percent of the total TAZs lie in the same range of differences (-100 to 100). Given the known problems in an unverified and unvalidated InfoUSA database, these statistics appear to be reasonable. In other words, UrbanSim appears to have performed reasonably. It is also worthwhile to point out that the business establishments in the 2005 InfoUSA database are geocoded on the street system without the necessary offsets to accurately locate them to the actual TAZs, which likely has introduced spatial errors into the aggregated employment at the TAZ level.

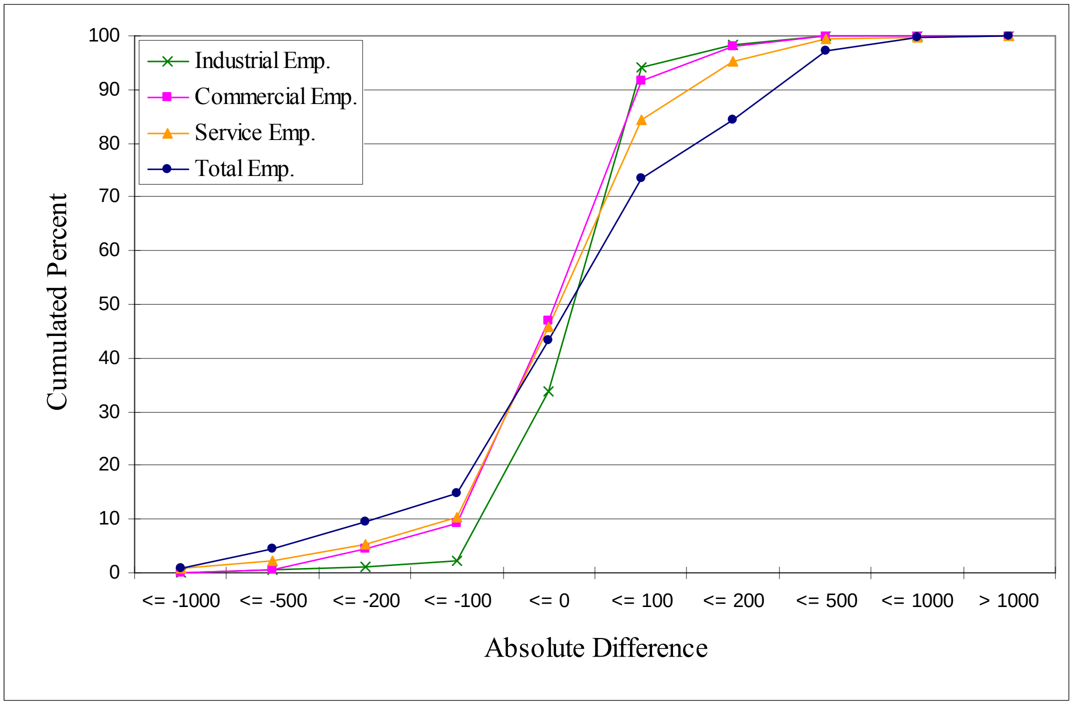


Figure 7.5 Cumulative Percentage of Employment Differences between UrbanSim and the 2005 InfoUSA Data

Figure 7.6 illustrates the spatial distribution of differences in the employment size from UrbanSim and the InfoUSA database at the TAZ level. The planning region boundaries are also shown on the map. The negative differences indicating under-estimation by UrbanSim are displayed in the blue color and the positive differences indicating over-estimation are in orange color. Most of the TAZs with large discrepancies are located in the northeast region, and some TAZs with large discrepancies are also found in the central west and southwest regions.

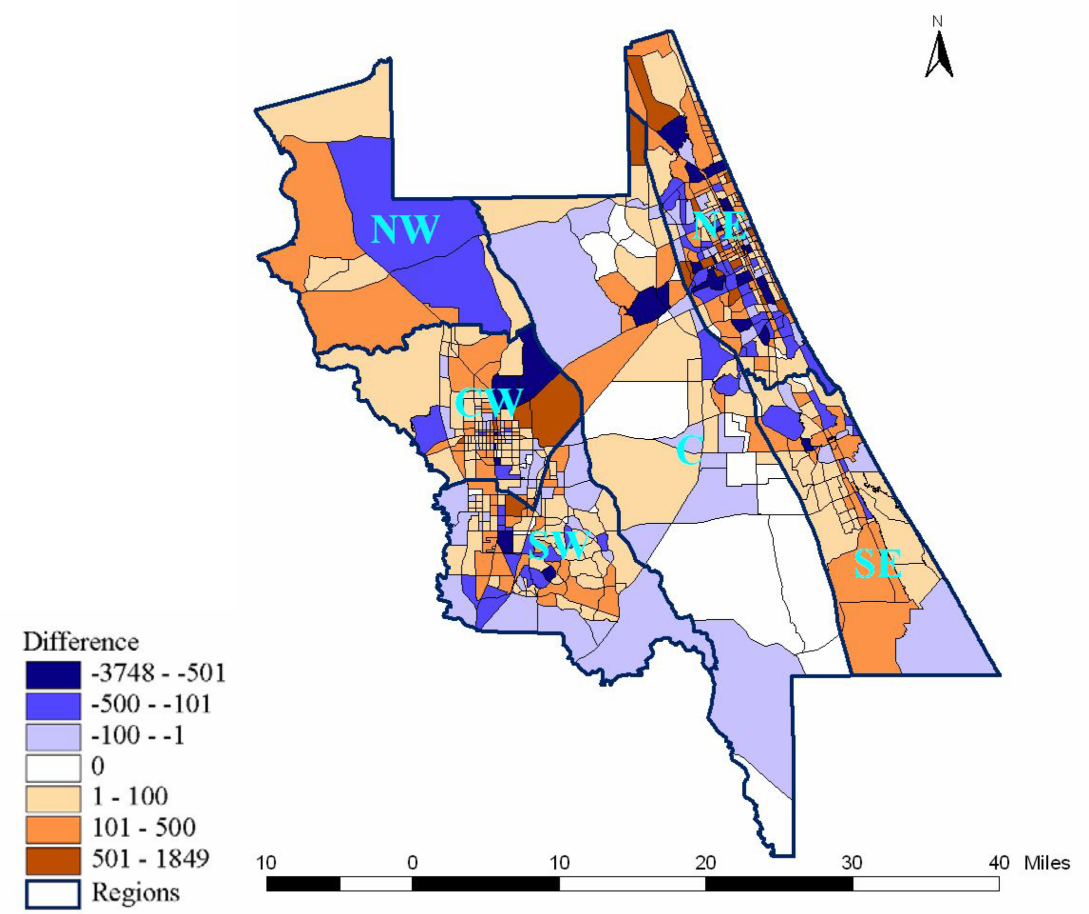


Figure 7.6 Spatial Distribution of Differences between UrbanSim and the 2005 InfoUSA Data

8. MODEL IMPLEMENTATION

The UrbanSim simulations for five scenarios have been successfully performed. The simulation process has been described in Chapter 6, with the interaction between UrbanSim and the FSUTMS model illustrated in Figure 6.2. The simulation results are presented in this chapter. The scenarios examined are the following:

- Scenario 1: The final plan in the 2020 LRTP with mid-range projection.
- Scenario 2: The alternative 2 in the 2020 LRTP with mid-range projection.
- Scenario 3: The alternative 3 in the 2020 LRTP with mid-range projection.
- Scenario 4: The final plan in the 2020 LRTP with low-range projection.
- Scenario 5: The final plan in the 2020 LRTP with high-range projection.

8.1 Scenario 1 (Final Transportation Plan and Mid-Range Projection)

Scenario 1 uses the final transportation improvement plan in the 2020 LRTP, as shown in Figure 8.1, and the mid-range projection of socioeconomic and demographic data adopted in the 2020 LRTP. Note that the countywide socioeconomic and demographic statistics are quite close between the UrbanSim input data and the socioeconomic and demographic data adopted in the 2020 LRTP, since both are use the mid-projection criteria. The interaction of UrbanSim and the FSUTMS take place for the years of 2002, 2005, 2007, 2010, 2013, 2017, and 2020.

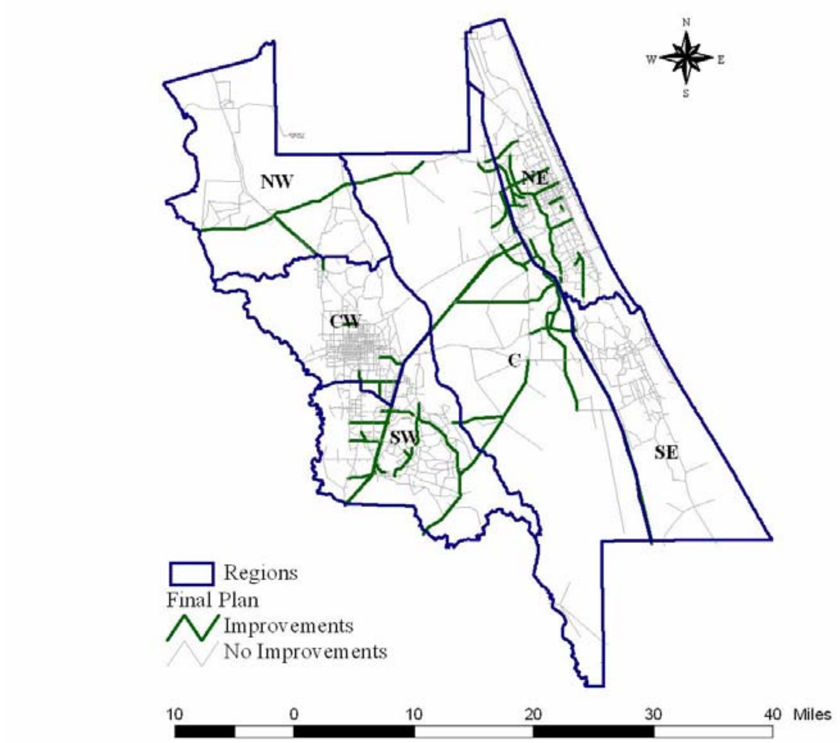


Figure 8.1 Final Plan in the 2020 LRTP

Figure 8.2 presents the results from the FSUTMS model for the forecast year 2020. The V/C ratio, defined as volume over capacity, for the 2020 LRTP and scenario 1 is shown in Figure 8.2 (a) and (b), respectively. Figure 8.2 (c) displays the volume difference in each link between the 2020 LRTP and scenario 1 results. In the northeast region, the volumes from scenario 1 are underestimated when compared to the volumes from the 2020 LRTP, whereas overestimation occurs in the central-west region.

Figures 8.3 through 8.6 illustrate the results of UrbanSim in terms of accessibility, population, households, and employment, respectively. Accessibility is defined as:

$$Accessibility_i = \sum_{j=1}^{J} (job_j \times \exp(LogSum_{ij})),$$

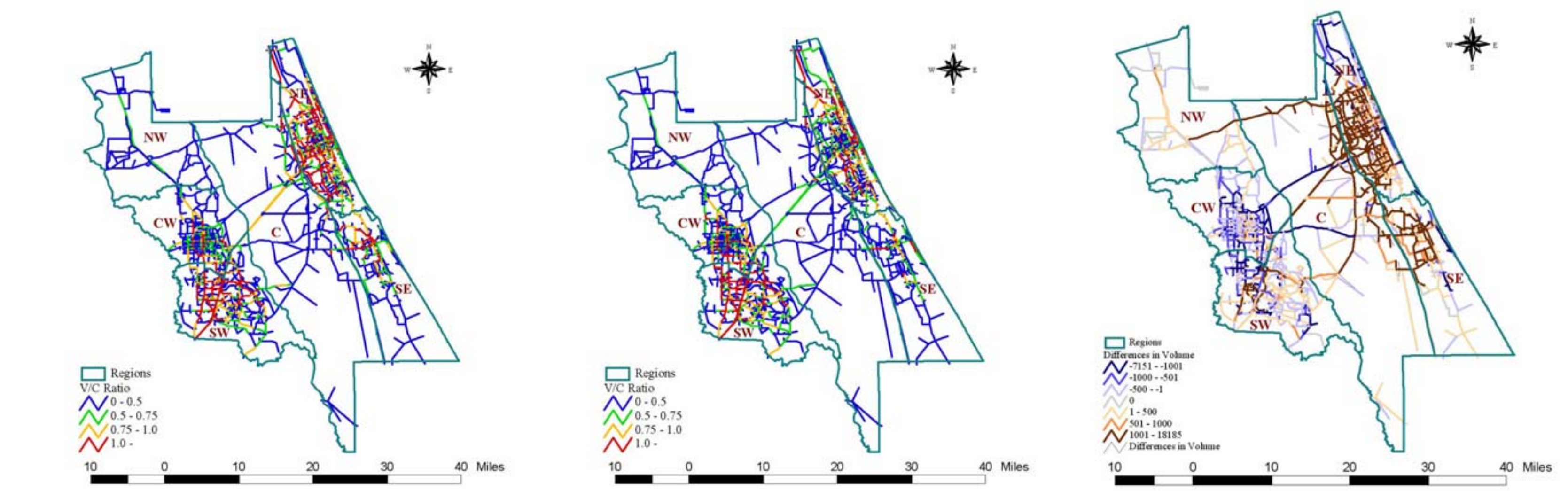
where job_j = employment in TAZ j, and
$LogSum_{ij}$ = logsum for one-vehicle households from TAZ i to TAZ j.

The accessibilities in 2010 and 2020 are relatively higher in the northeast region, especially around the City of Daytona Beach, than any of the other regions, as shown in Figure 8.3 (a) and (b), respectively. Figure 8.3 (c), however, indicates a higher increase in accessibility along the I-4 in the central, central-west, and southwest regions.

The population distributions in 2010 and in 2020 are shown in Figure 8.4 (a) and (b), respectively. The population changes between the base year to the year 2020 are displayed in Figure 8.4 (c). A large population growth is found near the urban growth boundaries, and a large population decline is evident in the southwest region.

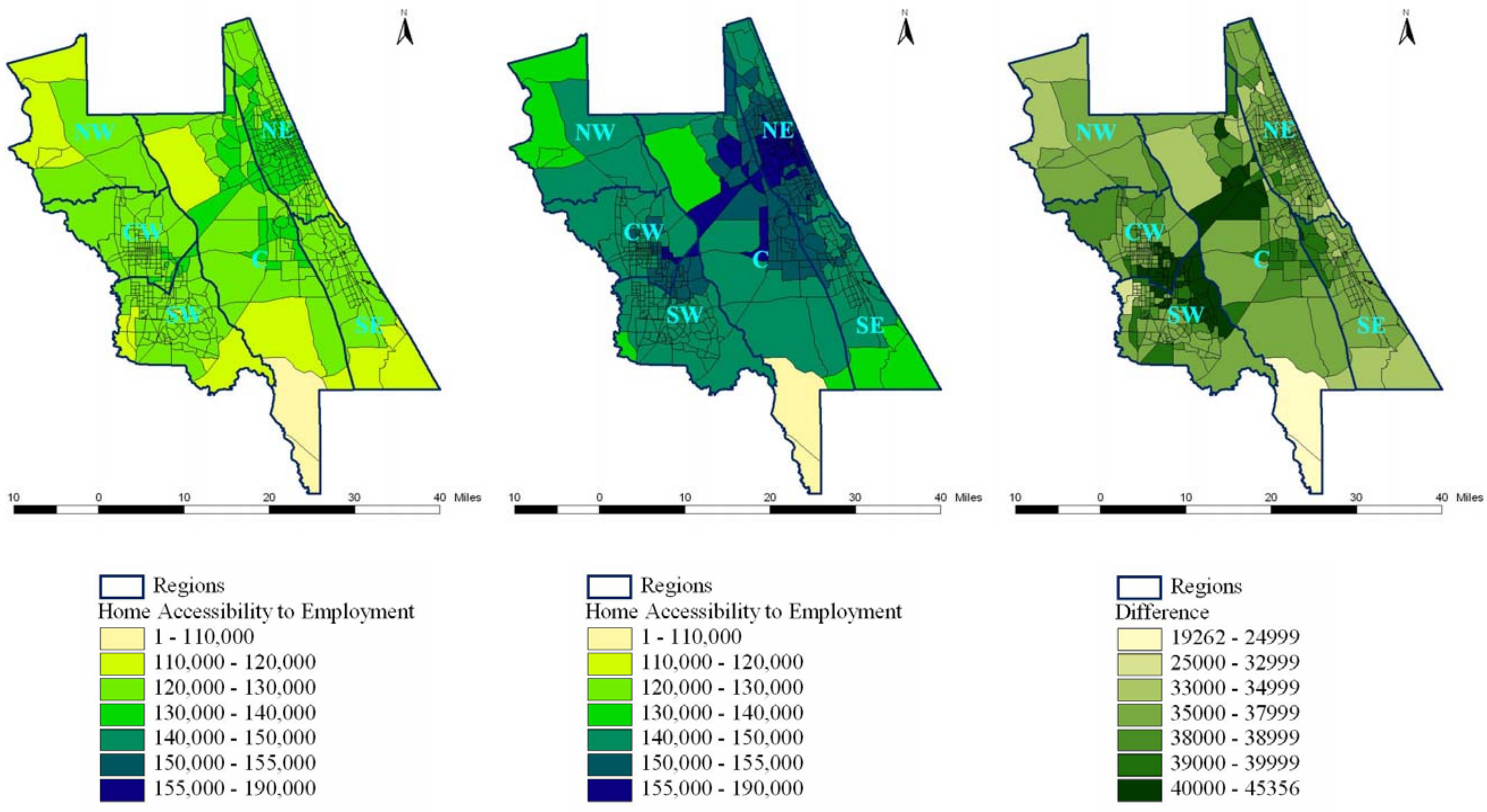
The household distribution in 2010 is similar to that in 2020, as shown in Figures 8.5 (a) and (b), respectively. The household changes between the base year and the year 2020, illustrated in Figure 8.5 (c), show positive growth in almost every TAZ. It is interesting because several TAZs in the southwest region would experience a large population loss, as shown in Figure 8.4 (c), but still gain households. One possible interpretation is that the residential developments in these TAZs are mostly multi-family homes, for which household size is smaller than single-family homes. A further investigation is necessary to interpret address this result.

Employment growth is found around the urban growth boundaries and in the southwest region, as shown in Figure 8.6. Tables 8.1 and 8.2 compare the number of households and employment, respectively, from UrbanSim with those for years of 1997, 2010, and 2020 based on the 2020 LRTP.



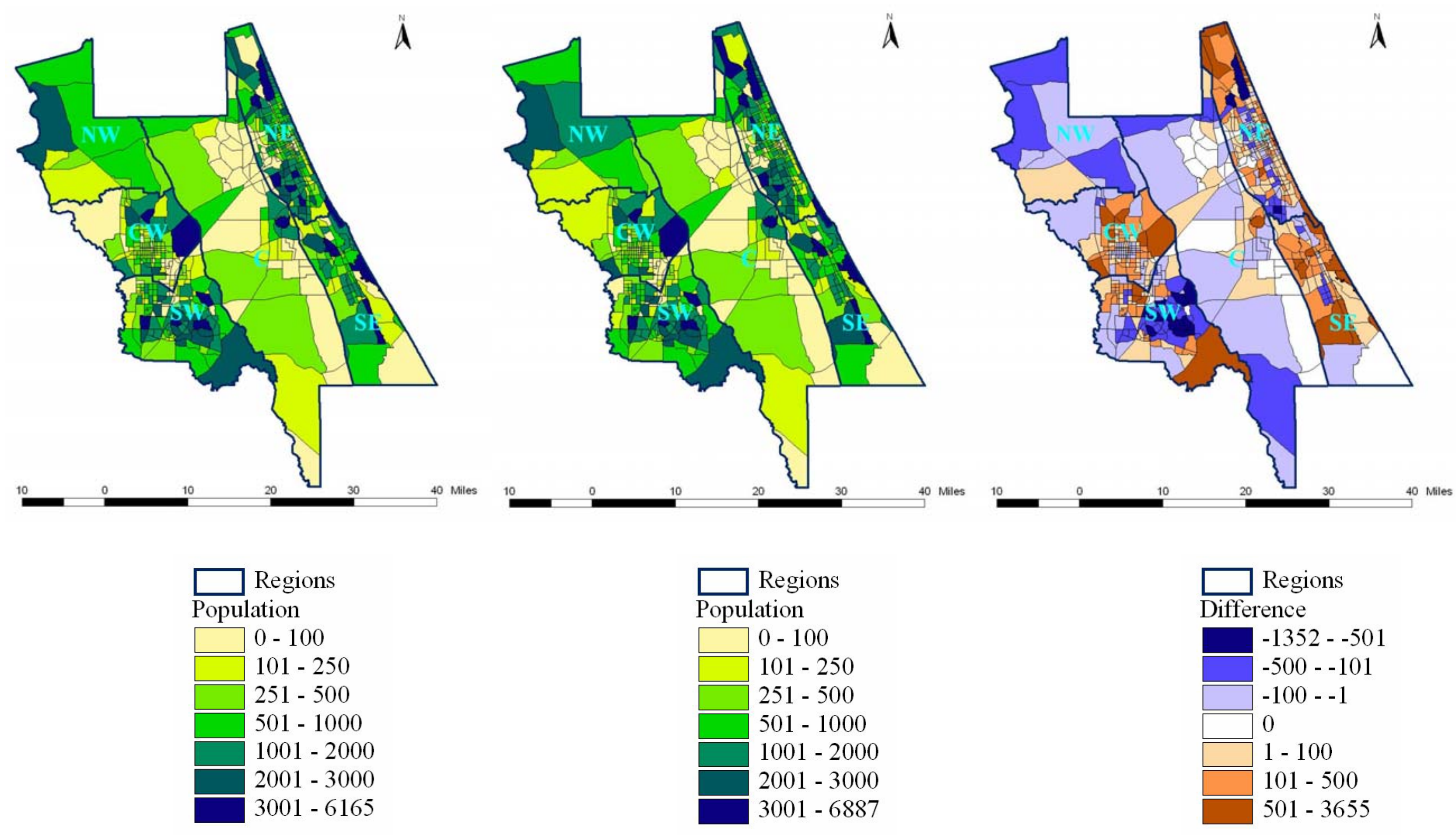
(a) 2020 LRTP (b) Scenario 1 (c) LRTP – Scenario 1

Figure 8.2 Comparison of V/C Ratios in 2020 between the LRTP and Scenario 1



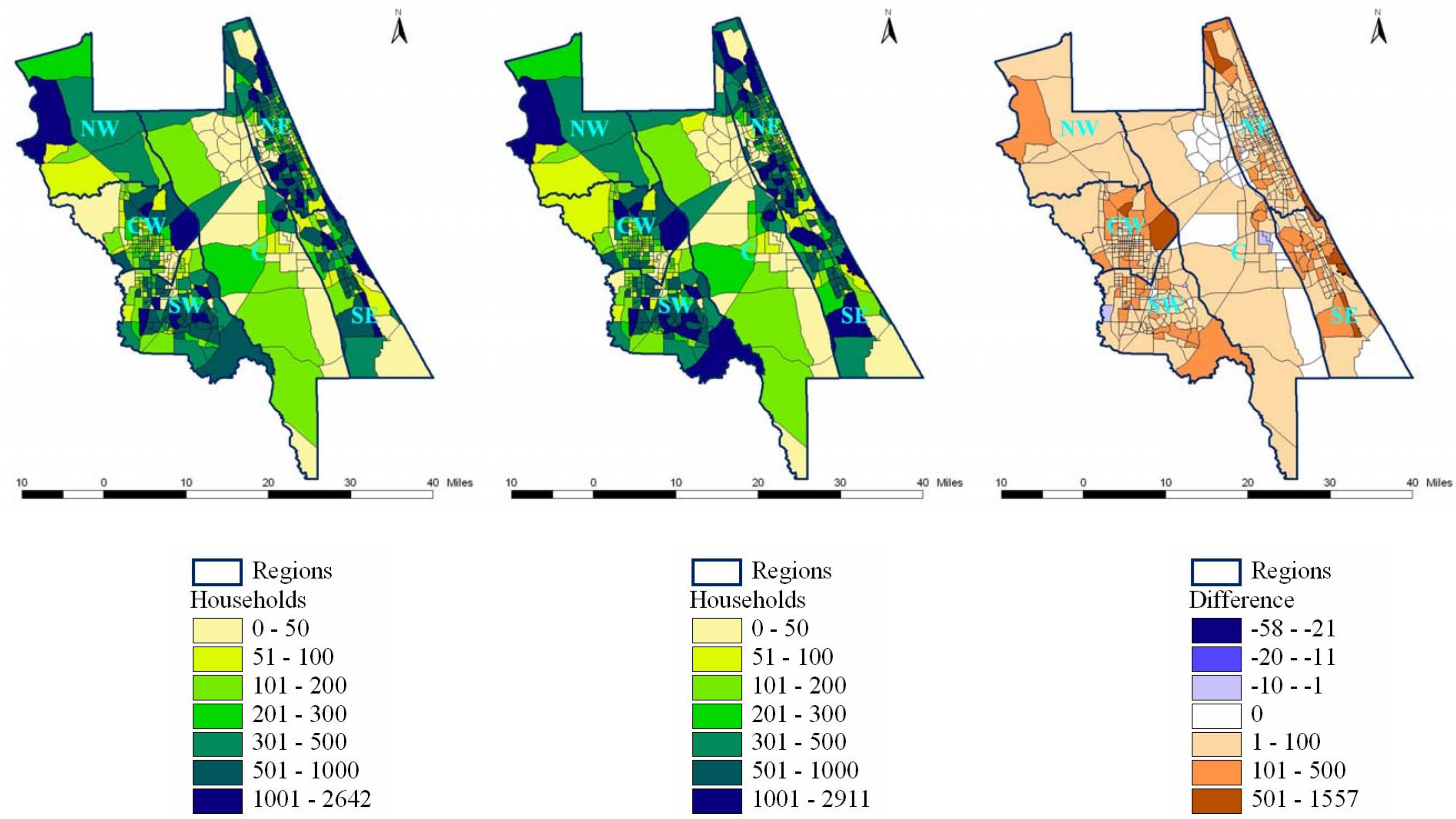
(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.3 Home Accessibility for One-Vehicle Households to Employment from Scenario 1



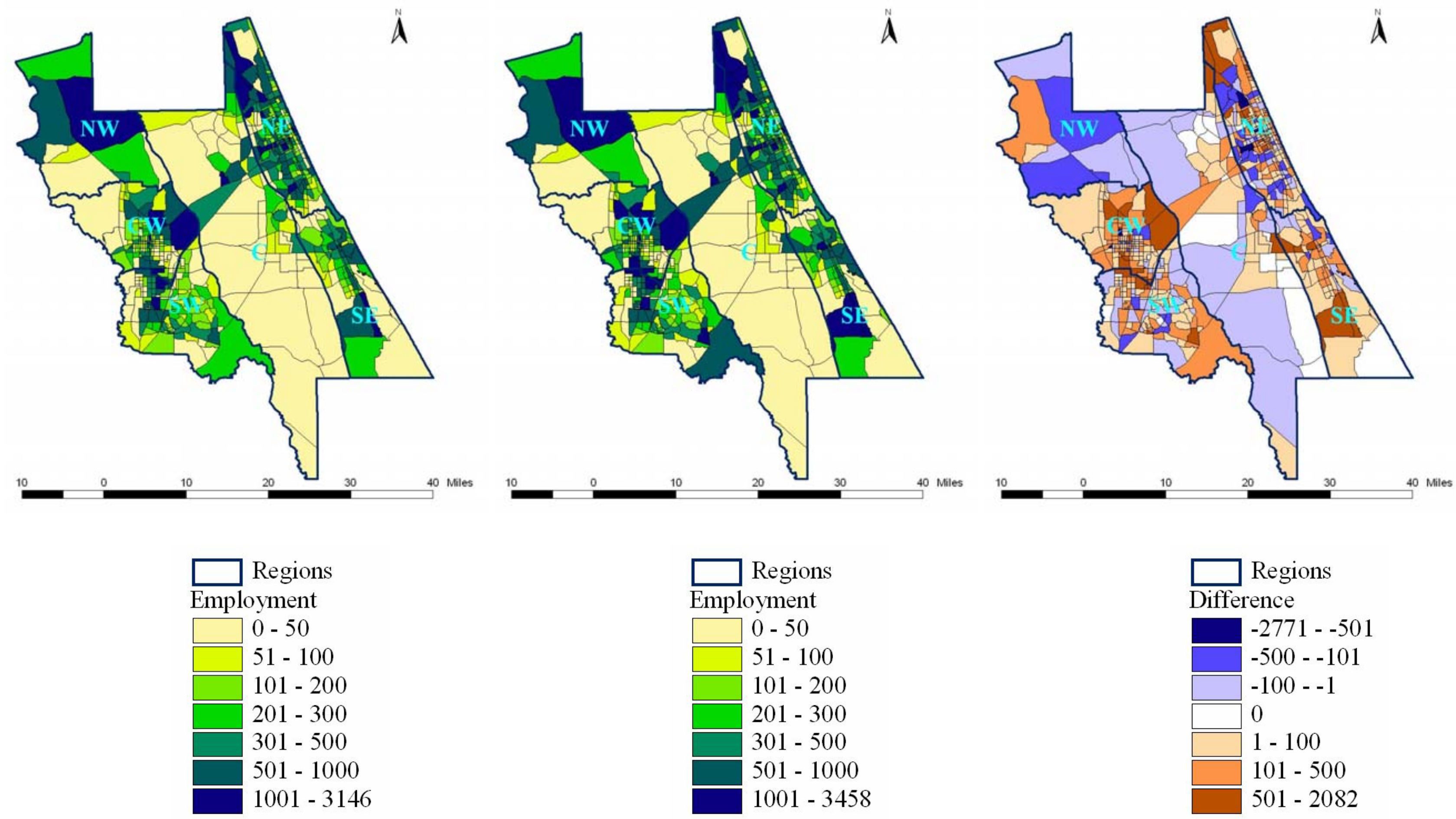
(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.4 Spatial Distribution of Population from Scenario 1



(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.5 Spatial Distribution of Households from Scenario 1



(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.6 Spatial Distribution of Employment from Scenario 1

Table 8.1 Comparison of Households between the 2020 LRTP and Scenario1

Planning Region	1997	2000	2010		2020	
	LRTP	Scenario 1	LRTP	Scenario 1	LRTP	Scenario 1
Northeast	84,022	80,899	106,863	96,383	119,592	97,987
Southeast	23,846	22,116	31,880	30,424	35,437	32,410
Central	5,689	6,754	18,312	8,087	29,739	8,505
Northwest	2,288	2,176	2,855	2,741	3,255	2,842
Central-west	19,178	19,145	24,590	25,992	27,705	28,937
Southwest	36,300	40,492	40,950	47,198	44,648	49,367

Table 8.2 Comparison of Employment between the 2020 LRTP and Scenario1

Planning Region	1997	2000	2010		2020	
	LRTP	Scenario 1	LRTP	Scenario 1	LRTP	Scenario 1
Northeast	100,648	101,315	130,583	106,643	130,583	118,548
Southeast	16,220	15,103	28,919	19,039	28,919	23,803
Central	2,626	6,778	9,761	8,369	9,761	10,565
Northwest	3,789	3,602	3,809	2,927	3,809	3,055
Central-west	20,434	25,440	30,587	36,443	30,587	46,432
Southwest	17,340	20,591	24,984	26,049	24,984	31,969

8.2 Scenario 2 (Transportation Plan Alternative 2 and Mid-Rang Projection)

Scenario 2 combines the second alternative transportation improvement plan in the 2020 LRTP with the mid-range projection of the socioeconomic and demographic data adopted in the 2020 LRTP. Figure 8.7 shows the second alternative transportation improvement plan and compares it to the final plan. Improvements in blue represent common improvements for both plans. The yellow and red colors indicate projects in the second alternative plan and the final plan, respectively. The difference between scenarios 1 and 2 is the transportation improvement plan.

The V/C ratios for scenarios 1 and 2 are displayed in Figures 8.8 (a) and (b), respectively, and the volume differences (scenario 2 – scenario 1) between the two scenarios are presented in Figure 8.8 (c). There is no obvious pattern of increased or decreased congestion. Analyses at the site level are required to identify the impact from different transportation improvements.

Accessibility, population, households, and employment from scenario 2 are displayed in Figures 8.9 through 8.13, respectively. Similar to the results from scenario 1, accessibilities in 2010 and 2020 are relatively better in the northeast region than in other regions, as shown in Figures 8.9 (a) and (b), respectively. A higher increase in accessibility is found in some TAZs in the central-west and southwest regions. Comparisons of scenarios 1 and 2 are made in terms of accessibility at the TAZ level in 2010 and 2020, as shown in Figure 8.10. In 2010, scenario 1 has some TAZs with higher or lower accessibility than scenario 2. The TAZs with higher accessibility are found in the central, central-west, and southwest regions, while TAZs with lower accessibility spread out in the northwest and northeast regions. In 2020, there are many TAZs that have lower accessibility in scenario 1 than in scenario 2.

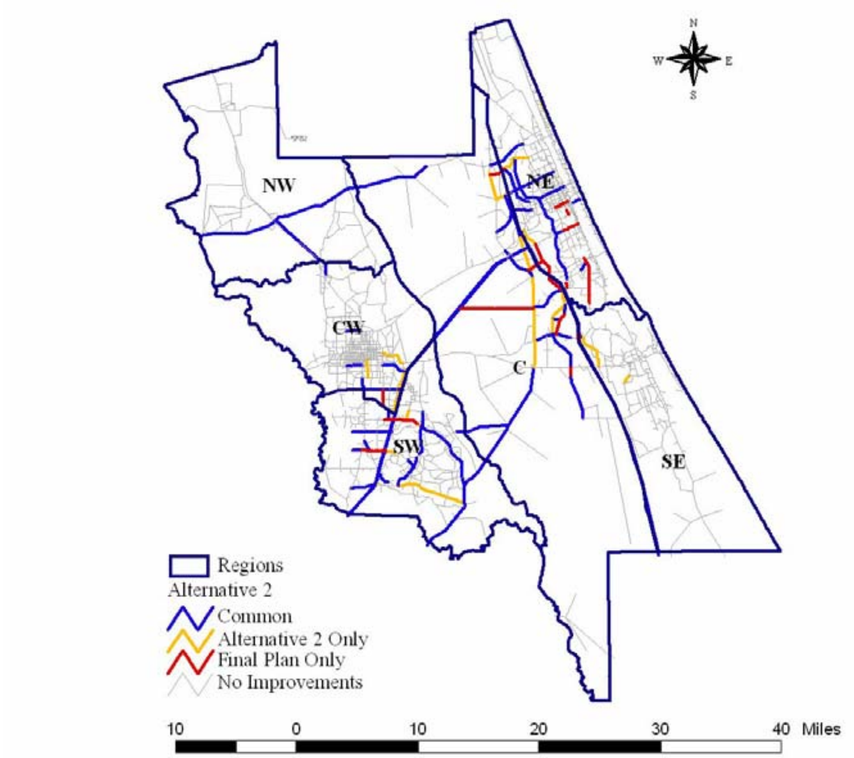
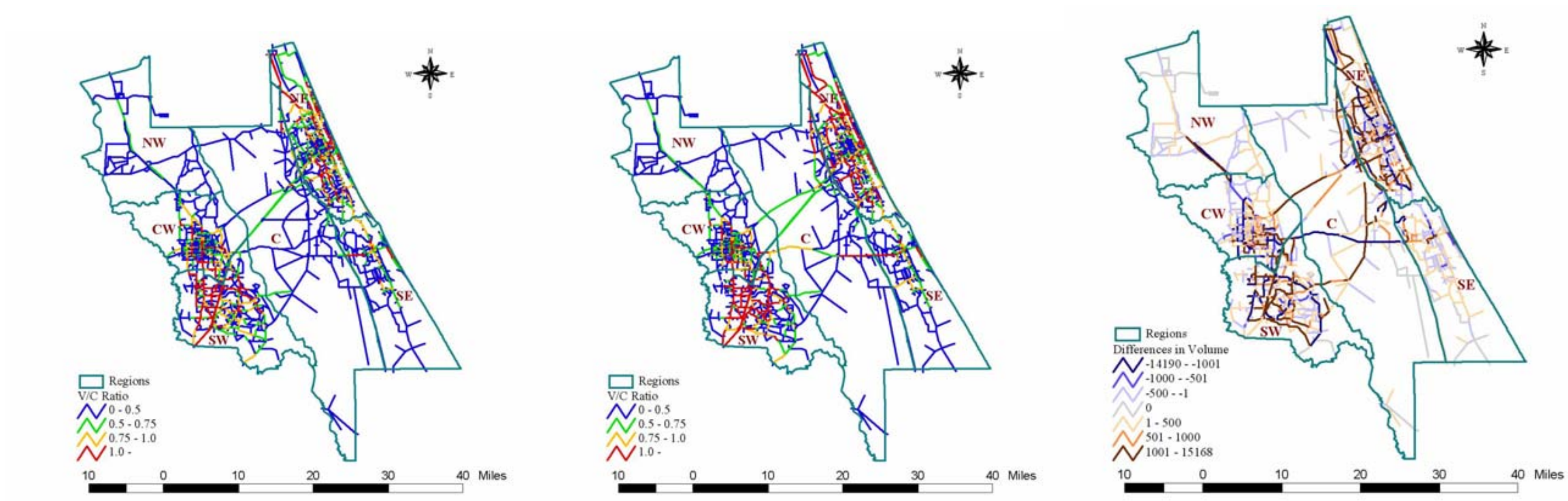


Figure 8.7 Comparison of Alternative 2 and the Final Plan in the 2020 LRTP

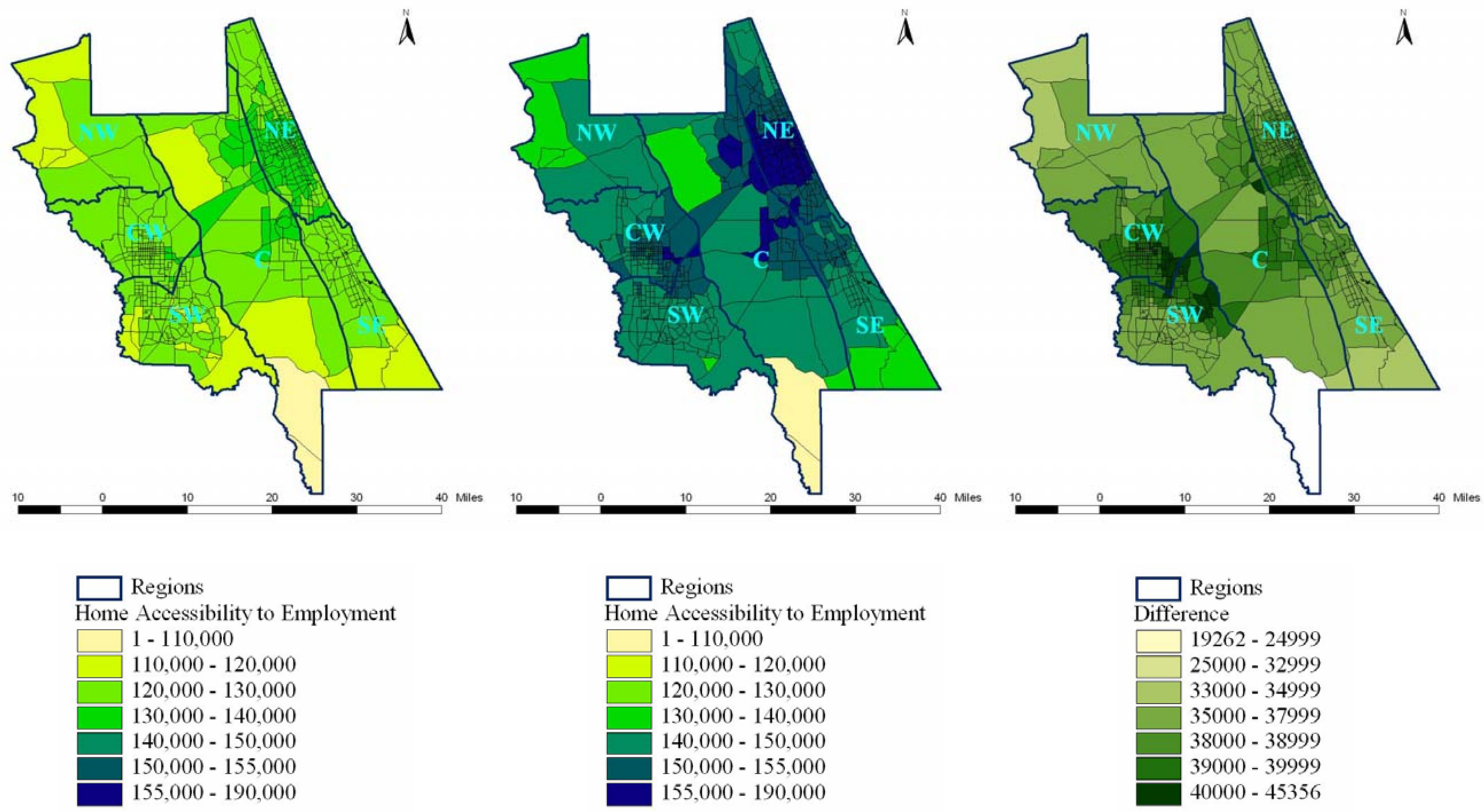
Figure 8.11 shows the population distribution in 2010 and 2020, as well as the differences in population between the base year and 2020. As shown in Figure 8.11 (c), the patterns of population changes based on scenarios 1 and 2 are similar. The large population growth is found in the northeast, southeast, central-west, and southwest regions, and the large population loss is found in the southwest region.

The household distributions in 2010 and 2020 are depicted in Figure 8.12 (a) and (b), respectively. The household changes from the base year to 2020 are illustrated in Figure 8.12 (c). The pattern of household changes is very similar to that of scenario 1. Figure 8.13 displays the employment distributions for 2010 and 2020, as well as the employment growth pattern; the latter is similar to that of scenario 1, even though some TAZs have growths of different magnitudes. Comparisons of the household and employment data from scenario 1 and scenario 2 for 2010 and 2020 are presented in Table 8.3 and 8.4, respectively.



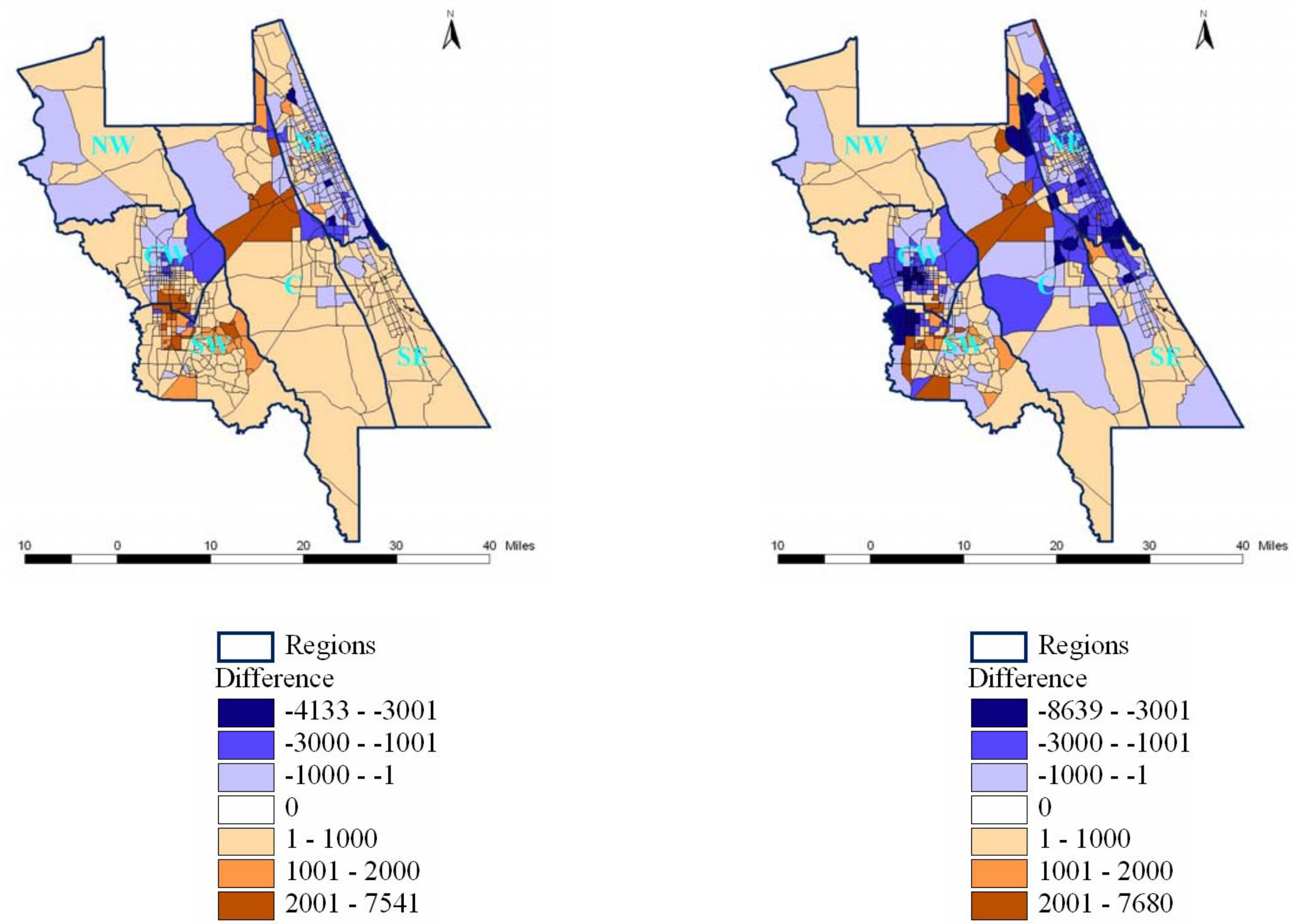
(a) Scenario 1 (b) Scenario 2 (c) Scenario 2 – Scenario 1

Figure 8.8 Comparison of V/C Ratios in 2020 between Scenario 1 and Scenario 2



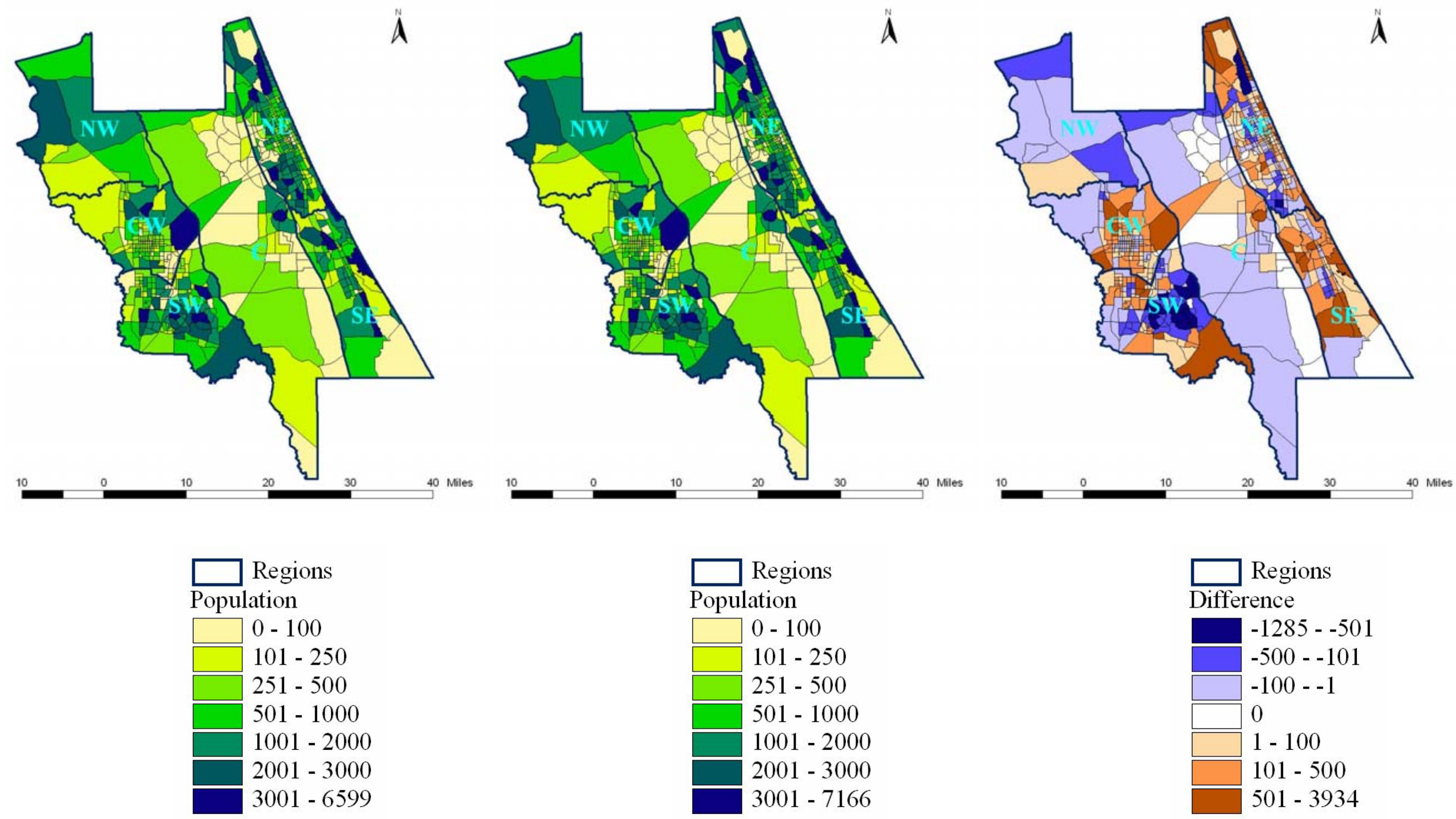
(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.9 Home Accessibility for One-Vehicle Households to Employment from Scenario 2



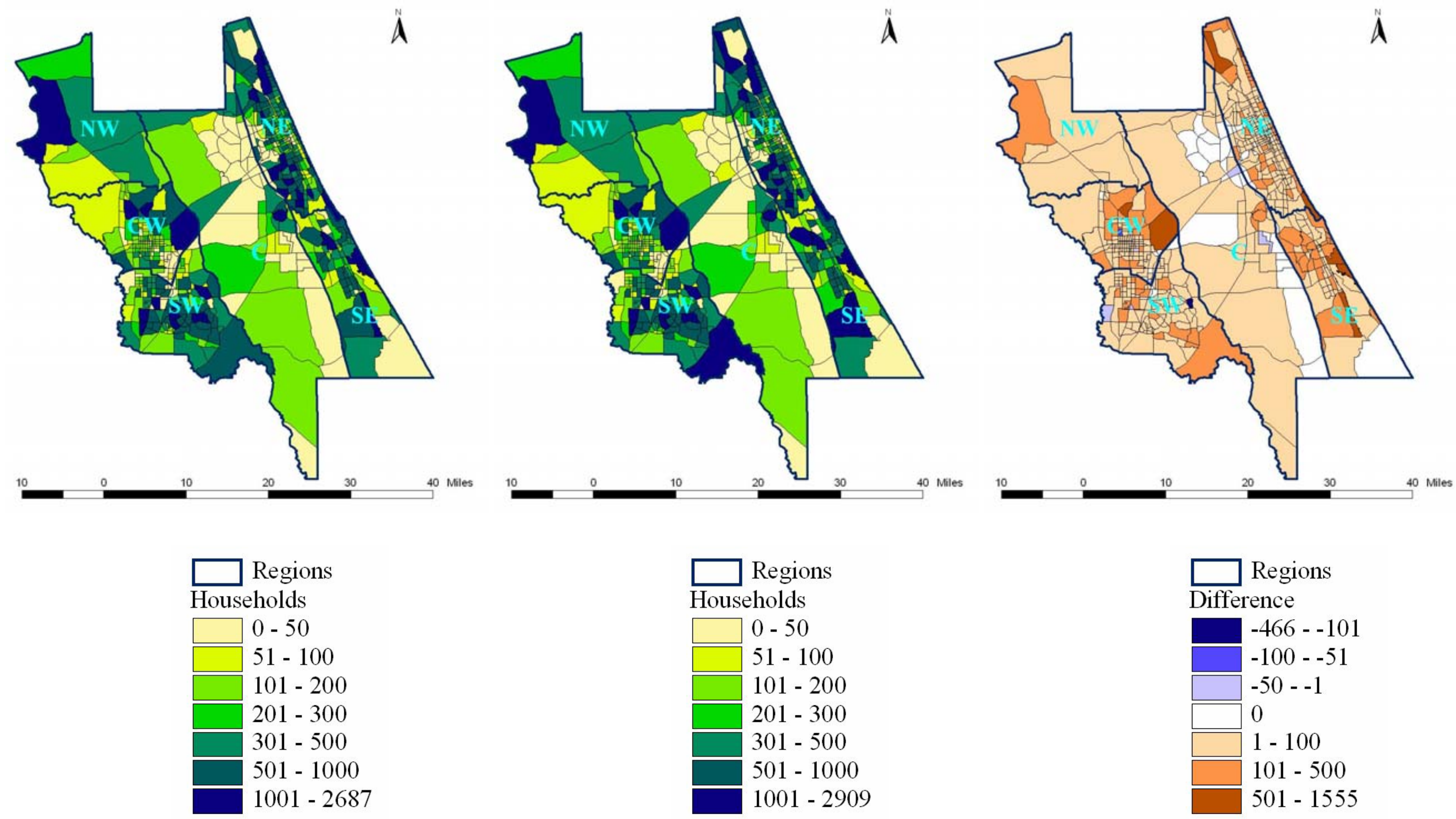
(a) Year 2010 (b) Year 2020

Figure 8.10 Accessibility Difference between Scenario 1 and Scenario 2 (Scenario 1 – Scenario 2)



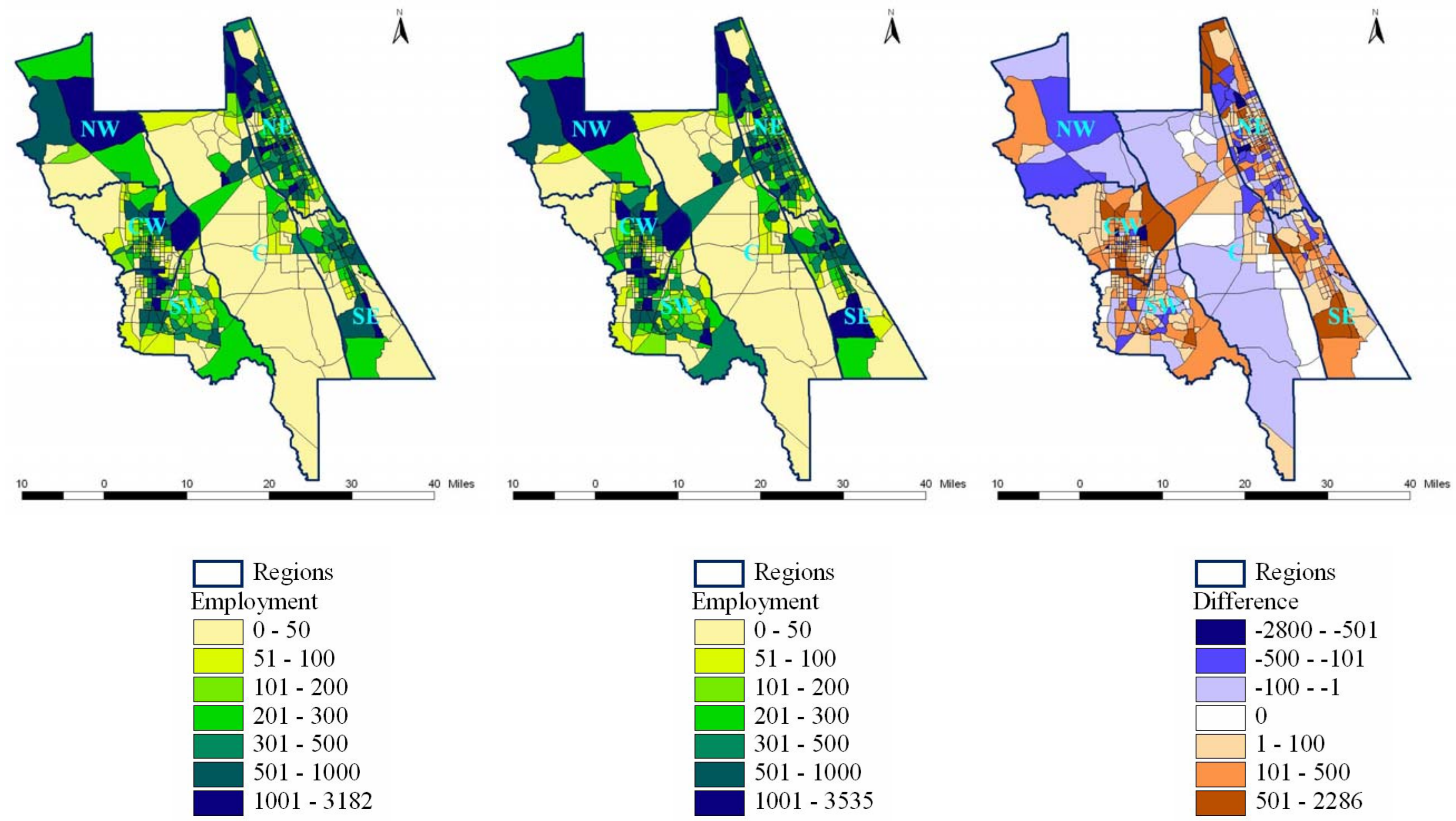
(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.11 Spatial Distribution of Population from Scenario 2



(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.12 Spatial Distribution of Households from Scenario 2



(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.13 Spatial Distribution of Employment from Scenario 2

Table 8.3 Comparison of Households between Scenario1 and Scenario 2

Planning Region	2010		2020	
	Scenario 1	Scenario 2	Scenario 1	Scenario 2
Northeast	96,383	96,515	97,987	98,145
Southeast	30,424	30,486	32,410	32,372
Central	8,087	8,077	8,505	8,492
Northwest	2,741	2,748	2,842	2,848
Central-west	25,992	25,993	28,937	28,766
Southwest	47,198	47,214	49,367	49,391

Table 8.4 Comparison of Employment between Scenario1 and Scenario 2

Planning Region	2010		2020	
	Scenario 1	Scenario 2	Scenario 1	Scenario 2
Northeast	106,643	106,420	118,548	118,318
Southeast	19,039	19,647	23,803	24,561
Central	8,369	8,133	10,565	10,096
Northwest	2,927	2,958	3,055	2,985
Central-west	36,443	36,427	46,432	46,676
Southwest	26,049	25,999	31,969	31,781

8.3 Scenario 3 (Transportation Plan Alternative 3 and Mid-Rang Projection)

Scenario 3 is based on the third alternative transportation improvement plan in the 2020 LRTP and mid-projection of the socioeconomic and demographic data. The third alternative transportation improvement plan is displayed in Figure 8.14 and compared with the final plan. The improvements, indicated in blue, represent common improvements for both the third and the final plans. The green and red colors indicate the third alternative plan and the final plan, respectively. Scenario 3 is only different from scenario 1 in terms of the transportation improvement plan.

Figure 8.15 shows the V/C ratios from scenarios 1 and 3, which illustrate the volume differences between the two scenarios. The volume differences are calculated by subtracting the volumes scenario 1 from the volumes in scenario 3. Site level analyses are necessary to identify the impact from different improvements of the third alternative plan and the final plan.

The accessibility distributions in 2010 and 2020 that are based on scenario 3 (shown in Figure 8.16 (a) and (b), respectively) have the same pattern as those based on scenarios 1 and 2. The pattern of accessibility changes in Figure 8.16 (c), which is different from that of scenarios 1 and 2. Scenarios 1 and 3 are compared in terms of accessibility at the TAZ level for years 2010 and 2020, as displayed in Figure 8.17. Scenario 1 has lower accessibility in many TAZs than scenario 3 in 2010 and 2020.

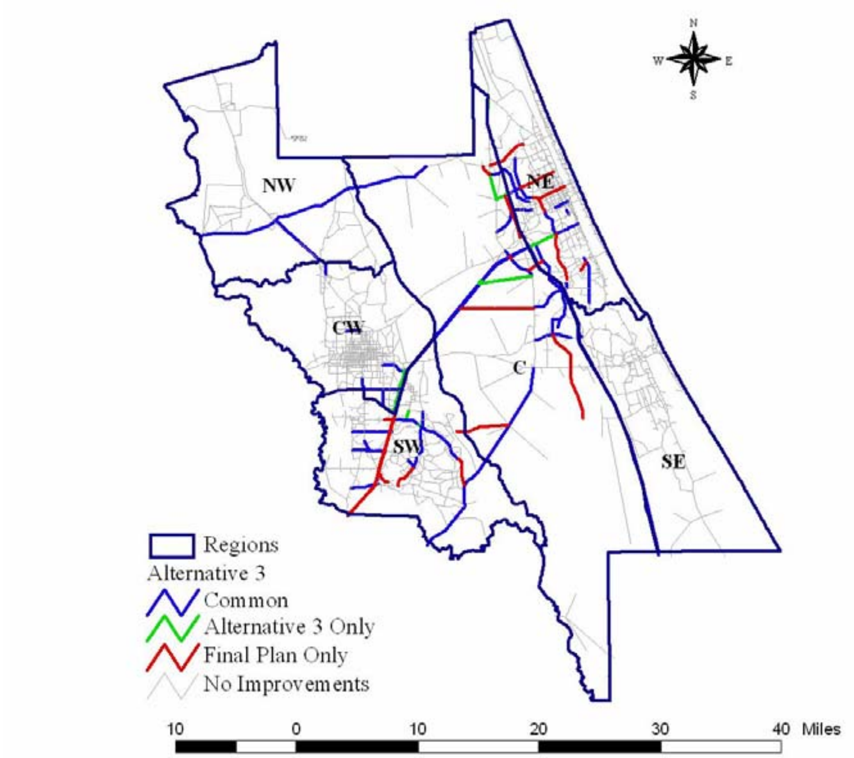
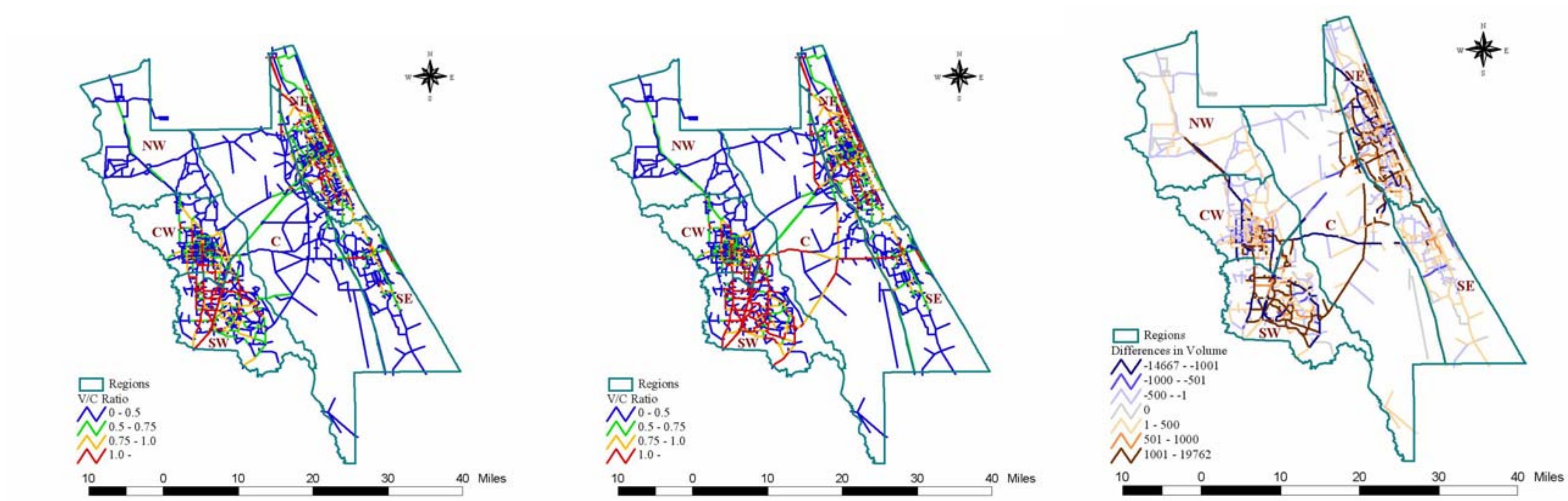


Figure 8.14 Comparison of Alternative 3 and the Final Plan in the 2020 LRTP

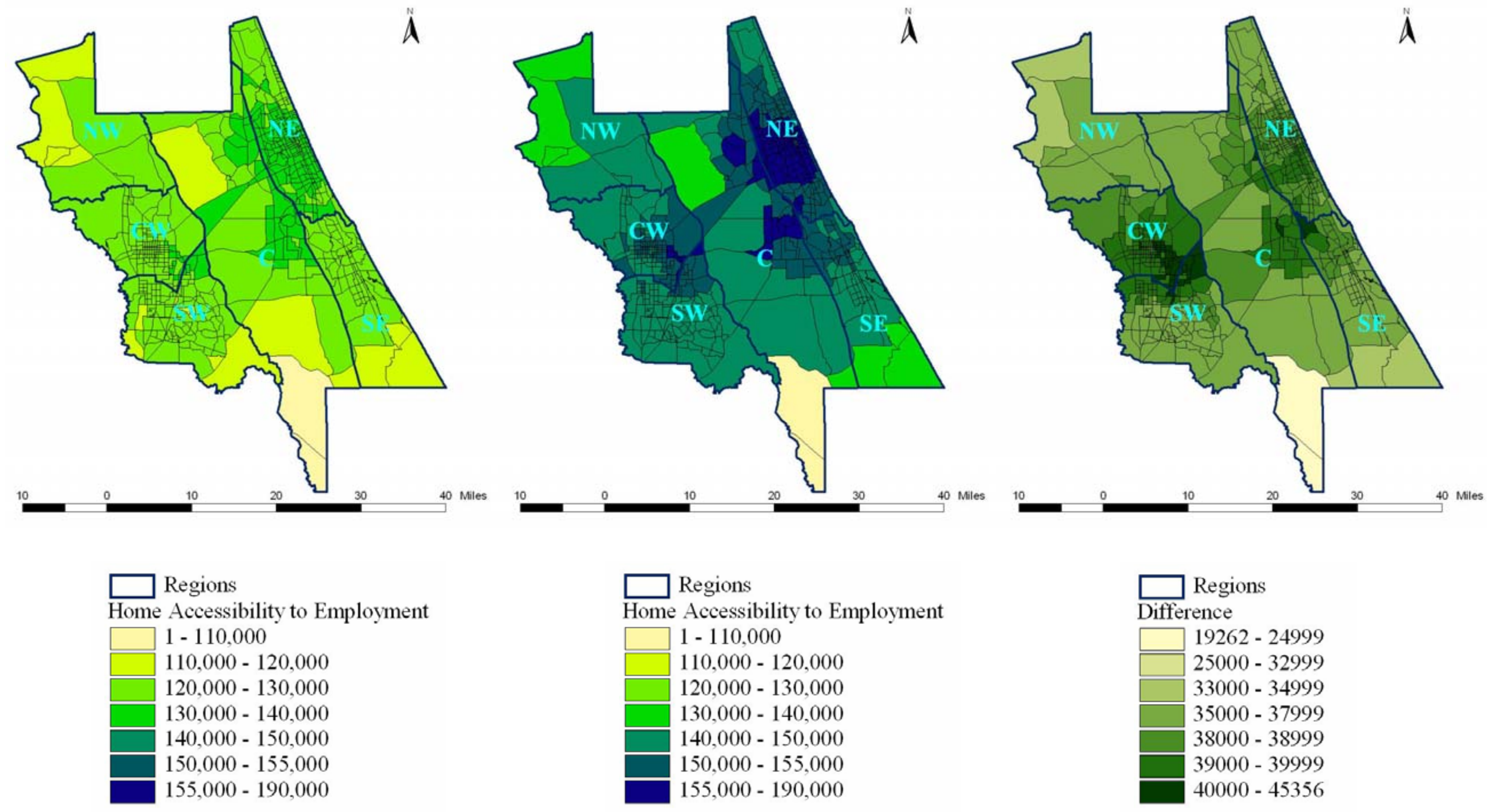
The population distributions in 2010 and 2020 are shown in Figures 8.18 (a) and (b), respectively, and the population growth between the base year and 2020 is displayed in Figure 8.18 (c). Many TAZs in the central west and southeast regions have population growths, while many TAZs in the central, northwest, and southwest regions experience population losses. In the northeast region, some TAZs have population growths and others losses.

Figures 8.19 and 8.20 show household and employment distributions, respectively, in 2010 and 2020, as well as their changes from the base year to 2020. Like scenario 2, the patterns of household and employment changes from the base year to 2020 are very similar to those from scenario 1. The reason for this may be that the transportation plans did not produce significant differences in accessibility and as a result, did not have a strong impact on land use changes. Table 8.5 and 8.6 compare regional household and employment distributions by regions from scenario 1 and scenario 3, respectively.



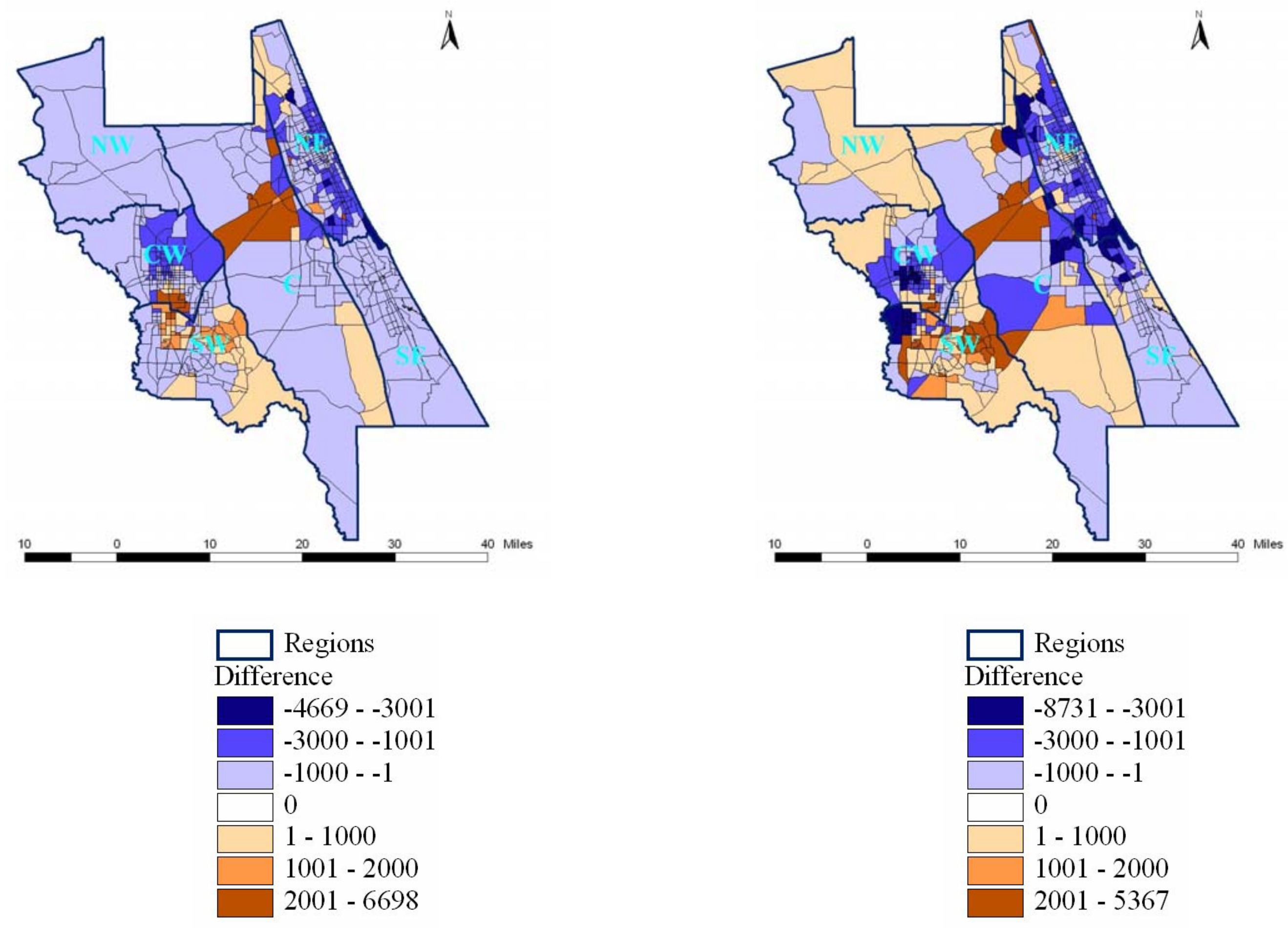
(a) Scenario 1 (b) Scenario 3 (c) Scenario 3 – Scenario 1

Figure 8.15 Comparison of V/C Ratios in 2020 between Scenario 1 and Scenario 3



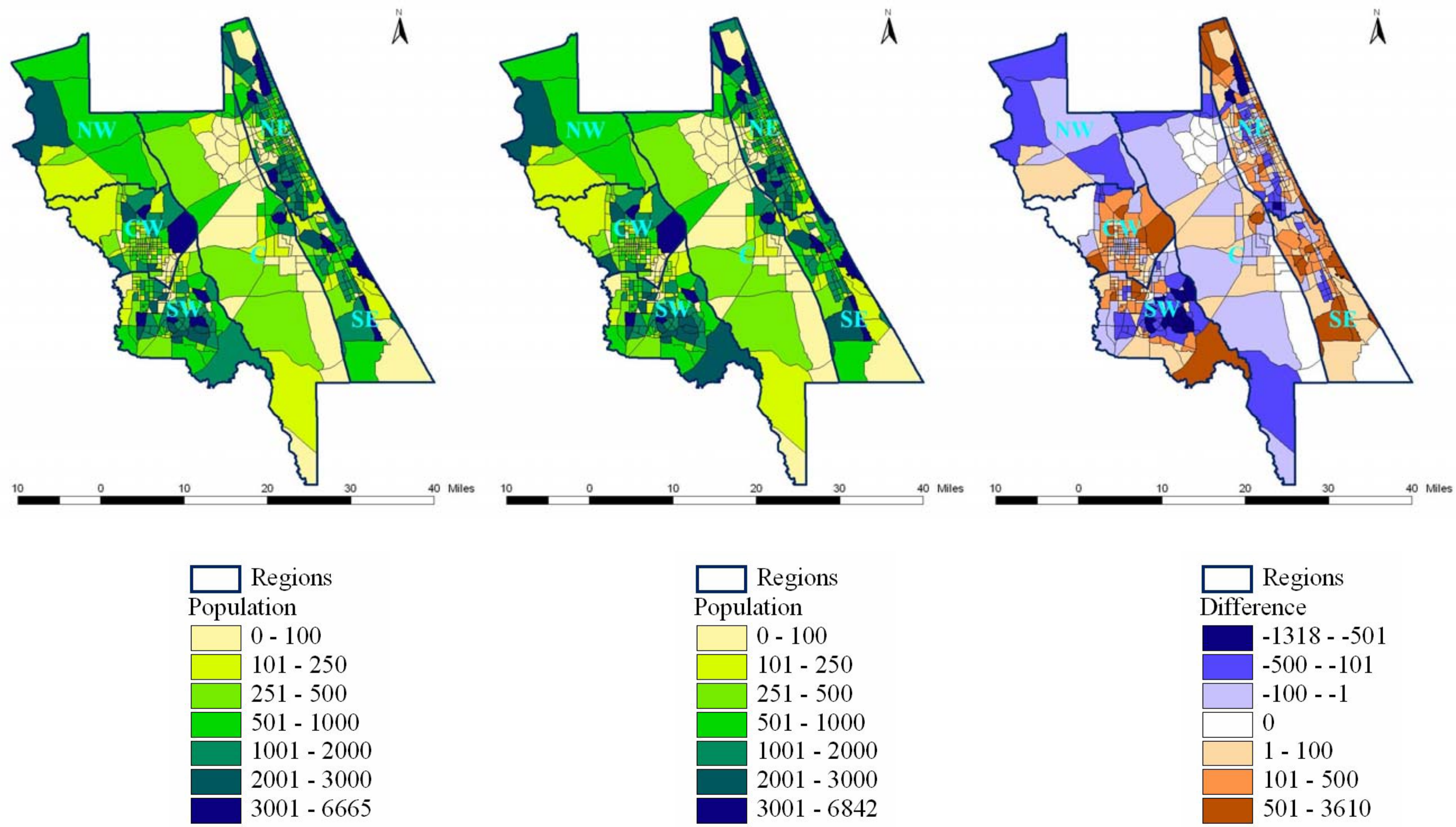
(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.16 Home Accessibility for One-Vehicle Households to Employment from Scenario 3



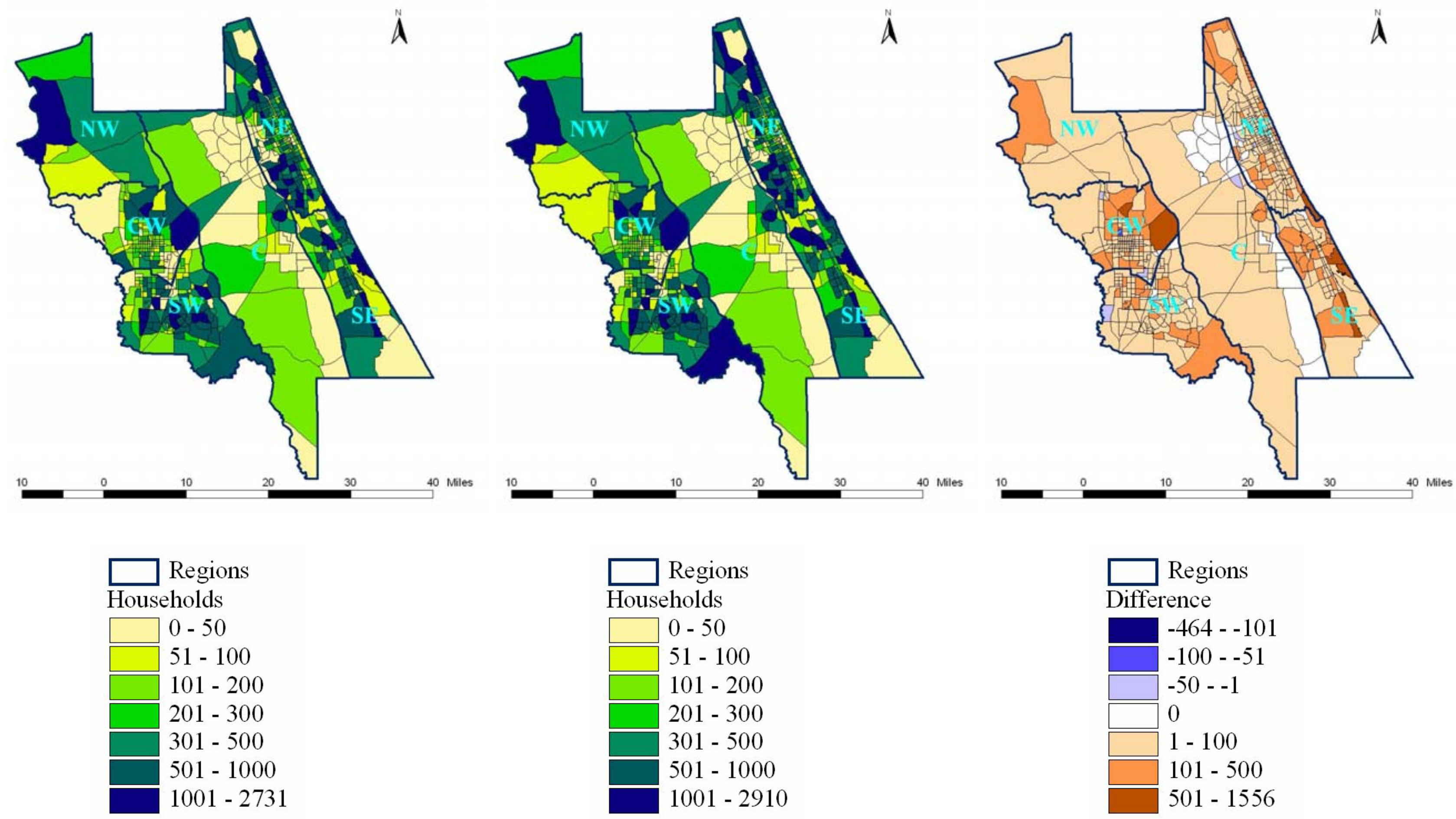
(a) Year 2010 (b) Year 2020

Figure 8.17 Accessibility Difference between Scenario 1 and Scenario 3 (Scenario 1 – Scenario 3)



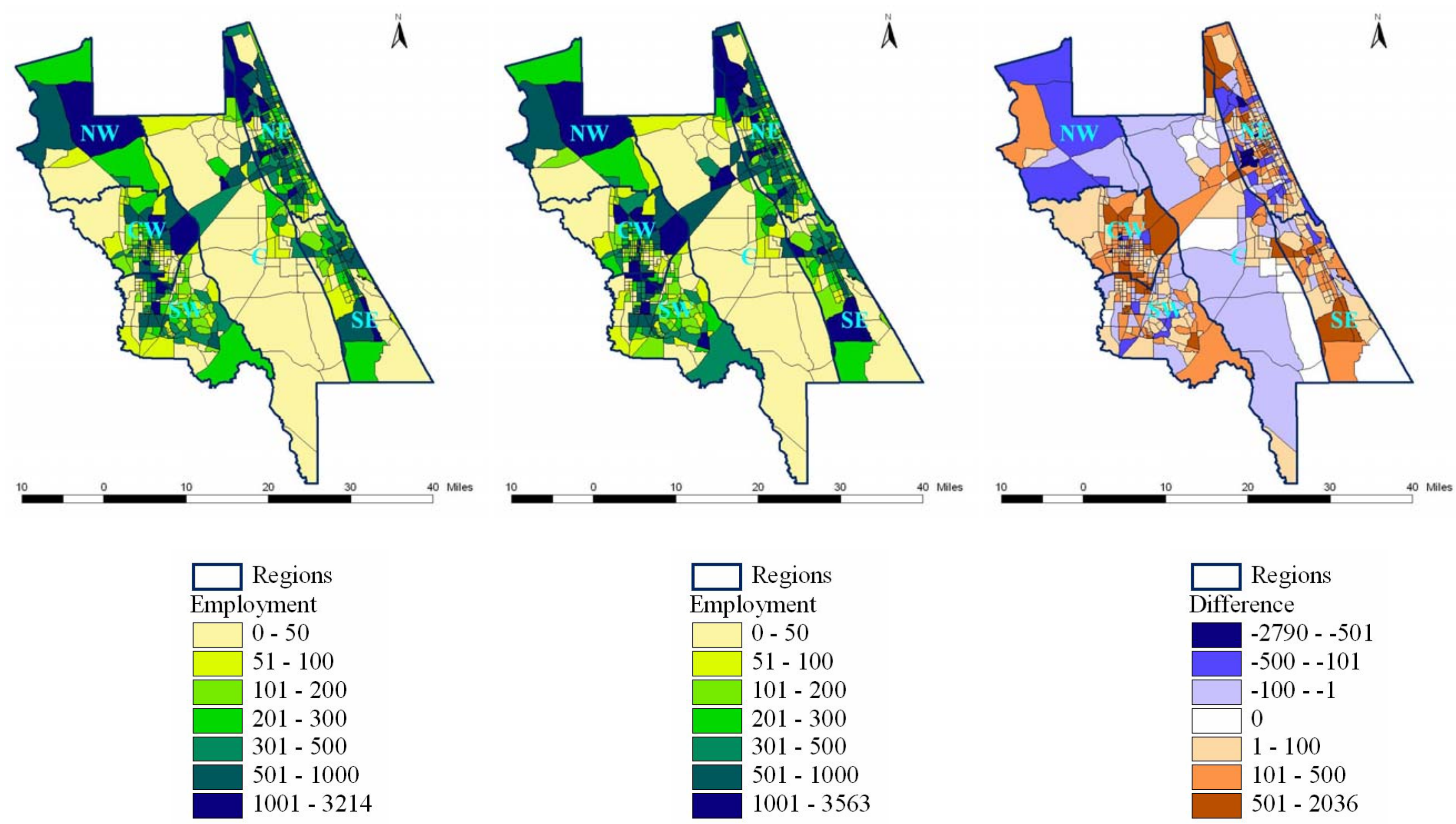
(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.18 Spatial Distribution of Population from Scenario 3



(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.19 Spatial Distribution of Households from Scenario 3



(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.20 Spatial Distribution of Employment from Scenario 3

Table 8.5 Comparison of Households between Scenario1 and Scenario 3

Planning Region	2010		2020	
	Scenario 1	Scenario 3	Scenario 1	Scenario 3
Northeast	96,383	96,397	97,987	97,899
Southeast	30,424	30,470	32,410	32,210
Central	8,087	8,051	8,505	8,448
Northwest	2,741	2,740	2,842	2,847
Central-west	25,992	26,003	28,937	28,916
Southwest	47,198	47,215	49,367	49,263

Table 8.6 Comparison of Employment between Scenario1 and Scenario 3

Planning Region	2010		2020	
	Scenario 1	Scenario 3	Scenario 1	Scenario 3
Northeast	106,643	107,154	118,548	118,100
Southeast	19,039	19,599	23,803	24,726
Central	8,369	8,594	10,565	11,172
Northwest	2,927	2,943	3,055	2,974
Central-west	36,443	36,576	46,432	46,243
Southwest	26,049	25,781	31,969	31,255

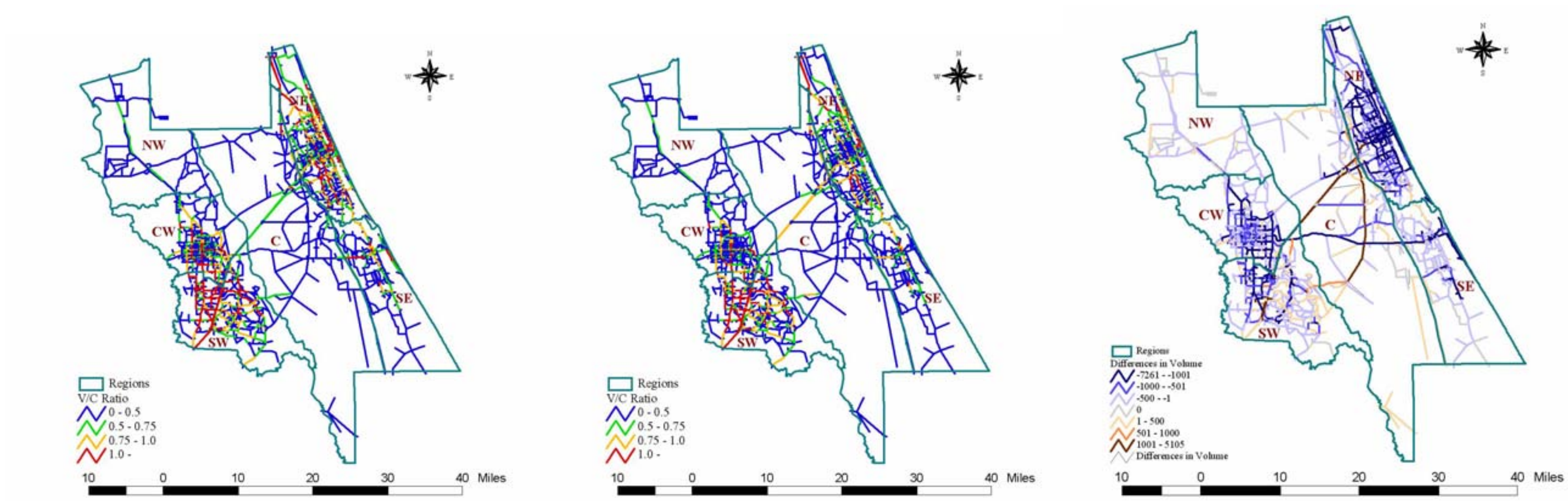
8.4 Scenario 4 (Final Transportation Plan and Low Projection)

Scenario 4 combines the final transportation improvement plan in the 2020 LRTP and the low-projection of socioeconomic and demographic data.

Figure 8.21 shows the V/C ratios from scenarios 1 and 4, along with the volume differences between the two scenarios. Volume differences, which are calculated by subtracting scenario 1 volumes from scenario 4 volumes, indicate that most links have lower volumes in scenario 4 than in scenario 1. This is expected because the socioeconomic and demographic figures used in scenario 4 are lower than those of scenario 1. A few links still have a higher volume in scenario 4 than in scenario 1.

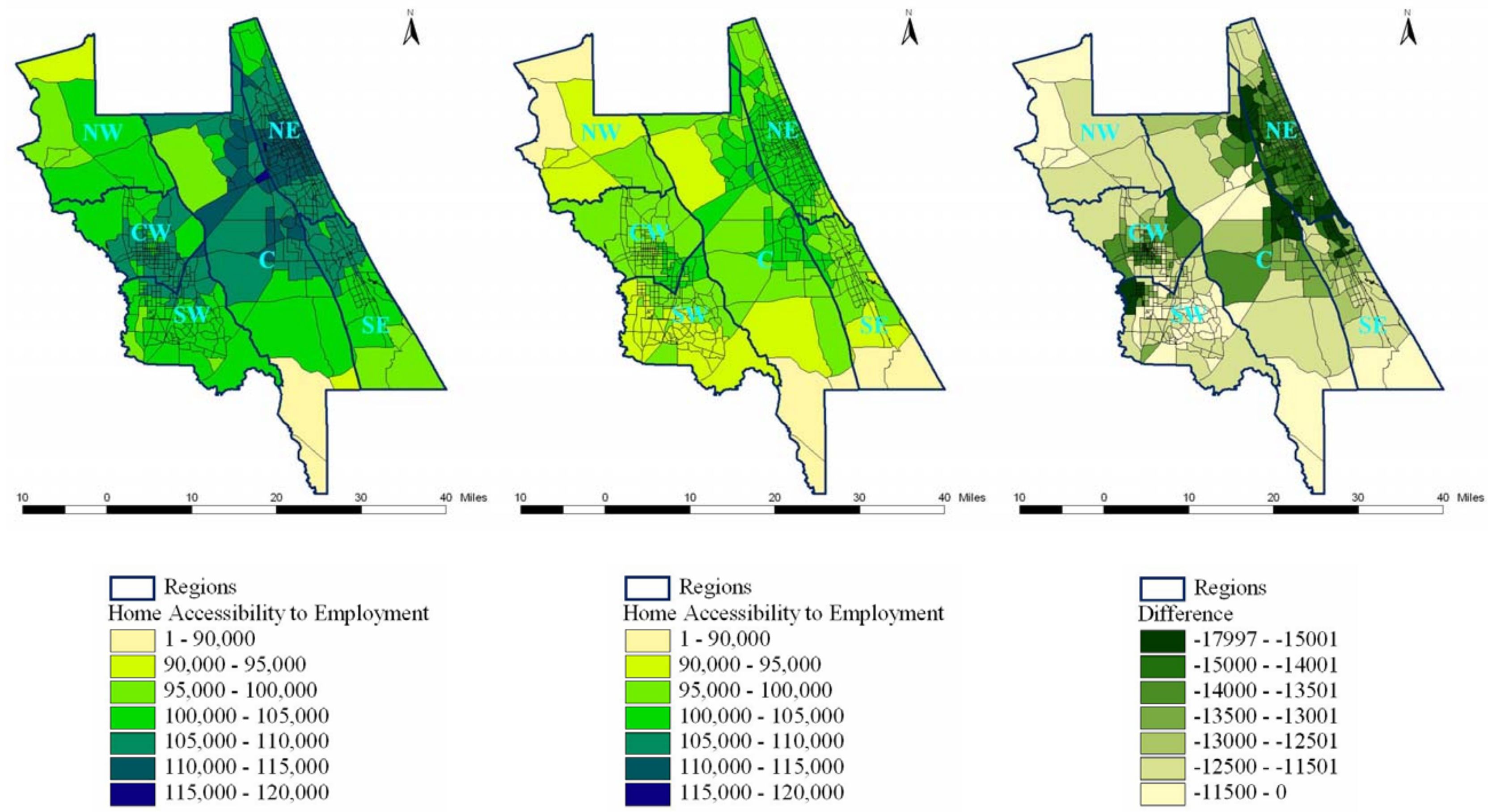
The accessibility from scenario 4 is lower than that from the other scenarios due to the negative growths in households and employment between 2010 and 2020 resulted from the use of the lower projections. Thus, accessibility changes from the base year to 2020 have negative values for all TAZs, as shown in Figure 8.22 (c).

Figure 8.23 depicts the population distributions in 2010 and 2020, as well as population changes from the base year to 2020. A significant population loss is found in the northeast and southwest regions. Some TAZs in the central west, northeast, and southeast regions still show population growths. The distributions of households in 2010 and 2020 and household changes between the base year and 2020 are displayed in Figure 8.24. The pattern of household changes is similar to the pattern of population changes. In Figure 8.25, employment distributions in 2010 and 2020 and the employment changes between the base year and 2020 are shown. The northeast region experiences a large loss of employment. Tables 8.7 and 8.8 present comparisons of regional households and employment from scenario 1 and scenario 4.



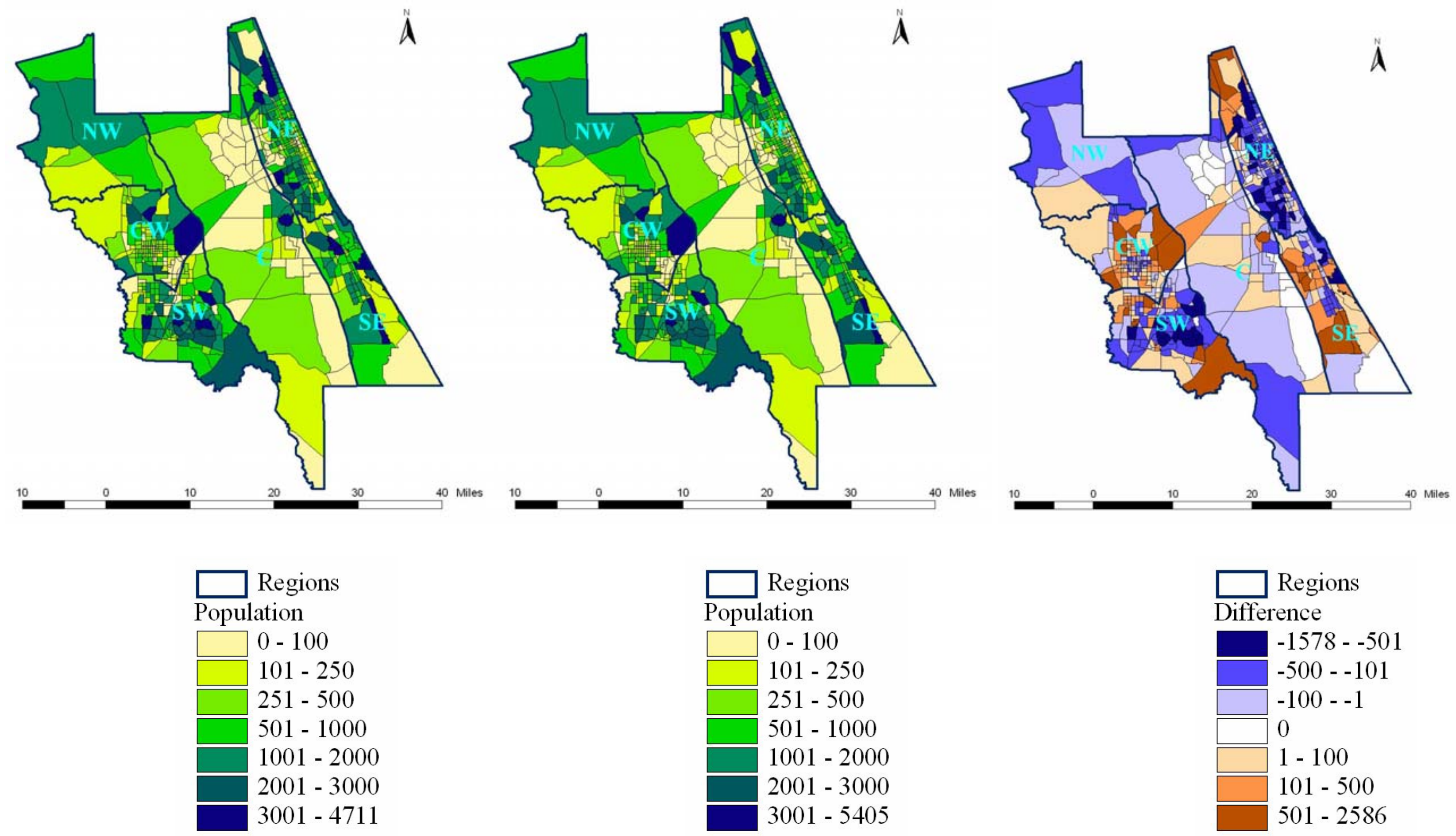
(a) Scenario 1 (b) Scenario 4 (c) Scenario 4 – Scenario 1

Figure 8.21 Comparison of V/C Ratios in 2020 between Scenario 1 and Scenario 4



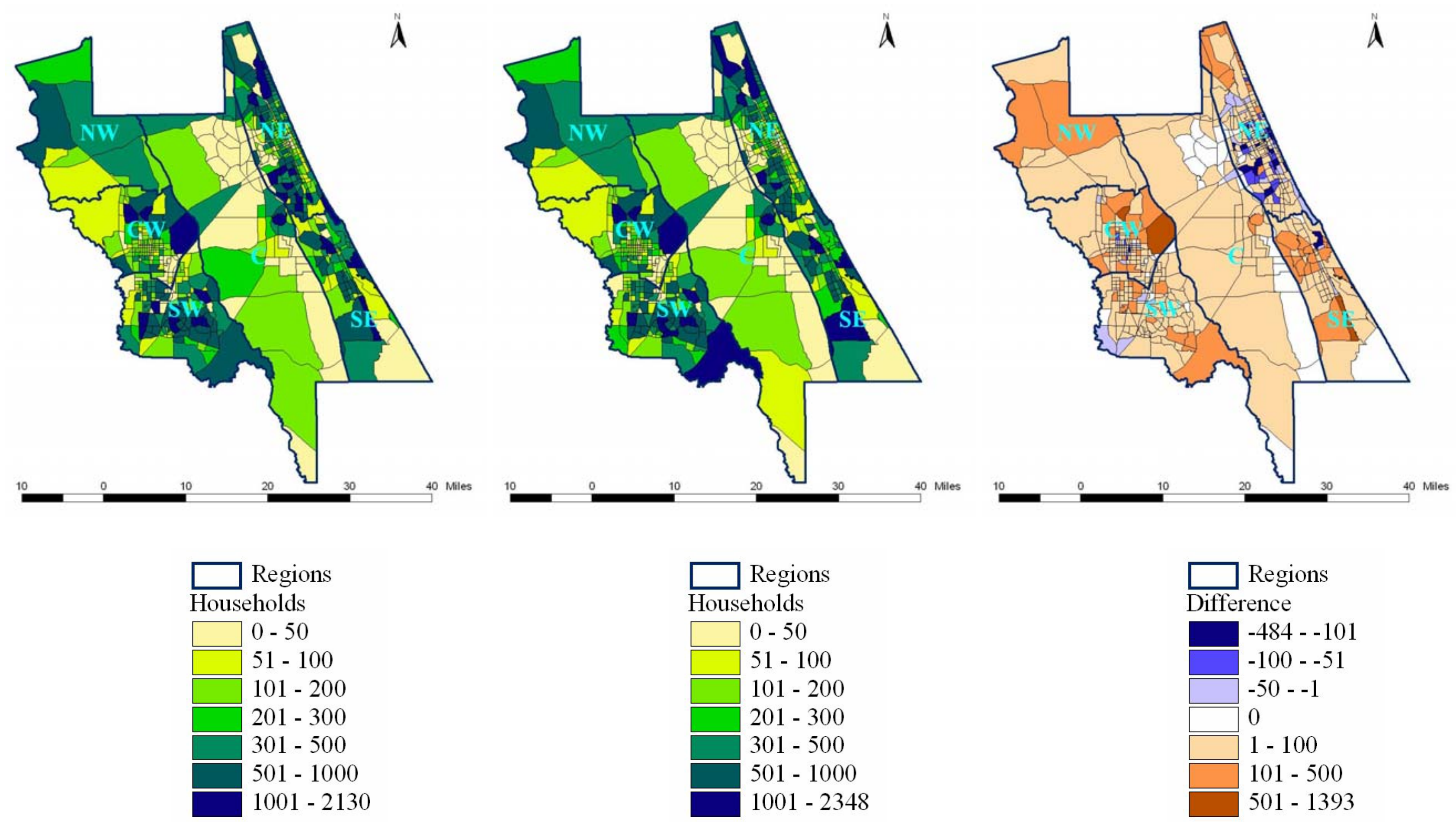
(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.22 Home Accessibility for One-Vehicle Households to Employment from Scenario 4



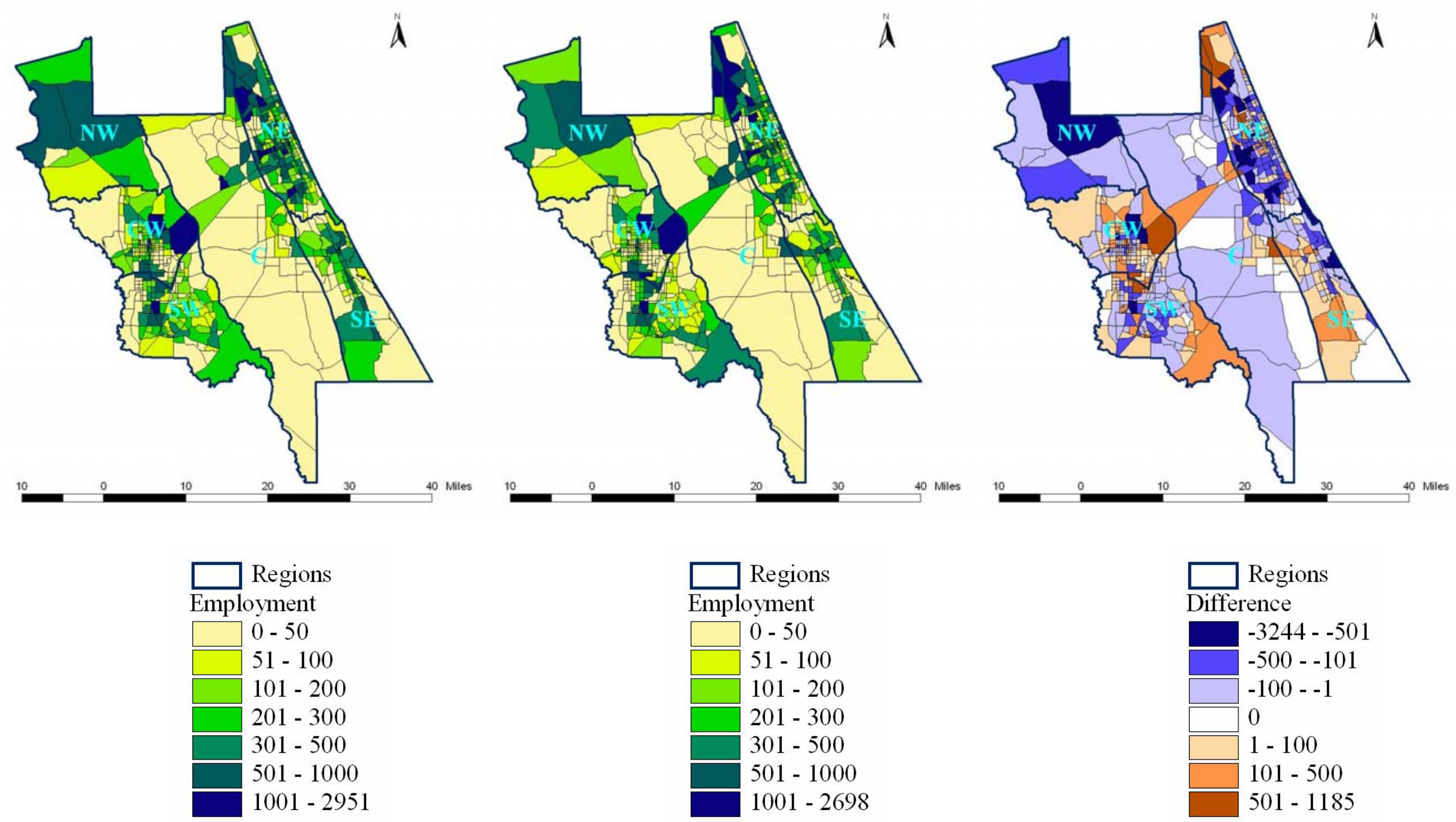
(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.23 Spatial Distribution of Population from Scenario 4



(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.24 Spatial Distribution of Households from Scenario 4



(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.25 Spatial Distribution of Employment from Scenario 4

Table 8.7 Comparison of Households between Scenario1 and Scenario 4

Planning Region	2010		2020	
	Scenario 1	Scenario 4	Scenario 1	Scenario 4
Northeast	96,383	78,728	97,987	63,487
Southeast	30,424	27,121	32,410	25,767
Central	8,087	8,053	8,505	8,135
Northwest	2,741	2,704	2,842	2,593
Central-west	25,992	25,311	28,937	26,405
Southwest	47,198	46,596	49,367	46,032

Table 8.8 Comparison of Employment between Scenario1 and Scenario 4

Planning Region	2010		2020	
	Scenario 1	Scenario 4	Scenario 1	Scenario 4
Northeast	106,643	91,757	118,548	82,841
Southeast	19,039	15,076	23,803	13,157
Central	8,369	6,402	10,565	7,373
Northwest	2,927	2,567	3,055	2,036
Central-west	36,443	28,679	46,432	27,235
Southwest	26,049	20,471	31,969	18,612

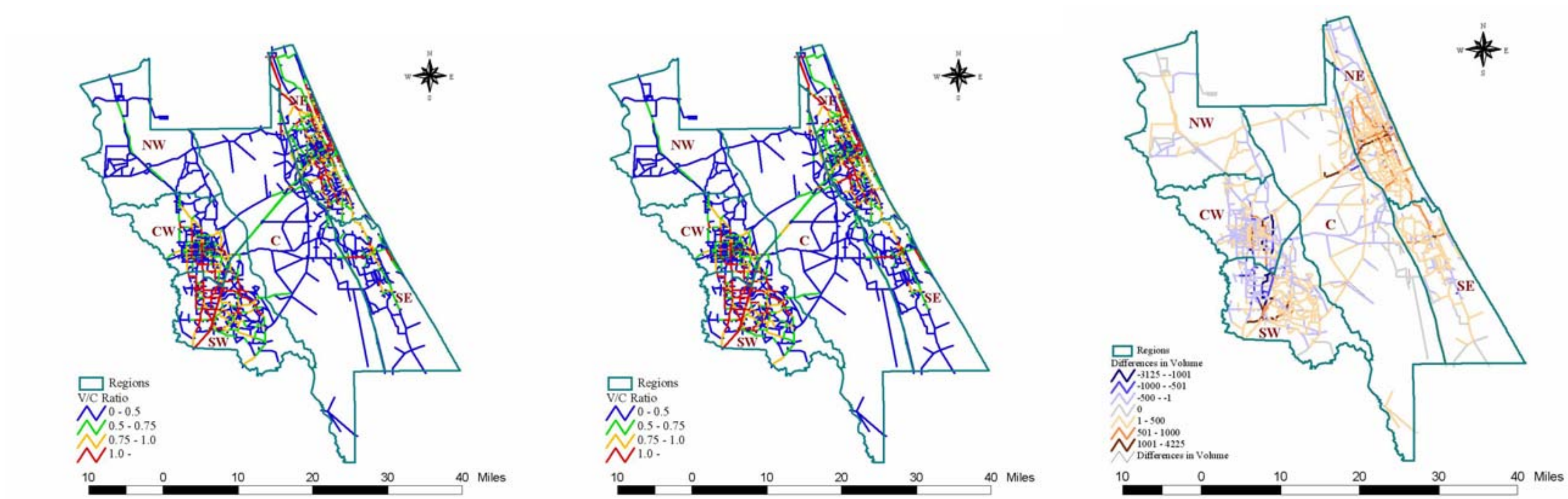
8.5 Scenario 5 (Final Transportation Plan and High Projection)

Scenario 5 is based on the final transportation improvement plan in the 2020 LRTP and high-projection of socioeconomic and demographic data.

Figure 8.26 (a) and (b) display the V/C ratios from scenarios 1 and 5, and Figure 8.26 (c) show the volume differences computed by subtracting the scenario 1 volumes from scenario 5 volumes. It may be seen that the links in the northeast region have higher volumes in scenario 5 than in scenario 1. Some links in the central-west and southwest regions, however, have lower volumes in scenario 5.

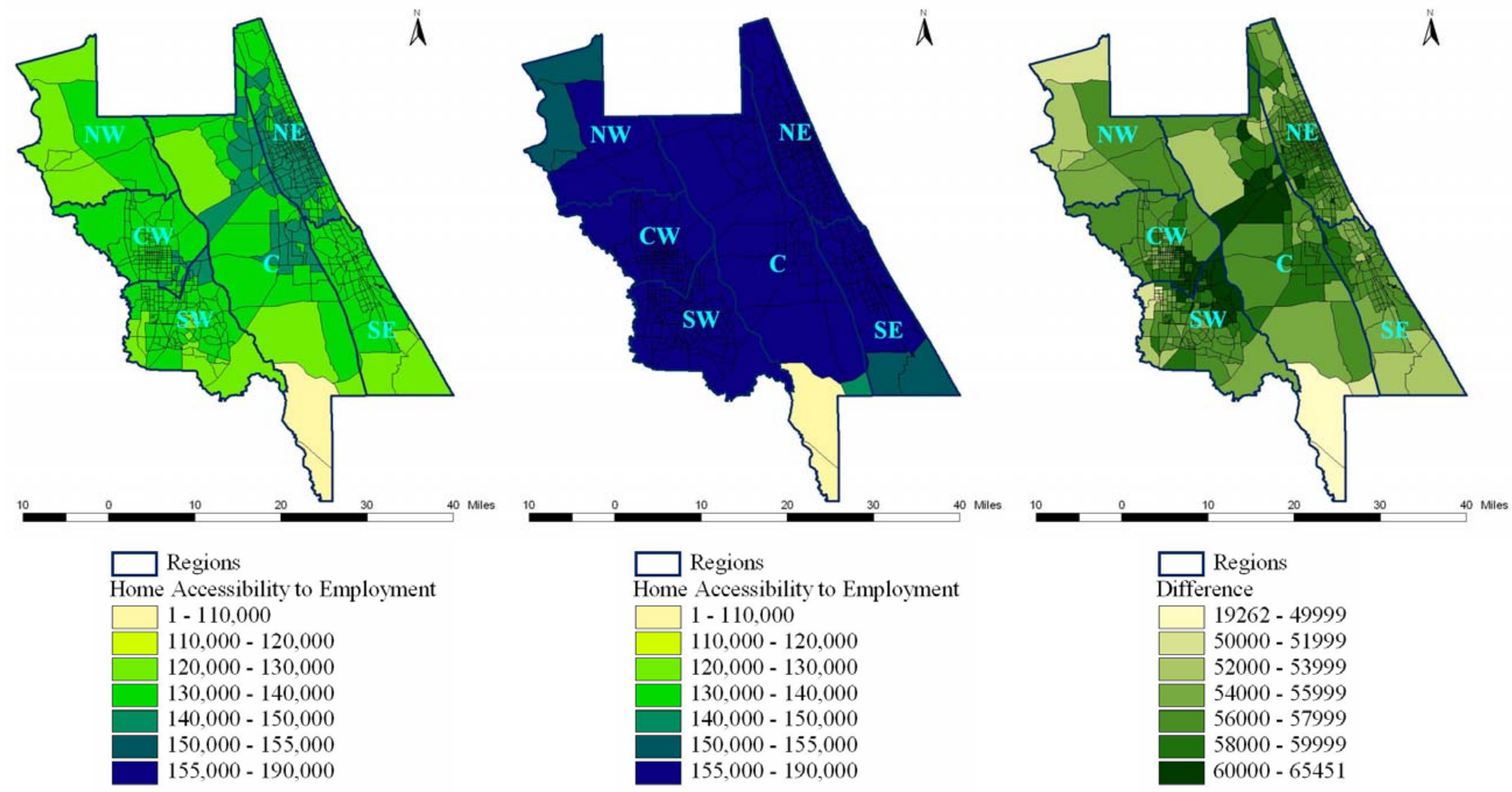
Compared to other scenarios, accessibilities in 2010 and 2020 from scenario 5 are relatively high, as shown in Figure 8.27. This is the result of using the high projection of socioeconomic and demographic data for control totals. Distributions of accessibility in 2010 and 2020 are displayed in Figure 8.27. Accessibility changes between the base year and 2020 are shown in Figure 8.27 (c). According to the figure, the accessibilities of TAZs around the I-4 in the central, central-west, and southwest regions increase significantly.

Population, household, and employment distributions are displayed in Figure 8.28 through 8.30. The spatial patterns of population, households, and employment for scenario 5 are similar to those for scenario 1. The reason may be that the location choice models distribute household and employment at favorable locations in terms of accessibility, land price, vacant land, and so on (refer to Chapter 5 for a detailed list of the factors). Comparisons of regional households and employment from scenarios 1 and 4 are shown in Tables 8.9 and 8.10, respectively.



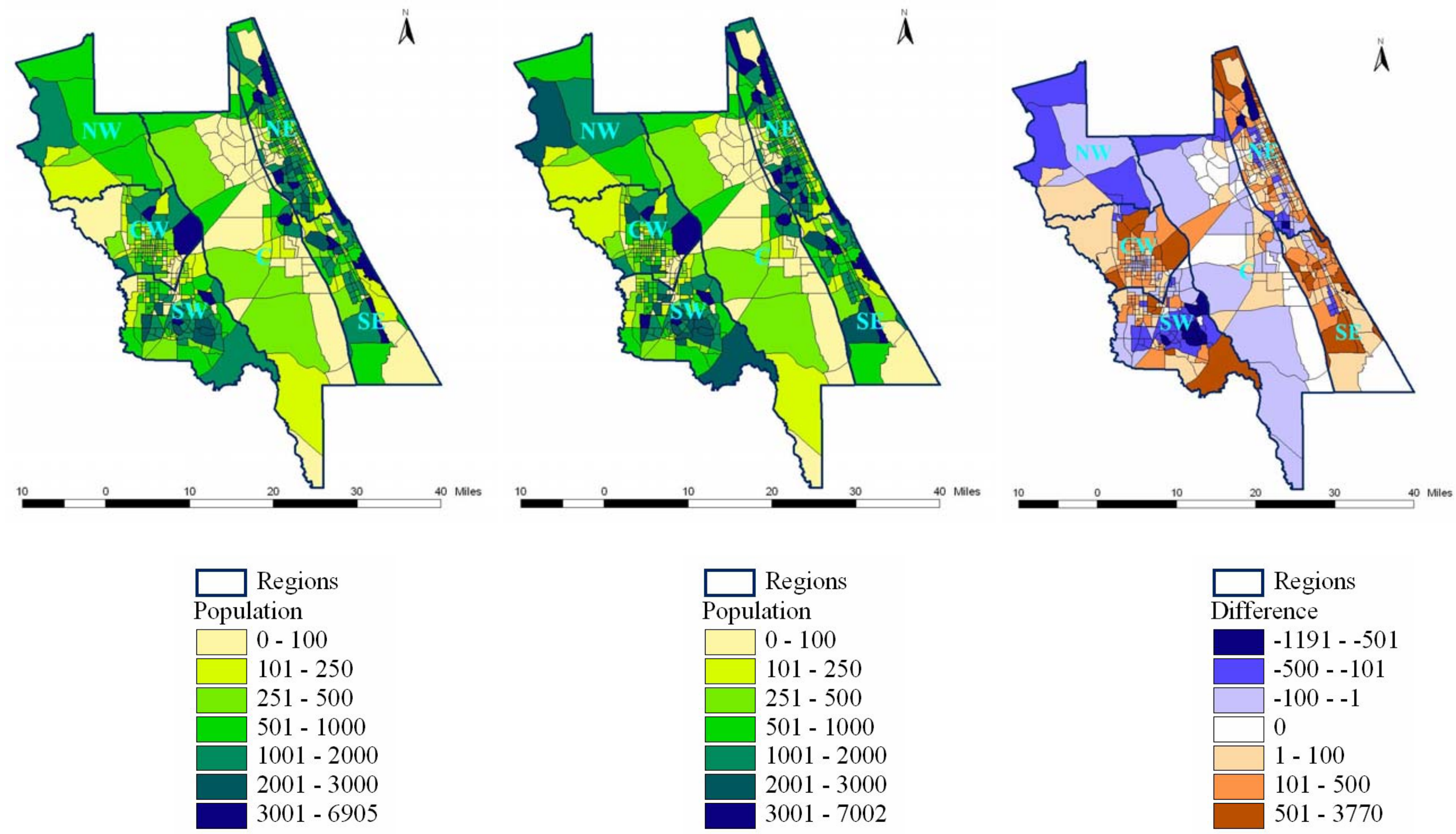
(a) Scenario 1 (b) Scenario 5 (c) Scenario 5 – Scenario 1

Figure 8.26 Comparison of V/C Ratio in 2020 between Scenario 1 and Scenario 5



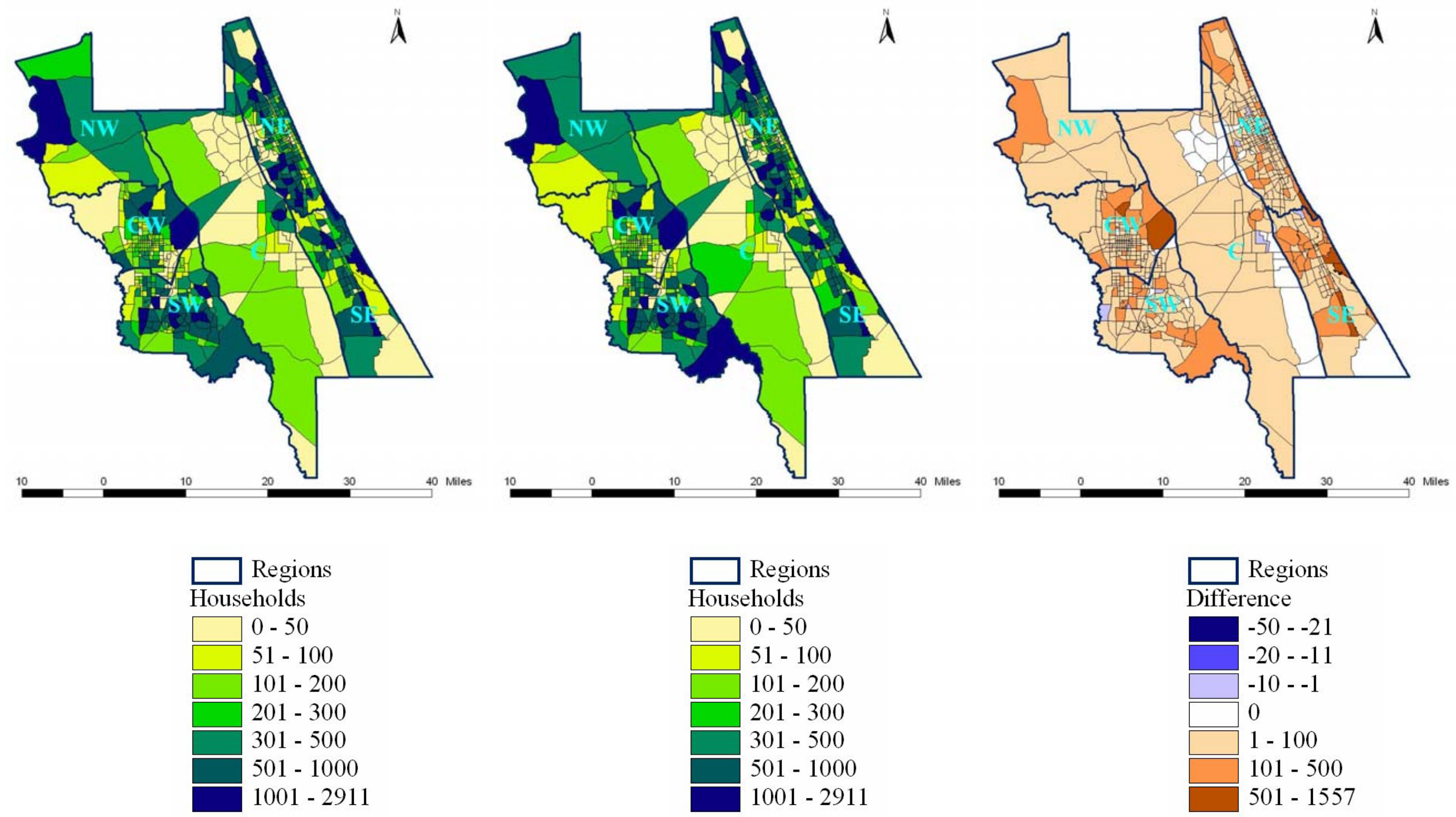
(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.27 Home Accessibility for One-Vehicle Households to Employment from Scenario 5



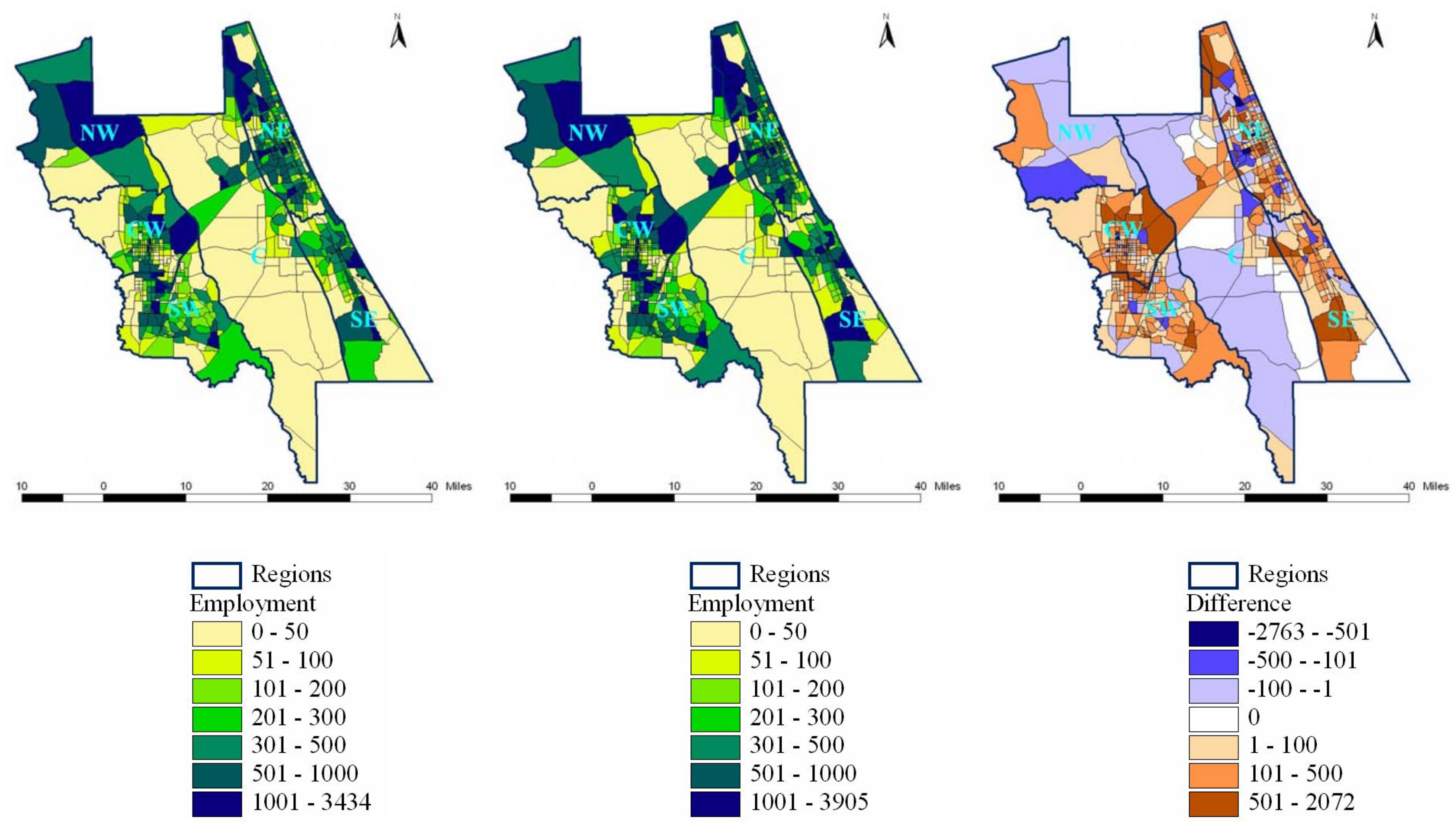
(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.28 Spatial Distribution of Population from Scenario 5



(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.29 Spatial Distribution of Households from Scenario 5



(a) Year 2010 (b) Year 2020 (c) 2020 – Base Year

Figure 8.30 Spatial Distribution of Employment from Scenario 5

Table 8.9 Comparison of Households between Scenario1 and Scenario 5

Planning Region	2010		2020	
	Scenario 1	Scenario 5	Scenario 1	Scenario 5
Northeast	96,383	96,325	97,987	97,766
Southeast	30,424	30,511	32,410	31,854
Central	8,087	8,002	8,505	8,310
Northwest	2,741	2,759	2,842	2,835
Central-west	25,992	25,687	28,937	28,183
Southwest	47,198	46,827	49,367	48,572

Table 8.10 Comparison of Employment between Scenario1 and Scenario 5

Planning Region	2010		2020	
	Scenario 1	Scenario 5	Scenario 1	Scenario 5
Northeast	106,643	115,927	118,548	136,304
Southeast	19,039	21,412	23,803	26,950
Central	8,369	8,969	10,565	12,147
Northwest	2,927	3,229	3,055	3,621
Central-west	36,443	38,156	46,432	50,922
Southwest	26,049	28,565	31,969	36,708

9. CONCLUSIONS AND RECOMMENDATIONS

Currently, there are many land use models proposed or in use. An extensive literature review reveals that UrbanSim, a micro-simulation land use model, is promising for the following reasons:

- It is based on theories of market economy and discrete choice behavior, therefore it captures both the impacts of market forces as well individuals' choices on land development processes.
- It is spatially disaggregated by using small grid cells and parcel data to model land use. The simulation is household and job-based, making it more realistic.
- It is temporally disaggregated by simulating land use changes on an annual basis.
- It models the dynamics of the land use and transportation interactions and the disequilibrium between the two systems, caused by the time lags before one system fully responds to changes in the other.
- It is designed for integration with a travel demand model. Its disaggregated nature also lends itself to activity-based travel demand modeling.

UrbanSim has been applied to Volusia County, Florida for the purpose of assessing the model's accuracy and investigating issues related to implementation. The UrbanSim model was run jointly with a FSUTMS/TRANPLAN model. Interfacing UrbanSim with the FSUTMS requires an additional program to convert the output from one model to the input of the other, which is accomplished by developing an ArcView conversion program. Although the conversion program was not designed to provide full automation to help feedback between the two models, it simplifies the data processing.

The following are the findings from this study:

- UrbanSim has been found to simulate land use changes reasonably well, although detailed analyses will help further understand the model's behavior and reasons for the differences between UrbanSim predictions and projections by Volusia County. It is necessary to point out that differences are expected since these projections are produced by using different methods. Therefore, attempts need to be directed at the possible causes of the differences, including the assumptions used.

- Many urban areas go through a "consensus building" process when allocating growth to different municipalities. In the current version of UrbanSim, such a process cannot be modeled. However, users are allowed to specify land development projects, including location, type, intensity, and implementation schedule as part of a scenario. Thus community visions need to be put in more concrete terms of possible developments so that they may be included in UrbanSim. This also points to the possibility that the existing "consensus building" processes adopted by many local governments may need to be improved to allow community visions to be better reflected through the model.

- Consultations with local government agencies are desirable in developing model specifications and estimating model parameters. Location choice models and developer model reflect the behavior of local activities. Consultations with local agencies will help improve model performance.

- Feedback from the travel model to UrbanSim influences the land development patterns. It will be useful to measure the sensitivity of UrbanSim to accessibility to determine the necessary frequency of the feedback. Through feedback, UrbanSim also has the potential of testing the effects of different project schedules on both land use and transportation.

- The most significant efforts in this project are related to data imputation and quality control, as well as model parameter estimation. The main problems with the data are missing or outdated information in the address database (used to locate businesses to parcels), missing information on the number of housing units for multi-family dwellings (including condominiums and apartments), and missing information on properties. While these problems may be addressed separately by the government agencies that created the data, a GIS-based database tool is necessary to facilitate data compilation. This tool needs to provide simple statistics functions to allow examination of the data. ArcGIS is a more suitable GIS platform for this purpose than ArcView; the latter is limited in its database management capabilities.

- UrbanSim is designed to be integrated with life-style travel models. As a result, output from UrbanSim based on life-style household structure needs to be summarized to support the classic travel demand models. On the other hand, it will be a natural fit to life-style models and activity-based models.

- High performance computers are required. For this study, a computer of 3.4 GHz Pentium with 2 GB of RAM is used. The computer time for a 10-year simulation is approximately three hours. The running time may vary by the study area.

- The development of an UrbanSim model requires expertise in both GIS and statistics, the latter for estimating discrete choice models. Some MPOs may not have in-house expertise and may need to rely on services provided by consultants.

- A detailed user manual on data processing is needed. The UrbanSim User manual and technical reports are provided on-line (http://www.urbansim.org/docs/), but information on data processing is inadequate. This may be because UrbanSim continues to be improved and it is not yet a commercial product.

- A TAZ-based UrbanSim model will reduce the amount of data processing involved. The University of Washington has been in the process of developing such a model. However, because the TAZs are usually much larger spatial units than the current grid system used in UrbanSim (150 meters by 150 meters), it is unclear whether model accuracy would be affected by the reduced spatial resolution.

REFERENCE

Allen, E. (2001), "INDEX: Software for Community Indicators," *Planning Support Systems: Integrating Geographic Information Systems, Models, and Visualization Tools*, R. Brail (ed.), Rutgers University, Center for Urban Policy Research and ESRI Press, Available from http://www.crit.com/documents/ESRIChapter.pdf.

Anas, A., R. Arnott and K. A. Small (1998), "Urban Spatial Structure," *Journal of Economics Literature*, Vol. XXXVI, pp. 1426-1464.

Berry, M. W., R. O. Flamm, B. C. Hazen, and R. L. MacIntyre (1996), "Lucas: A System for Modeling Land-Use Change," *IEEE Computational Science & Engineering*, Vol. 3. No. 1, pp. 24-35.

Batty, M. and Y. Xie (1994), "From Cells to Cities," *Environment and Planning B*, Vol. 21, pp.31-48.

Cambridge Systematics, Inc. (1991), *Making the Land Use, Transportation, Air Quality Connection*, Technical Report Volume 1, Modeling Practices, prepared for 1000 Friends of Oregon, prepared by Cambridge Systematics, Inc., Cambridge, MA and Hague Consulting Group, Netherlands.

Cambridge Systematics, Inc. (1994), *Making the Land Use, Transportation, Air Quality Connection*, Technical Report Volume 4, Model Modifications, prepared for 1000 Friends of Oregon, prepared by Cambridge Systematics, Inc., Parsons, Brinckerhoff, Quade & Douglas, and S.H.Putman Associates, Inc.

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.

Canin Associates (2000), *Data Development and Analysis with GIS Support*, Year 1 & 2 Deliverables, prepared for Metroplan Orlando, prepared by Canin Associates, Orlando, FL.

Corradino Group (2001), *Update of the Indianapolis Regional Transportation Plan for 2025*, Final Report, prepared for Indianapolis Metropolitan Planning Organization, prepared by the Corradino Group, Knoxville, TN.

Corradino Group (2004), *Simplified Land Allocation Model*, prepared by the Corradino Group, Knoxville, TN. Available from http//www.corradino.com/.

Clarke, K. C., S. Hoppen, and L. Gaydos (1996), "Methods and Techniques for Rigorous Calibration of a Cellular Automaton Model of Urban Growth," *the Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, Santa Fe, New Mexico, January 21-25, 1996.

Couclelis, H. (1996), "Verbal Directions for Way-Finding: Space, Cognition, and Language," *The Construction of Cognitive Maps*, Ed. J. Portugali, pp. 133-153, Kluwer Academic Publishers, Dordrecht, Netherlands.

Croteau, K. G., B. G. Faber, and V. L. Thomas, (1997), "Smart Places: A Tool for Design and Evaluation of Land Use Scenarios," Proceedings of the 1997 ESRI User Conference, Environmental and System Research Institute, San Diego, California, July 8-11, 1997, Available from http://gis.esri.com/library/userconf/proc97/home.htm.

Dickey, J. W. and C. Leiner (1983), "Use of Topaz for Transportation-Land Use Planning in a Suburban County," *Transportation Research Record 931*, Transportation Research Board, National Research Council, Washington, D.C., pp. 20-26.

Dowling, R., R. Ireson, A. Skabardonis, D. Gillen, P. Stopher, A. Horowitz, J. Bowman, E. Deakin, and R. Dulla (2000), *Predicting Short-term and Long-term Air Quality Effects of Traffic-flow Improvement Projects*, Technical Report NCHRP 25-21, Transportation Research Board, National Research Council, Washington, D.C.

Emmi, P. and L. Magnusson (1995), "Opportunity and Mobility in Urban Housing Markets," in *Progress in Planning*, Pergammon Press, New York, NY.

Federal Highway Administration (FHWA) (2004), *Toolbox for Regional Policy Analysis*, Available from http://www.fhwa.dot.gov/planning/toolbox/index.htm.

Herbert, J. and B. H. Stevens (1960), "A Model for the Distribution of Residential Activity in Urban Areas," *Journal of Regional Science*, Vol. 2, pp. 21-36.

Hunt, J. D. (1997), "A Description of the MEPLAN Framework for Land Use and Transport Interaction Modeling," Compendium of Papers of the 77th Annual Meeting of the Transportation Research Board, National Research Council, Washington, D.C.

Johnston, R. A., D. R. Shabazian, and S. Gao (2003), "UPLAN: A Versatile Urban Growth Model for Transportation Planning," Compendium of Papers of the 82nd Annual Meeting of the Transportation Research Board, National Research Council, Washington, D.C.

Kain, J. F. (1986), "Computer Simulation Models of Urban Location," Handbook of Regional and Urban Economics, Vol. 2, E.S. Mills (ed), Amsterdam, North-Holland, pp. 847-875.

Kain, J. F. and W. C. Apgar, Jr. (1985), *Housing and Neighborhood Dynamics: A Simulation Study*, Harvard University Press, Cambridge, MA.

Kaltenbach, K.D. (2003), "FSUTMS: How Did We Get Here?" *Florida Transportation Modeling Newsletter*, Vol. 24, October.

Klosterman, R. E. (1987), "Politics of Computer-Aided Planning," *Town Planning Review*, Vol. 58, pp. 441-452.

Klosterman, R. E. (1999), "The What If? Collaborative Planning Support System," *Environment and Planning B: Planning and Design*, Vol. 26, pp. 393 - 408.

Krishnamurthy, S. and K. Kockelman (2002), *Uncertainty in Integrated Land Use-Transport Models: Simulation and Propagation*, Research Report 167523-1, Center for Transportation Research, University of Texas, Austin, TX.

Landis, J. D. (1994), "The California Urban Futures Model: A New Generation of Metropolitan Simulation Models," *Environment and Planning B*, Vol. 21, pp. 399-420.

LANL (2002), *Transportation Analysis Simulation System (Transims) Portland Study Reports*, Los Alamos National Laboratory, Los Alamos, NM, Available from http://transims.tsasa.lanl.gov/.

Mackett, R. L. (1979), *A model of the Relationships between Transport and Land-Use*, Working Paper 122, Institute for Transport Studies, University of Leeds, Leeds, England.

Mackett, R. L. (1983), *The Leeds Integrated Transport Model (LILT)*, Supplementary Report 805, Transport and Road Research Laboratory, University of Leeds, Crowthorne, England.

Mackett, R.L. (1990a), "The Systematic Application of the LILT Model to Dortmund, Leeds and Tokyo," *Transportation Reviews 10*, pp. 323-338.

Mackett, R. L. (1990b), "Comparative Analysis of Modeling Land-Use Transport Interaction at the Micro and Macro Levels," *Environment and Planning*, Vol. 22 A, pp 459-475.

Maffii, S. and A. Martino (1999), "The Integrated Land-Use-Transport Model of Naples: From the Master Transport Plan to the EU-Policies," *Proceedings of the 6th International Conference of the Computers in Urban Planning and Urban Management*, Venice, Italy.

Mann, W. (1995), "Land Use/Transportation Integrated Model," Paper No. 950158, Compendium of Papers, the 74th Annual Meeting of the Transportation Research Board, National Research Council, Washington, D.C.

Miller, E., D. Kriger, and J. D. Hunt (1998), "Integrated Urban Models for Simulation of Transit and Land Use Policies," TCRP Web Document 9, Final Report, TCRP Project H-12, Transportation Research Board, National Research Council, Washington, D.C.

PBSJ (1999), *Land Use Impacts of Transportation: A Guidebook*, Technical Report 423A, NCHRP 8-32, prepared by Parsons Brinckerhoff, Quade and Douglas Inc., prepared for Transportation Research Board, National Research Council, Washington, D.C.

Prastacos, P. (1985), "Urban Development Models for the San Francisco Region: From Plum to Polis," *Transportation Research Record 1046*, Transportation Research Board, National Research Council, Washington, D.C., pp. 37-44.

Putman, S. H. (1983), *Integrated Urban Models*, Pion Limited, London. England.

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

Southworth, F. (1995), *A Technical Review of Urban Land Use-Transportation Models as Tools for Evaluating Vehicle Travel Reduction Strategies*. Technical Report ORNL-6881, U.S. Department of Energy, Washington, D.C.

Takeyama M. (1996), *Geo-Algebra: A Mathematical Approach to Integrating Spatial Modeling and GIS*, PhD Thesis, Department of Geography, University of California, Santa Barbara, CA.

Timmermans, H. (2003), "The Saga of Integrated Land Use-Transport Modeling: How Many More Dreams Before We Wake Up?" *Proceedings of the 10th International Conference on Travel Behavior Research*, Lucerne, Switzerland, August 2003.

U.S. DOT (2002), *Enhancement of DVRPC's Travel Simulation Models*, U.S. DOT, Washington, D.C., Available from http://tmip.fhwa.dot.gov/clearinghouse/docs/landuse/compendium/dvrpc_toc.stm.

U.S. EPA (2000), *Projecting Land-Use Change: A Summary of Models for Assessing the Effects of Community Growth and Change on Land-Use Patterns*, EPA/600/R-00/098, U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, OH.

ULAM (2004), *Urban Land Use Allocation Model Description*, Available from http//www.ulam.org/.

UrbanSim (2001), *UrbanSim Description*, Available from http://www.urbansim.org/.

Waddell, P and G. F. Ulfarsson (2004), "Introduction to Urban Simulation: Design and Development of Operational Models," *Handbook of Transport*, *Volume 5: Transport Geography and Spatial Systems*, P. Stopher, K. Button, K. Haynes, and D. Hensher, eds., Pergammon Press, New York, NY, Preprint available from http://www.urbansim.org/papers.

Walton, R., M. Corlett, C. Arthur, and A. Bagley, *Land Use and Transportation Modeling with the ArcView Spatial Analyst: The MAG Subarea Allocation Model*, Available from http://www.mag.maricopa.gov/archive/Newpages/on-line.htm.

Webster, F. V., P. H. Bly, and N. Paulley (1988), *Urban Land-Use and Transport Interaction: Policies and Models*, Report on the International Study on Land Use and Transport Interaction (ISGLUTI), Avebury, Aldershot.

Wegener, M. (1982), "Modeling Urban Decline: a Multilevel Economic-Demographic Model of the Dortmund region," *International Regional Science Review*, Vol. 7, pp. 21-41.

Wegener, M. (1983), *Description of the Dortmund Region Model*, Working Paper 8, Institut für Raumplanung, Dortmund, Germany.

Wegener, M. (1985), "The Dortmund Housing Market Model: a Monte Carlo Simulation of a Regional Housing Market," *Microeconomic Models of Housing Markets*, Stahl, K. ed., Lecture

Notes in Economics and Mathematical Systems 239, Berlin/Heidelberg/New York, Springer Verlag, pp. 144-191.

Wegener, M. (1995), "Current and Future Land Use Models," *Proceedings of the Land Use Modeling Conference*, Travel Model Improvement Program, U.S. Department of Transportation, Washington, D.C., pp. 13-40.

Wegener, M. (1998), *The IRPUD Model: Overview*, Available from http://irpud.raumplanung.uni-dortmund.de/irpud/pro/mod/mod_e.htm.

Wegener, M., C. Schürmann, and K. Spiekermann (2001), "From Macro to Micro: Evolution of an Urban Simulation Model," *7th International Conference on Computers in Urban Planning and Urban Management*, Honolulu, Hawaii, July 2001.

APPENDICES

Appendix I. AML Code to Create Grid Cells

```
&args routine
&if [null %routine%] &then
  &return &inform &run uagrid <routine(grid to process)>

/* set path variables...
&s data e:/workspace/cntygrid
&s home [show workspace]

&echo &on
&watch %routine%.watch
&call %routine%
&echo &off
&watch &off
&return
/* --------------------------------------------
&routine idgrid
&workspace %data%
grid
setcell 492.1245
setwindow %data% %data%
val = scalar(0)
docell
 if(%data% gt 0)
  begin
   val := val + 1
   id150 = val
  end
end
quit /* grid

&return
```

Appendix II. Land Use Lookup Table

Parcel Code	DESCRIPTION	LANDUSE1	LANDUSE2
00	Residential Vacant Land	Vacant	NR
01	Residential Single Family	Single Family Residential	R
02	Residential Mobile Home	Mobile Home	R
03	Multi-Family > 5 Units	Multi-Family Residential	R
04	Condominiums/Timeshare	Multi-Family Residential	R
05	Residential Co-operatives	Multi-Family Residential	R
06	Retirement Homes	Group Quarters	GQ
07	Multi-Family Common Areas	Multi-Family Residential	R
08	Multi-Family Less Than 5 Units	Multi-Family Residential	R
09	Interval Ownership Timeshares	Multi-Family Residential	R
10	Commerical Vacant Land	Commerical	C
11	One-Story Store	Commerical	C
12	Store/Office/SFR	Commerical	C
13	Department Store	Commerical	C
14	Supermarket	Commerical	C
15	Shopping Center Regional	Commerical	C
16	Local Shopping Center	Commerical	C
17	One Story Office	Office	C
18	Multi-Story Office	Office	C
19	Professional Office	Office	C
20	Airport	Transportation Communication Utility	NR
21	Restaurant	Commerical	C
22	Drive-in Restaurants	Commerical	C
23	Financial Institutions	Office	C
24	Mobile Home Park	Mobile Home	R
25	Service Shop	Commercial	C
26	Service Station	Commercial	C
27	Auto Sales Repair, etc.	Commercial	C
28	Parking Lots	Parking	NR
29	Wholesale Outlet	Commercial	C
30	Florist, Greenhouses	Commercial	C
31	Drive-in Theaters - Open	Commercial	C
32	Enclosed Theaters, Auditorium	Commercial	C
33	Nightclubs, Lounges, Bars	Commercial	C
34	Bowling Alleys	Recreation	C
35	Tourist Attractions	Commercial	C
36	Camps, Campgrounds	Recreation	NR
37	Race Tracks - Horse / Auto / Dog	Recreation	NR
38	Golf Courses	Recreation	C
39	Hotels / Motels	Commercial	C

40	Industrial Vacant Land	Vacant	NR
41	Light Manufacturing	Industrial	I
42	Heavy Industrial	Industrial	I
43	Lumber Yards	Industrial	I
44	Packing Plants	Industrial	I
45	Breweries, Wineries, etc.	Industrial	I
46	Food Processing	Industrial	I
47	Mineral Processing	Industrial	I
48	Warehousing	Warehousing	C
49	Open Storage	Warehousing	C
50	AG Improved, Rural Homesite	Single Family Residential	R
51	AG Cropland, Feed Grains	Agriculture	C
52	AG Cropland, Vegetables	Agriculture	C
53	AG Cropland, Special	Agriculture	C
54	AG Timberland # 1	Agriculture	NR
55	AG Timberland # 2	Agriculture	NR
56	AG Timberland # 3	Agriculture	NR
57	AG Timberland # 4	Agriculture	NR
58	AG Timberland # 5	Agriculture	NR
59	Wasteland - AG Timberland Not Classified	Vacant	NR
60	AG Pastures, Class 1	Agriculture	NR
61	AG Pastures, Improved, Class 2	Agriculture	NR
62	AG Pastures, Semi-Improved, Class 3	Agriculture	NR
63	AG Pastures, Native, Class 4	Agriculture	NR
64	AG Pastures, Native, Class 5	Agriculture	NR
65	AG Pastures, Native, Class 6	Agriculture	NR
66	AG Citrus Groves, Orchards, etc	Agriculture	C
67	AG Poultry Farms, Bees, Fish, etc.	Agriculture	C
68	AG Daries, Feed Lots	Agriculture	C
69	AG Nurseries, Ornamentals, etc.	Commercial	C
70	Institutional, Vacant Land	Vacant	NR
71	Institutional, Churches	Civic and Quasi-Public	G
72	Institutional, Private Schools	School	C
73	Institutional, Hospitals, Private	Hospital, Convalescent Center	G
74	Homes for the Aged	Group Quarters	GQ
75	Orphanages	Group Quarters	GQ
76	Mortuaries, Cemetaries, etc.	Civic and Quasi-Public	G
77	Clubs, Lodges, Halls	Commercial	C
78	Sanitariums, Convalescent, etc.	Hospital, Convalescent Center	GQ
79	Cultural Facilities	Civic and Quasi-Public	G
80	Undefined	No Land Use Code	
81	Military	Military	G

82	Forest, Parks, etc.	Forest	NR
83	Schools, Public	School	G
84	Colleges	School	G
85	Hospitals	Hospital, Convalescent Center	G
86	Other County	Government	G
87	Other State	Government	G
88	Other Federal	Government	G
89	Other Municipal	Government	G
90	Leasehold Interests	Recreation	NR
91	Utilities	Transportation, Communication, Utilities	NR
92	Mining & Production of Petroleum & Gas	Mining	I
93	Subsurface Rights	Right-Of-Way	NR
94	Right of Way, Street Roads, Ditch, etc	Right-Of-Way	NR
95	Rivers, Lakes, Submerged Lands	Water	NR
96	Sewage, Solid Waste, Borrow Pit	Transportation, Communication, Utilities	NR
97	Outdoor Recreational or Park, Classifie	Recreation	G
98	Centrally Assessed	Government	G
99	Acreage Not Zoned Commercial	Vacant	NR

Appendix III. Input Files to UrbanSim

annual_employment_control_totals

SECTOR_ID	INTEGER	Index into the employment_sectors table
YEAR	INTEGER	
TOTAL_HOME_BASED_EMPLOYMENT	INTEGER	*(optional)* Target employment in this sector.
TOTAL_NON_HOME_BASED_EMPLOYMENT	INTEGER	*(optional)* for this year
TOTAL_EMPLOYMENT	INTEGER	*(optional)* (see below)

Notes:

- YEAR must be between ABSOLUTE_MIN_YEAR and ABSOLUTE_MAX_YEAR.
- SECTOR_ID must be a valid index into the employment_sectors table.
- TOTAL_EMPLOYMENT must be greater than or equal to zero.
- A control total must be provided for each sector in employment_sectors for every year in the scenario, including the start and end years.
- If the TOTAL_HOME_BASED_EMPLOYMENT and TOTAL_NON_HOME_BASED_EMPLOYMENT columns are present, the TOTAL_EMPLOYMENT must not be present. And if the TOTAL_EMPLOYMENT column is present, the TOTAL_HOME_BASED_EMPLOYMENT and TOTAL_NON_HOME_BASED_EMPLOYMENT columns must not be present.

annual_relocation_rates_for_jobs

SECTOR_ID	INTEGER	Index into the EMPLOYMENT_SECTORS table
JOB_RELOCATION_PROBABILITY	FLOAT	Probability that a job in this sector will relocate within the time span of one year

Notes:

- There must be a single entry for every employment sector in the employment_sectors table.
- JOB_RELOCATION_PROBABILITY must be between 0 and 1, inclusive.

employment_sectors

SECTOR_ID	INTEGER	Unique identifier
NAME	VARCHAR	Unique name of the Sector

Notes:

- SECTOR_ID must be unique, greater than zero, and less than or equal to 99.
- It's recommended that SECTOR_IDs do not exceed 1000 for efficiency reasons.
- NAME must be unique. We recommend that NAMEs follow the style guide.

employment_adhoc_sector_groups

GROUP_ID	INTEGER	Unique identifier
NAME	VARCHAR	Unique name of the Group

Notes:

- GROUP_ID must be unique and greater than zero

- NAME must be unique. The required employment ad hoc sector groups must be lower case with underscores between words, e.g. lower_case_with_underscores_between_words. We recommend that all NAMEs follow this style.

employment_adhoc_sector_group_definitions

SECTOR_ID	INTEGER	Index into the employment_sectors table
GROUP_ID	INTEGER	Index into the employment_adhoc_sector_groups table

Notes:

- SECTOR_ID must be a valid index into the employment_sectors table
- GROUP_ID must be a valid index into the employment_adhoc_sector_groups table
- The combination of SECTOR_ID+GROUP_ID must be unique

jobs

JOB_ID	INTEGER	Unique identifier
GRID_ID	INTEGER	Grid cell this job exists in; zero if currently not assigned to a grid cell
HOME_BASED	BOOLEAN	True if home-based
SECTOR_ID	INTEGER	Sector this job belongs to

Notes:

- GRID_ID must be a valid id in the gridcells table (may not be zero, since all baseyear jobs must be placed)
- JOB_ID must be unique and greater than zero
- SECTOR_ID must be a valid id in the employment_sectors table
- If HOME_BASED = false, then this job's GRID_ID must correspond to a non-vacant cell whose DEVELOPMENT_TYPE_ID has a positive entry in the sqft_for_non_home_based_jobs table.

jobs_for_estimation, jobs_for_estimation_home_based, jobs_for_estimation_non_home_based

The schema and structure of this table is identical to the basic jobs table.

residential_units_for_home_based_jobs

DEVELOPMENT_TYPE_ID	INTEGER	Unique identifier
RATIO	FLOAT	The ratio of residential units that can accommodate a home-based job of this development type to the total number of residential units

Notes:

- There must be one row for each DEVELOPMENT_TYPE_ID.
- RATIO must be between 0 and 1 (0 means that we're not allowing any home-based jobs in this development type; 1 means that every unit can accommodate a job)

sqft_for_non_home_based_jobs

DEVELOPMENT_TYPE_ID	INTEGER	Unique identifier
SQFT	FLOAT	Number of square feet needed to place one job in a grid cell with this development type. A SQFT of -1 means that no non-home-based jobs can be placed in a grid cell with this development type.

Notes:

- There must be one row for each DEVELOPMENT_TYPE_ID.
- SQFT must be greater than zero or else -1.
- SQFT is usually between 200 and 2000.

annual_household_control_totals

YEAR	INTEGER	Year for the total
AGE_OF_HEAD	INTEGER	*(optional)* Household characteristic bin number of age of head of household
CARS	INTEGER	*(optional)* Household characteristic bin number of number of cars in household
CHILDREN	INTEGER	*(optional)* Household characteristic bin number of number of children in household
INCOME	INTEGER	*(optional)* Household characteristic bin number of household income
PERSONS	INTEGER	*(optional)* Household characteristic bin number of size of household in number of people
RACE_ID	INTEGER	*(optional)* Household characteristic bin number of race of head of household
WORKERS	INTEGER	*(optional)* Household characteristic bin number of employed people in household
TOTAL_NUMBER_OF_HOUSEHOLDS	INTEGER	Target number of households of this household type and year

Notes:

- TOTAL_NUMBER_OF_HOUSEHOLDS must be greater than or equal to zero.
- For each year in the scenario, including the start and end years, the entries should be complete.
- Complete entries for a year mean that there is a row for that year for the cross product of all specified bins.
- If a year is not complete, the household types that are not specified will not be modified.
- YEAR must be between ABSOLUTE_MIN_YEAR and ABSOLUTE_MAX_YEAR

annual_relocation_rates_for_households

AGE_MIN	INTEGER	The minimum age for which this probability is valid.
AGE_MAX	INTEGER	The maximum age for which this probability is valid, -1 means no maximum
INCOME_MIN	INTEGER	The minimum income for which this probability is valid.
INCOME_MAX	INTEGER	The maximum income for which this probability is valid, -1 means no maximum
PROBABILITY_OF_RELOCATING	FLOAT	The probability of relocating in a year.

Notes:

- AGE_MIN must be >= 0
- AGE_MAX must be > AGE_MIN or else -1
- INCOME_MIN must be >= 0 and must be a multiple of 1000.
- INCOME_MAX must be > INCOME_MIN and a multiple of 1000 -1 (e.g. 200,999) or else -1
- PROBABILITY_OF_RELOCATING must be >= 0.0 and <= 1.0
- The ranges must be disjoint and cover the entire space (from zero to infinity in the two-dimensional space produced by age and income).

households

HOUSEHOLD_ID	INTEGER	Unique identifier
GRID_ID	INTEGER	Grid cell this household resides in; zero if currently not residing in a housing unit
PERSONS	INTEGER	Total number of people living in this household.
WORKERS	INTEGER	Total number of workers living in this household.
AGE_OF_HEAD	INTEGER	Age of head of the household
INCOME	INTEGER	Income of this household
CHILDREN	INTEGER	Number of children living in this household
RACE_ID	INTEGER	Race of head of household
CARS	INTEGER	Number of cars in this household

Notes:

- HOUSEHOLD_ID must be unique and greater than zero
- GRID_ID must be a valid id in the gridcells table or zero
- PERSONS must be greater than zero and less than or equal to ABSOLUTE_MAX_HOUSEHOLD_SIZE
- WORKERS must be between zero and ABSOLUTE_MAX_HOUSEHOLD_SIZE
- AGE_OF_HEAD greater than zero and less than or equal to ABSOLUTE_MAX_PERSON_AGE
- INCOME must be greater than zero and less than or equal to ABSOLUTE_MAX_INCOME
- CHILDREN must be greater than or equal to zero and less than or equal to ABSOLUTE_MAX_HOUSEHOLD_SIZE
- RACE_ID must be a valid id in the race_names table
- CARS must be greater than or equal to zero and less than or equal to ABSOLUTE_MAX_HOUSEHOLD_SIZE

- The total number of households in a single grid cell should be no greater than that cell's residential units.

households_for_estimation

The schema and structure of this table is identical to the basic households table.

household_characteristics_for_ht

CHARACTERISTIC	VARCHAR	See above for valid values
MIN	INTEGER	Minimum value for this bin for this characteristic. Values are placed in a bin if MIN <= value <= MAX
MAX	INTEGER	Maximum value for this bin for this characteristic; -1 means infinity / no maximum

Notes:

- MIN must be greater than or equal to zero
- MAX must be greater than MIN or else -1
- Bins for each CHARACTERISTIC may not overlap.
- Bins for each CHARACTERISTIC must cover all values from zero to infinity
- At least one bin for each of the five required CHARACTERISTICS. Preferably, bins for all seven CHARACTERISTICS, but for backward compatibility, Persons and Race have defaults if they are not specified.

household_characteristics_for_hlc

CHARACTERISTIC	VARCHAR	The valid values are: "Income", "Age of Head", "Children", "Workers", "Cars"
MIN	INTEGER	Minimum value for this bin for this characteristic. Values are placed in a bin if MIN <= value <= MAX
MAX	INTEGER	Maximum value for this bin for this characteristic; -1 means infinity / no maximum

Notes:

- MIN must be greater than or equal to zero
- MAX must be greater than MIN or else -1
- Bins for each CHARACTERISTIC may not overlap.
- Bins for each CHARACTERISTIC must cover all values from zero to infinity
- At least one bin for each of the three CHARACTERISTICS
- The table MUST include the bin 0..0 for the "Workers" CHARACTERISTIC, so can distinguish between "employed" and "unemployed" households

race_names

RACE_ID	INTEGER	Unique identifier
NAME	VARCHAR	Name of the race
MINORITY	BOOLEAN	True if the race is a minority

Notes:

- RACE_ID must be unique and greater than zero
- There must be at least one non-minority group listed

base_year

YEAR	INTEGER	Year of base data

Notes:
- YEAR must be >= 1900.

cities

CITY_ID	INTEGER	Unique identifier
CITY_NAME	VARCHAR	

Notes:
- CITY_ID must be unique and greater than zero.

counties

COUNTY_ID	INTEGER	Unique identifier
COUNTY_NAME	VARCHAR	

Notes:
- COUNTY_ID must be unique and greater than zero.

models

JAVA_CLASS	VARCHAR	The Java class implementing this model, e.g., 'org.urbansim.models.AccessibilityModel'
RUN_ORDER	INTEGER	The order in which to run this model. Lower numbered models are run before higher numbered models, e.g. -1 is run before 0 which is run before 1.

Notes:
- JAVA_CLASS must extend the Model class.
- RUN_ORDER must be unique.

sampling_rates

HLC_SAMPLING_RATE	FLOAT	Gridcell sampling rate for the Household Location Choice Model
ELC_SAMPLING_RATE	FLOAT	Gridcell sampling rate for the Employment Location Choice Model

Notes:
- The table must have exactly one row
- The rates must be between 0 and 1, inclusive.

scenario_information

END_YEAR	INTEGER	The scenario runs from the base_year of the data up to, and including, this end year.
DESCRIPTION	VARCHAR	*(optional)* Human readable description
PARENT_DATABASE_URL	VARCHAR	The next database in the chain of scenario databases.
CONTINUATION	BOOLEAN	True if this scenario is the continuation of a larger group of scenarios and thus the output database should be appended to; false if this scenario is a new scenario and thus the output database should be cleared before use.

Notes:

- PARENT_DATABASE_URL must be empty in the baseyeardata database.

urbansim_constants

CELL_SIZE	FLOAT	Width and height of each grid cell in UNITS
UNITS	VARCHAR	Units of measurement, eg. "meters" or "feet"
WALKING_DISTANCE_CIRCLE_RADIUS	FLOAT	Walking distance in meters, e.g., 600 m, R in the descriptions
YOUNG_AGE	INTEGER	Max age for a person to be considered young
PROPERTY_VALUE_TO_ANNUAL_COST_RATIO	FLOAT	Ratio of the total property value to an annual rent for that property
LOGIT_CHOICE_SET_SIZE_FOR_ESTIMATION	INTEGER	The size of the choice sets used to output estimation data tables for the household and employment location choice models. Each model describes how this choice set is selected in its estimation data table specification.
NUMBER_OF_DEVELOPER_MODEL_HISTORY_YEARS	INTEGER	The number of years the user wants the developer model estimation data writer to roll back and use in the estimation data.
DEVELOPER_MODEL_ESTIMATION_THRESHOLD_COUNT	INTEGER	The number of times a transition from starting development type to ending development type must occur in development_event_history events (for the NUMBER_OF_DEVELOPER_MODEL_HISTORY_YEARS years before the baseyear) in order for those events to be used by the developer model estimation writer when constructing choice sets and when writing estimation data. This value is used to restrict the developer model estimation data writer to only use "frequent" development events. Note that all historical development events, regardless of their frequency, are used for the roll back process of the developer model estimation data writer.
LOW_INCOME_FRACTION	FLOAT	Fraction of the total number of households considered to have low incomes, e.g., 0.1
MID_INCOME_FRACTION	FLOAT	Fraction of the total number of households considered to have mid-level incomes, e.g., 0.5
NEAR_ARTERIAL_THRESHOLD	FLOAT	Line distance from the centroid of a cell to an arterial for it to be considered nearby, e.g., 300. NearArterialThreshold in description.
NEAR_HIGHWAY_THRESHOLD	FLOAT	Line distance from the centroid to a highway for it to be considered nearby, e.g., 300. NearHighwayThreshold in description.
PERCENT_COVERAGE_THRESHOLD	INTEGER	The threshold above which a grid cell's percent_*, e.g. percent_wetland, must be to be considered "covered" for that attribute. So, if

		percent_coverage_threshhold is 50 percent and percent_wetland is 60 percent, the grid cell would be considered "covered" by wetland.
RECENT_YEARS	INTEGER	Maximum number of years to look back when considering recent transitions. For example, if RECENT_YEARS = 3, then the value COMMERICIAL_SQFT_RECENTLY_ADDED in the gridcells table would refer to the number of square feet of commercial space built in the last 3 years.
MAX_PERSONS_PER_HOUSEHOLD_FOR_CONTROL_TOTALS	INTEGER	Maximum number of persons listed in the annual_household_control_totals table. Households with more than this number of PERSONs will map to this number as described for the annual_household_control_totals table.

Notes:

- CELL_SIZE must be greater than zero
- UNITS must be one of the following: meters, feet, miles, kilometers.
- WALKING_DISTANCE_CIRCLE_RADIUS must be greater than zero and <= ABSOLUTE_MAX_DISTANCE
- WALKING_DISTANCE_CIRCLE_RADIUS must result in less than 250 grid cells within walking distance
- YOUNG_AGE must be greater than zero and <= ABSOLUTE_MAX_PERSON_AGE
- PROPERTY_VALUE_TO_ANNUAL_COST_RATIO must be greater than zero
- LOGIT_CHOICE_SET_SIZE_FOR_ESTIMATION must be greater than 1.
- DEVELOPER_MODEL_ESTIMATION_THRESHOLD_COUNT must be greater than 0.
- LOW_INCOME_FRACTION must be between 0 and 1
- MID_INCOME_FRACTION must be between 0 and 1
- MID_INCOME_FRACTION + LOW_INCOME_FRACTION must be at most 1.
- NEAR_ARTERIAL_THRESHOLD must be greater than zero and <= ABSOLUTE_MAX_DISTANCE
- NEAR_HIGHWAY_THRESHOLD must be greater than zero and <= ABSOLUTE_MAX_DISTANCE
- PERCENT_COVERAGE_THRESHOLD must be between zero and 100. Note that this value is exclusive; for example, if the value is set to 45 and a grid cell is 45% covered by roads, the cell will not be considered to be "covered" by roads.
- RECENT_YEARS must be > 0 and <= ABSOLUTE_MAX_RECENT_YEARS
- MAX_PERSONS_PER_HOUSEHOLD_FOR_CONTROL_TOTALS must be > 0 and <= ABSOLUTE_MAX_HOUSEHOLD_SIZE

model_variables

VARIABLE_NAME	VARCHAR	Unique name of a model variable.
DEFINITION	VARCHAR	*(optional)* Optional text definition of the variable (for documentation only). Do not use quotes characters or special characters.
DESCRIPTION	VARCHAR	*(optional)* Optional text description of the variable (for documentation only). Do not use quotes characters or special characters.
SHORT_NAME	VARCHAR	A name of at most six characters for the variable, to help users who want to export data to programs with name length limitations.
JAVA_CLASS	VARCHAR	Name of the Java class implementing this variable
LOG_THIS_VALUE	BOOLEAN	If true, the model variable logger model will write this variable value to the output database.

Notes:

- VARIABLE_NAME must be unique and non-empty.
- SHORT_NAME must be unique and non-empty.
- JAVA_CLASS must implement either org.urbansim.models.term.Term or org.urbansim.models.term.TermFactory.
- If JAVA_CLASS implements org.urbansim.models.term.TermFactory, VARIABLE_NAME must have a single "?" in it.
- If JAVA_CLASS implements org.urbansim.models.term.TermFactory, SHORT_NAME must have one or more consecutive "?" in it. Each "?" is replaced by one upper case character from the replacement string. If the replacement string is too short, it is left padded with zeros. Thus if the replacement string is "commercial", a single "?" would be replace by "C" whereas two "??"s would be replaced by "CO".
- DEFINITION and DESCRIPTION must not include special characters or quotes.

developer_model_coefficients

SUB_MODEL_ID	INTEGER	Identifier for a submodel, if used by the model, or -2 if not used by this model. See Developer model and Employment Location Choice model for more details.
COEFFICIENT_NAME	VARCHAR	Unique name of a coefficient
ESTIMATE	DOUBLE	The estimated value of this coefficient
STANDARD_ERROR	DOUBLE	*(optional)* The standard error of this estimated value. This is for reference only and is not used by UrbanSim.
T_STATISTIC	DOUBLE	*(optional)* The t-statistic of this coefficient for the test of significance from 0. This is for reference only and is not used by UrbanSim.
P_VALUE	DOUBLE	*(optional)* The p-value of this t-statistic, gives the Prob(|x|>|estimated coefficient|) when x is drawn from a t-distribution with mean 0. This is for reference only and is not used by UrbanSim.

Notes:

- SUB_MODEL_ID must be in the appropriate cross-reference table for each model.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must be unique.

employment_location_choice_model_coefficients

SUB_MODEL_ID	INTEGER	
COEFFICIENT_NAME	VARCHAR	
ESTIMATE	DOUBLE	
STANDARD_ERROR	DOUBLE	*(optional)*
T_STATISTIC	DOUBLE	*(optional)*
P_VALUE	DOUBLE	*(optional)*

Notes:

- SUB_MODEL_ID must be in the appropriate cross-reference table for each model.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must be unique.

employment_non_home_based_location_choice_model_coefficients

SUB_MODEL_ID	INTEGER	
COEFFICIENT_NAME	VARCHAR	
ESTIMATE	DOUBLE	
STANDARD_ERROR	DOUBLE	*(optional)*
T_STATISTIC	DOUBLE	*(optional)*
P_VALUE	DOUBLE	*(optional)*

Notes:

- SUB_MODEL_ID must be in the appropriate cross-reference table for each model.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must be unique.

employment_home_based_location_choice_model_coefficients

SUB_MODEL_ID	INTEGER	
COEFFICIENT_NAME	VARCHAR	
ESTIMATE	DOUBLE	
STANDARD_ERROR	DOUBLE	*(optional)*
T_STATISTIC	DOUBLE	*(optional)*
P_VALUE	DOUBLE	*(optional)*

Notes:

- SUB_MODEL_ID must be in the appropriate cross-reference table for each model.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must be unique.

household_location_choice_model_coefficients

SUB_MODEL_ID	INTEGER	
COEFFICIENT_NAME	VARCHAR	
ESTIMATE	DOUBLE	
STANDARD_ERROR	DOUBLE	*(optional)*
T_STATISTIC	DOUBLE	*(optional)*
P_VALUE	DOUBLE	*(optional)*

Notes:

- SUB_MODEL_ID must be in the appropriate cross-reference table for each model.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must be unique.

land_price_model_coefficients

SUB_MODEL_ID	INTEGER	
COEFFICIENT_NAME	VARCHAR	
ESTIMATE	DOUBLE	
STANDARD_ERROR	DOUBLE	*(optional)*
T_STATISTIC	DOUBLE	*(optional)*
P_VALUE	DOUBLE	*(optional)*

Notes:

- SUB_MODEL_ID must either be -2 for all entries, or be an appropriate cross-reference to the development_types table.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must be unique.

residential_land_share_model_coefficients

SUB_MODEL_ID	INTEGER	
COEFFICIENT_NAME	VARCHAR	
ESTIMATE	DOUBLE	
STANDARD_ERROR	DOUBLE	*(optional)*
T_STATISTIC	DOUBLE	*(optional)*
P_VALUE	DOUBLE	*(optional)*

Notes:

- SUB_MODEL_ID must be in the appropriate cross-reference table for each model.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must be unique.

developer_model_specification

SUB_MODEL_ID	INTEGER	The starting development_type for a gridcell.
EQUATION_ID	INTEGER	The ending development_type for a gridcell.
COEFFICIENT_NAME	VARCHAR	The name of the model coefficient on this variable (in this equation in this sub-model)
VARIABLE_NAME	VARCHAR	The name of a model variable (in this equation in this sub-model)

Notes:

- SUB_MODEL_ID must be a valid development_type.
- EQUATION_ID must be a valid development_type.

- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must exist in the Model Coefficients table.
- VARIABLE_NAME must exist in model_variables table.
- Each comibination of (SUB_MODEL_ID, EQUATION_ID, VARIABLE_NAME) must be unique.
- Specific models have additional requirements: see Developer Model and Employment Location Choice Model.

employment_location_choice_model_specification

SUB_MODEL_ID	INTEGER	Foreign key into the employment_sectors table.
EQUATION_ID	INTEGER	Each sub-model does not have multiple equations, so use '-2' in this column
COEFFICIENT_NAME	VARCHAR	
VARIABLE_NAME	VARCHAR	

Notes:
- SUB_MODEL_ID must exist in the employment_sectors table.
- EQUATION_ID must be -2.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must exist in the Model Coefficients table.
- VARIABLE_NAME must exist in model_variables table.
- Each comibination of (SUB_MODEL_ID, EQUATION_ID, VARIABLE_NAME) must be unique.
- Specific models have additional requirements: see Developer Model and Employment Location Choice Model.

employment_non_home_based_location_choice_model_specification

SUB_MODEL_ID	INTEGER	Foreign key into the employment_sectors table.
EQUATION_ID	INTEGER	
COEFFICIENT_NAME	VARCHAR	Each sub-model does not have multiple equations, so use '-2' in this column
VARIABLE_NAME	VARCHAR	

Notes:
- SUB_MODEL_ID must exist in the employment_sectors table.
- EQUATION_ID must be -2.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must exist in the Model Coefficients table.
- VARIABLE_NAME must exist in model_variables table.
- Each comibination of (SUB_MODEL_ID, EQUATION_ID, VARIABLE_NAME) must be unique.
- Specific models have additional requirements: see Developer Model and Non-home-based Employment Location Choice Model.

employment_home_based_location_choice_model_specification

SUB_MODEL_ID	INTEGER	Foreign key into the employment_sectors table.
EQUATION_ID	INTEGER	
COEFFICIENT_NAME	VARCHAR	Each sub-model does not have multiple equations, so use '-2' in this column
VARIABLE_NAME	VARCHAR	

Notes:

- SUB_MODEL_ID must exist in the employment_sectors table.
- EQUATION_ID must be -2.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must exist in the Model Coefficients table.
- VARIABLE_NAME must exist in model_variables table.
- Each comibination of (SUB_MODEL_ID, EQUATION_ID, VARIABLE_NAME) must be unique.
- Specific models have additional requirements: see Developer Model and Employment Home-Based Location Choice Model.

household_location_choice_model_specification

SUB_MODEL_ID	INTEGER	This model does not have multiple sub-models, so use '-2' for every row.
EQUATION_ID	INTEGER	This model does not have multiple equations per sub-model, so use '-2' for every row.
COEFFICIENT_NAME	VARCHAR	
VARIABLE_NAME	VARCHAR	

Notes:

- SUB_MODEL_ID must be -2.
- EQUATION_ID must be -2.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must exist in the Model Coefficients table.
- VARIABLE_NAME must exist in model_variables table.
- Each comibination of (SUB_MODEL_ID, EQUATION_ID, VARIABLE_NAME) must be unique.

land_price_model_specification

SUB_MODEL_ID	INTEGER	
EQUATION_ID	INTEGER	
COEFFICIENT_NAME	VARCHAR	
VARIABLE_NAME	VARCHAR	

Notes:

- SUB_MODEL_ID must be either be -2 (if the land price model is not using sub models) or a DEVELOPMENT_TYPE_ID value from the development_types table (if the land price model is using sub models).
- EQUATION_ID must be -2, since the land price model does not use separate equations.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must exist in the Model Coefficients table.

- VARIABLE_NAME must exist in model_variables table.
- Each comibination of (SUB_MODEL_ID, EQUATION_ID, VARIABLE_NAME) must be unique.

residential_land_share_model_specification

SUB_MODEL_ID	INTEGER	This model does not have multiple sub-models, so use '-2' for every row.
EQUATION_ID	INTEGER	This model does not have multiple equations per sub-model, so use '-2' for every row.
COEFFICIENT_NAME	VARCHAR	
VARIABLE_NAME	VARCHAR	

Notes:

- SUB_MODEL_ID must be -2.
- EQUATION_ID must be -2.
- Each combination of (SUB_MODEL_ID, COEFFICIENT_NAME) must exist in the Model Coefficients table.
- VARIABLE_NAME must exist in model_variables table.
- Each comibination of (SUB_MODEL_ID, EQUATION_ID, VARIABLE_NAME) must be unique.

development_constraints

CONSTRAINT_ID	INTEGER	Unique rule identification number
CITY_ID	INTEGER	Value or *don't care*
COUNTY_ID	INTEGER	Value or *don't care*
IS_IN_WETLAND	INTEGER	Boolean integer value or *don't care*
IS_OUTSIDE_URBAN_GROWTH_BOUNDARY	INTEGER	Boolean integer value or *don't care*
IS_IN_STREAM_BUFFER	INTEGER	Boolean integer value or *don't care*
IS_ON_STEEP_SLOPE	INTEGER	Boolean integer value or *don't care*
IS_IN_FLOODPLAIN	INTEGER	Boolean integer value or *don't care*
PLANTYPE_X	VARCHAR	String listing of Plan Type id, e.g. "1,2,3", or *don't care*
DEVTYPE_X	VARCHAR	String listing of Development Types, e.g. "1,2,3", that rule prohibits, or "ALL" to prohibit all development.

Notes:

- CONSTRAINT_ID must be a unique positive integer.
- CITY_ID must be a valid CITY_ID or -1 for *don't care.*
- COUNTY_ID must be a valid COUNTY_ID or -1 for *don't care.*
- IS_IN_WETLAND must be 0, 1, or -1 for *don't care.*
- IS_OUTSIDE_URBAN_GROWTH_BOUNDARY must be 0, 1, or -1 for *don't care.*
- IS_IN_STREAM_BUFFER must be 0, 1, or -1 for *don't care.*
- IS_ON_STEEP_SLOPE must be 0, 1, or -1 for *don't care.*
- IS_IN_FLOODPLAIN must be 0, 1, or -1 for *don't care.*
- PLANTYPE_X must be a comma-delimited list of valid PLAN_TYPE_IDs or -1 for *don't care.*

- DEVTYPE_X must be a comma-delimited list of valid DEVELOPMENT_TYPE_IDs or "ALL".

transition_types

TRANSITION_ID	INTEGER	Unique identifier
INCLUDE_IN_DEVELOPER_MODEL	BOOLEAN	True if this transition type is considered a viable transition in the Developer Model. Some transition types may never be used in the Developer Model and some may be turned on and off as needed.
STARTING_DEVELOPMENT_TYPE_ID	INTEGER	Index into the DevelopmentType table
ENDING_DEVELOPMENT_TYPE_ID	INTEGER	Index into the DevelopmentType table
HOUSING_UNITS_MEAN	FLOAT	The average number of new houses to build during this transition.
HOUSING_UNITS_STANDARD_DEVIATION	FLOAT	The standard deviation in the above.
HOUSING_UNITS_MIN	INTEGER	The minimum number of new houses to build during this transition.
HOUSING_UNITS_MAX	INTEGER	The maximum number of new houses to build during this transition.
COMMERCIAL_SQFT_MEAN	FLOAT	The average number of new commercial square footage (COMMERCIAL_SQFT) added.
COMMERCIAL_SQFT_STANDARD_DEVIATION	FLOAT	
COMMERCIAL_SQFT_MIN	INTEGER	
COMMERCIAL_SQFT_MAX	INTEGER	
INDUSTRIAL_SQFT_MEAN	FLOAT	The average number of industrial square footage (INDUSTRIAL_SQFT) added.
INDUSTRIAL_SQFT_STANDARD_DEVIATION	FLOAT	
INDUSTRIAL_SQFT_MIN	INTEGER	
INDUSTRIAL_SQFT_MAX	INTEGER	
GOVERNMENTAL_SQFT_MEAN	FLOAT	The average number of governmental square footage (GOVERNMENTAL_SQFT) added.
GOVERNMENTAL_SQFT_STANDARD_DEVIATION	FLOAT	
GOVERNMENTAL_SQFT_MIN	INTEGER	
GOVERNMENTAL_SQFT_MIN	INTEGER	
GOVERNMENTAL_SQFT_MAX	INTEGER	
HOUSING_IMPROVEMENT_VALUE_MEAN	FLOAT	The average increase in the residential improvement value per unit (RESIDENTIAL_IMPROVEMENT_VALUE).
HOUSING_IMPROVEMENT_VALUE_STANDARD_DEVIATION	FLOAT	
HOUSING_IMPROVEMENT_VALUE_MIN	INTEGER	

HOUSING_IMPROVEMENT_VALUE_MAX	INTEGER	
COMMERCIAL_IMPROVEMENT_VALUE_MEAN	FLOAT	
COMMERCIAL_IMPROVEMENT_VALUE_STANDARD_DEVIATION	FLOAT	
COMMERCIAL_IMPROVEMENT_VALUE_MIN	INTEGER	
COMMERCIAL_IMPROVEMENT_VALUE_MAX	INTEGER	
INDUSTRIAL_IMPROVEMENT_VALUE_MEAN	FLOAT	The average increase in the industrial improvement value per sq. ft. (INDUSTRIAL_IMPROVEMENT_VALUE).
INDUSTRIAL_IMPROVEMENT_VALUE_STANDARD_DEVIATION	FLOAT	
INDUSTRIAL_IMPROVEMENT_VALUE_MIN	INTEGER	
INDUSTRIAL_IMPROVEMENT_VALUE_MAX	INTEGER	
GOVERNMENTAL_IMPROVEMENT_VALUE_MEAN	FLOAT	The average increase in the governmental improvement value per sq. ft. (GOVERNMENTAL_IMPROVEMENT_VALUE).
GOVERNMENTAL_IMPROVEMENT_VALUE_STANDARD_DEVIATION	FLOAT	
GOVERNMENTAL_IMPROVEMENT_VALUE_MIN	INTEGER	
GOVERNMENTAL_IMPROVEMENT_VALUE_MAX	INTEGER	
YEARS_TO_BUILD	INTEGER	Number of years it takes to complete this build event. The additional sq. ft. and/or units will only appear at once this many years after the build event was created.

Notes:

- TRANSITION_ID must be unique and greater than zero.
- STARTING_DEVELOPMENT_ID must be a valid index into the DevelopmentTypes table.
- ENDING_DEVELOPMENT_ID must be a valid index into the DevelopmentTypes table.
- There must be at most one entry for each transition type (all possible starting_development_type_id * all possible ending_development_type_id) and some transitions will not have an entry.
- There must be an entry for each transition defined in the developer_model_coefficients table.
- YEARS_TO_BUILD must be a positive integer.
- For each of the distributions in this table, the following order must hold: min <= mean <= max.
- For each of the distributions in this table, the standard deviation must be >= 0.
- If the ENDING_DEVELOPMENT_ID is a development type in the "developed" DevelopmentTypeGroups, at least one of the following must be greater than zero: HOUSING_UNITS_MAX, COMMERCIAL_SQFT_MAX, INDUSTRIAL_SQFT_MAX, GOVERNMENTAL_SQFT_MAX.

- If the ENDING_DEVELOPMENT_ID is a development type that is NOT in the "developed" DevelopmentTypeGroups, all of the following must be zero: HOUSING_UNITS_MAX, COMMERCIAL_SQFT_MAX, INDUSTRIAL_SQFT_MAX, GOVERNMENTAL_SQFT_MAX.

target_vacancies

YEAR	INTEGER	Year of the simulation for which the vacancy targets apply
TARGET_TOTAL_RESIDENTIAL_VACANCY	FLOAT	Ratio of unused residential units to total residential units
TARGET_TOTAL_NON_RESIDENTIAL_VACANCY	FLOAT	Ratio of unused nonresidential sqft to total nonresidential sqft

Notes:

- There must be exactly one row for each year in the scenario, including the base year.
- YEAR must be between the ABSOLUTE_MIN_YEAR and ABSOLUTE_MAX_YEAR values defined in Constants.java.
- TARGET_TOTAL_RESIDENTIAL_VACANCY must be between 0 and 1, inclusive.
- TARGET_TOTAL_NON_RESIDENTIAL_VACANCY must be between 0 and 1, inclusive.

development_constraint_events

CONSTRAINT_ID	INTEGER	Unique rule identification number
CITY_ID	INTEGER	Value or *don't care*
COUNTY_ID	INTEGER	Value or *don't care*
IS_IN_WETLAND	INTEGER	Boolean integer value or *don't care*
IS_OUTSIDE_URBAN_GROWTH_BOUNDARY	INTEGER	Boolean integer value or *don't care*
IS_IN_STREAM_BUFFER	INTEGER	Boolean integer value or *don't care*
IS_ON_STEEP_SLOPE	INTEGER	Boolean integer value or *don't care*
IS_IN_FLOODPLAIN	INTEGER	Boolean integer value or *don't care*
PLANTYPE_X	VARCHAR	String listing of Plan Type id, e.g. "1,2,3", or *don't care*
DEVTYPE_X	VARCHAR	String listing of Development Types, e.g. "1,2,3", that rule prohibits, or "ALL" to prohibit all development.
SCHEDULED_YEAR	SHORT	Year when the event will be implemented
CHANGE_TYPE	VARCHAR	A (for Add) or D (for Delete)

- SCHEDULED_YEAR must be between ABSOLUTE_MIN_YEAR and ABSOLUTE_MAX_YEAR.
- You cannot add a constraint with a CONSTRAINT_ID that already exists in the SCHEDULED_YEAR.
- You may only delete a constraint with a CONSTRAINT_ID that already exists in the SCHEDULED_YEAR.
- For CHANGE_TYPE "D", only SCHEDULED_YEAR and CONSTRAINT_ID are used; all other columns are ignored.

development_events

GRID_ID	INTEGER	Grid cell where the event takes place
SCHEDULED_YEAR	SHORT	Year where the event will be implemented
DEVELOPMENT_TYPE_CHANGE_TYPE	VARCHAR	Type of change for the DEVELOPMENT_TYPE attribute, either N (no change), or R (replace)
STARTING_DEVELOPMENT_TYPE_ID	INTEGER	*(optional)* This field is ignored. It is here so that the schema is the same for the development_events and development_event_history tables.
ENDING_DEVELOPMENT_TYPE_ID	INTEGER	This grid cell's development type at the ending of the SCHEDULED_YEAR. Index into the development_types table
RESIDENTIAL_UNITS_CHANGE_TYPE	VARCHAR	N (no change), A (add), or R (replace)
RESIDENTIAL_UNITS	INTEGER	
COMMERCIAL_SQFT_CHANGE_TYPE	VARCHAR	N, A, or R
COMMERCIAL_SQFT	INTEGER	
INDUSTRIAL_SQFT_CHANGE_TYPE	VARCHAR	N, A, or R
INDUSTRIAL_SQFT	INTEGER	
GOVERNMENTAL_SQFT_CHANGE_TYPE	VARCHAR	N, A, or R
GOVERNMENTAL_SQFT	INTEGER	
RESIDENTIAL_IMPROVEMENT_VALUE_ CHANGE_TYPE	VARCHAR	N, A, or R
RESIDENTIAL_IMPROVEMENT_VALUE	INTEGER	See description, above
COMMERCIAL_IMPROVEMENT_VALUE_ CHANGE_TYPE	VARCHAR	N, A, or R
COMMERCIAL_IMPROVEMENT_VALUE	INTEGER	See description, above
INDUSTRIAL_IMPROVEMENT_VALUE_ CHANGE_TYPE	VARCHAR	N, A, or R
INDUSTRIAL_IMPROVEMENT_VALUE	INTEGER	See description, above
GOVERNMENTAL_IMPROVEMENT_VALUE_ CHANGE_TYPE	VARCHAR	N, A, or R
GOVERNMENTAL_IMPROVEMENT_VALUE	INTEGER	See description, above
FRACTION_RESIDENTIAL_LAND_VALUE_ CHANGE_TYPE	VARCHAR	N or R
FRACTION_RESIDENTIAL_LAND_VALUE	FLOAT	Fraction of residential land in this cell

Notes:

- If a CHANGE_TYPE column is "N", its corresponding value column should be zero.
- FRACTION_RESIDENTIAL_LAND_VALUE must be between 0 and 1.
- If a CHANGE_TYPE column is "N", its corresponding value column should be zero.

- GRID_ID must be a valid ID in the gridcells table or zero.
- DEVELOPMENT_TYPE_ID must be a valid index into the development_types table or zero.
- SCHEDULED_YEAR must be between ABSOLUTE_MIN_YEAR and ABSOLUTE_MAX_YEAR.
- RESIDENTIAL_UNITS must be greater than or equal to zero and <= ABSOLUTE_MAX_CELL_RESIDENTIAL_UNITS
- COMMERCIAL_SQFT must be greater than or equal to zero and <= ABSOLUTE_MAX_CELL_SQFT
- INDUSTRIAL_SQFT must be greater than or equal to zero and <= ABSOLUTE_MAX_CELL_SQFT
- GOVERNMENTAL_SQFT must be greater than or equal to zero and <= ABSOLUTE_MAX_CELL_SQFT
- RESIDENTIAL_IMPROVEMENT_VALUE must be greater than or equal to zero and <= ABSOLUTE_MAX_CELL_DOLLARS
- COMMERCIAL_IMPROVEMENT_VALUE must be greater than or equal to zeroand <= ABSOLUTE_MAX_CELL_DOLLARS
- INDUSTRIAL_IMPROVEMENT_VALUE must be greater than or equal to zero and <= ABSOLUTE_MAX_CELL_DOLLARS
- GOVERNMENTAL_IMPROVEMENT_VALUE must be greater than or equal to zero and <= ABSOLUTE_MAX_CELL_DOLLARS

development_event_history

GRID_ID	INTEGER	Grid cell where the event takes place
SCHEDULED_YEAR	SHORT	Year where the event will be implemented
DEVELOPMENT_TYPE_CHANGE_TYPE	VARCHAR	Type of change for the DEVELOPMENT_TYPE attribute, either N (no change), or R (replace)
STARTING_DEVELOPMENT_TYPE_ID	INTEGER	*(optional)* This field is ignored. It is here so that the schema is the same for the development_events and development_event_history tables.
ENDING_DEVELOPMENT_TYPE_ID	INTEGER	This grid cell's development type at the ending of the SCHEDULED_YEAR. Index into the development_types table
RESIDENTIAL_UNITS_CHANGE_TYPE	VARCHAR	N (no change), A (add), or R (replace)
RESIDENTIAL_UNITS	INTEGER	
COMMERCIAL_SQFT_CHANGE_TYPE	VARCHAR	N, A, or R
COMMERCIAL_SQFT	INTEGER	
INDUSTRIAL_SQFT_CHANGE_TYPE	VARCHAR	N, A, or R
INDUSTRIAL_SQFT	INTEGER	
GOVERNMENTAL_SQFT_CHANGE_TYPE	VARCHAR	N, A, or R
GOVERNMENTAL_SQFT	INTEGER	
RESIDENTIAL_IMPROVEMENT_VALUE_CHANGE_TYPE	VARCHAR	N, A, or R

RESIDENTIAL_IMPROVEMENT_VALUE	INTEGER	Total improvement value for this development event.
COMMERCIAL_IMPROVEMENT_ VALUE_CHANGE_TYPE	VARCHAR	N, A, or R
COMMERCIAL_IMPROVEMENT_VALUE	INTEGER	Total improvement value for this development event.
INDUSTRIAL_IMPROVEMENT_VALUE_ CHANGE_TYPE	VARCHAR	N, A, or R
INDUSTRIAL_IMPROVEMENT_VALUE	INTEGER	Total improvement value for this development event.
GOVERNMENTAL_IMPROVEMENT_ VALUE_CHANGE_TYPE	VARCHAR	N, A, or R
GOVERNMENTAL_IMPROVEMENT_ VALUE	INTEGER	Total improvement value for this development event.
FRACTION_RESIDENTIAL_LAND_ VALUE_CHANGE_TYPE	VARCHAR	N or R
FRACTION_RESIDENTIAL_LAND_ VALUE	FLOAT	Fraction of residential land in this cell

Notes:

- DEVELOPMENT_TYPE_CHANGE_TYPE is ignored on input, and updated (to be 'N' for the no-build events, and 'R' for all others) as the developer model estimation data writer runs.
- STARTING_DEVELOPEMENT_TYPE_ID is ignored on input, and updated (to be the ending development type for that gridcell in that year) as the developer model estimation data writer runs. It must *not* be a member of the vacant_undevelopable group.
- ENDING_DEVELOPMENT_TYPE_ID must *not* be a member of the vacant_developable or vacant_undevelopable development type groups.
- Value of 'R' is only allowed for DEVELOPMENT_TYPE_CHANGE_TYPE and FRACTION_RESIDENTIAL_LAND_VALUE_CHANGE_TYPE. All other *_CHANGE_TYPE columns must have values of either 'N' or 'A'.
- FRACTION_RESIDENTIAL_LAND_VALUE_CHANGE_TYPE and FRACTION_RESIDENTIAL_LAND_VALUE are ignored, as the system currently does not have enough information to use this data.

employment_events

GRID_ID	INTEGER	Grid cell where the event takes place
SCHEDULED_YEAR	SHORT	Year when the event will be implemented
HOME_BASED_JOBS	INTEGER	Number of home-based jobs affected by the event.
NON_HOME_BASED_JOBS	INTEGER	Number of non-home-based jobs affected by the event
SECTOR_ID	INTEGER	Index into the employment_sectors table

Notes:

- SECTOR_ID must be a valid index into the employment_sectors table.
- GRID_ID must be a valid ID in the gridcells table or zero.
- The combination of GRID_ID and SCHEDULED_YEAR must be unique for each employment event.
- SCHEDULED_YEAR must be between ABSOLUTE_MIN_YEAR and ABSOLUTE_MAX_YEAR.
- The number of jobs can be positive or negative to indicate addition or removal of employment.

land_use_events

GRID_ID	INTEGER	Grid cell where the event takes place
SCHEDULED_YEAR	SHORT	Year where the event will be implemented
PLAN_TYPE_CHANGE_TYPE	VARCHAR	Type of change for the PLAN_TYPE attribute, either N (no change) or R (replace)
PLAN_TYPE_ID	INTEGER	Index into the plan_types table
IS_OUTSIDE_UGB_CHANGE_TYPE	VARCHAR	N or R
IS_OUTSIDE_UGB	BOOLEAN	See gridcells table
PERCENT_WATER_CHANGE_TYPE	VARCHAR	N, A or R
PERCENT_WATER	INTEGER	See gridcells table
PERCENT_FLOODPLAIN_CHANGE_TYPE	VARCHAR	N, A or R
PERCENT_FLOODPLAIN	INTEGER	See gridcells table
PERCENT_WETLAND_CHANGE_TYPE	VARCHAR	N, A or R
PERCENT_WETLAND	INTEGER	See gridcells table
PERCENT_SLOPE_CHANGE_TYPE	VARCHAR	N, A or R
PERCENT_SLOPE	INTEGER	See gridcells table
PERCENT_OPEN_SPACE_CHANGE_TYPE	VARCHAR	N, A or R
PERCENT_OPEN_SPACE	INTEGER	See gridcells table
PERCENT_PUBLIC_SPACE_CHANGE_TYPE	VARCHAR	N, A or R
PERCENT_PUBLIC_SPACE	INTEGER	See gridcells table
PERCENT_ROADS_CHANGE_TYPE	VARCHAR	N, A or R
PERCENT_ROADS	INTEGER	See gridcells table
PERCENT_STREAM_BUFFER_CHANGE_TYPE	VARCHAR	N, A or R
PERCENT_STREAM_BUFFER	INTEGER	See gridcells table

Notes:

- PLAN_TYPE_ID must be a valid index into the plan_types table.
- GRID_ID must be a valid ID in the gridcells table or zero.
- The combination of GRID_ID and SCHEDULED_YEAR must be unique for each land use event.
- SCHEDULED_YEAR must be between ABSOLUTE_MIN_YEAR and ABSOLUTE_MAX_YEAR.
- The 'PERCENT' value columns must be between -100 and 100 for a corresponding CHANGE_TYPE of 'A', or between 0 and 100 for a corresponding CHANGE_TYPE of 'R'.

gridcell_fractions_in_zones

GRID_ID	INTEGER	The ID of a grid cell that is partially within this zone.
ZONE_ID	INTEGER	The ID of a traffic analysis zone that intersects this grid cell.
FRACTION	DOUBLE	The fraction of this grid cell that lies in this zone.

Notes:

- GRID_ID must match a grid cell in gridcells.

- ZONE_ID must match a zone in zones and it must indicate a zone that partially covers a grid cell.
- FRACTION must be > 0 and < 1, e.g. 0.4 for 40%.
- Sum of fraction for each grid cell should be 1.

travel_data

FROM_ZONE_ID	INTEGER	"From" traffic analysis zone
TO_ZONE_ID	INTEGER	"From" traffic analysis zone
LOGSUM0	FLOAT	*(Not checked by Schema Checker)* Logsum value for 0 vehicle households, transit logsum
LOGSUM1	FLOAT	*(Not checked by Schema Checker)* Logsum value for 1 vehicle households, transit logsum
. . .	FLOAT	*(Not checked by Schema Checker)* . . .
LOGSUMN	FLOAT	*(Not checked by Schema Checker)* Logsum value for N+ vehicle households, transit logsum

Notes:

- There must be a row for each combination of FROM_ZONE_ID and TO_ZONE_ID for all zones in the zones table. For instance, if zones contains 3 zones (1, 2, and 5) there must be at least the following 9 entries in travel_data: (1,1), (1,2), (1,5), (2,1), (2,2), (2,5), (5,1), (5,2), (5,5). When the TravelDataMatrix is instantiated it allocates an array based on the maximum of FROM_ZONE_ID and TO_ZONE_ID squared times the number of logsums. Later when this array is accessed, it is accessed by the zone id's read from the zones table. Thus, if the numbers of zones represented in these two tables are inconsistent, this will cause an array access violation.
- There may be additional entries in travel_data for zones for FROM_ZONE_ID or TO_ZONE_ID values that are not zones in the zones table. These are not used by the model computation.
- All LOGSUM* values must be less than or equal to zero.
- There must be at least one LOGSUM* column.
- The LOGSUM* column names must start at LOGSUM0 and increase consecutively through the integers. The names must match the Perl regular expression /LOGSUM\d+/.
- HouseholdLocationChoiceModel requires LOGSUM0-LOGSUM2 to be defined. If LOGSUM3 is not defined then LOGSUM2 is used for all computations that would otherwise have used LOGSUM3.
- If you have positive logsum values, subtract the maximum logsum value from all logsums in your table. This will correctly shift the logsums so that none are greater than zero.

zones

ZONE_ID	INTEGER	Unique identifier
TRAVEL_TIME_TO_AIRPORT	INTEGER	Units: Minutes
TRAVEL_TIME_TO_CBD	INTEGER	Units: Minutes
FAZ_ID	INTEGER	*(optional)* Foreign key of the FAZ (forecast analysis zone) containing this zone. This field is required by the Opus implementation of the home-based employment location choice model, which uses this field instead of the zones_in_faz table. This field is not required for the Java version of the model.

Notes:
- ZONE_ID must be unique and greater than zero.
- TRAVEL_TIME_TO_AIRPORT must be >= 0 and <= ABSOLUTE_MAX_TRAVEL_TIME.
- TRAVEL_TIME_TO_CBD must be >= 0 and <= ABSOLUTE_MAX_TRAVEL_TIME.

development_types

DEVELOPMENT_TYPE_ID	INTEGER	Unique identifier
NAME	VARCHAR	Name of the development type
MIN_UNITS	INTEGER	Minimum number of units to be in this development type.
MAX_UNITS	INTEGER	Maximum number of units to be in this development type.
MIN_SQFT	INTEGER	Minimum square feet to be in this development type.
MAX_SQFT	INTEGER	Maximum square feet to be in this development type.
PRIMARY_USE_ID	INTEGER	The primary use of this development type. Index into primary_uses table.

Notes:
- DEVELOPMENT_TYPE_ID must be unique, greater than zero, and less than or equal to 99.
- We recommend that DEVELOPMENT_TYPE_IDs start at 1 and be sequential.
- MIN_UNITS must be >= 0.
- MAX_UNITS must be >= MIN_UNITS.
- MIN_SQFT must be >= 0.
- MAX_SQFT must be >= MIN_SQFT.
- PRIMARY_USE must be a valid index in the primary_uses table.
- The development types should not overlap, and should completely cover the space.

development_type_groups

GROUP_ID	INTEGER	Unique identifier
NAME	VARCHAR	Unique name of the development type group
NON_OVERLAPPING_GROUPS	VARCHAR	Name of the non-overlapping-group or empty for no non-overlapping-group

Notes:
- GROUP_ID must be unique, greater than zero, and less than or equal to 99.
- NAME must be unique. The required development type groups must be lower case with underscores between words e.g. high_density_residential. We recommend that all NAMEs follow this style.
- NAMEs and NON_OVERLAPPING_GROUPS names must not contain spaces.
- Development types must not overlap across the groups in the same NON_OVERLAPPING_GROUPS.

development_type_group_definitions

DEVELOPMENT_TYPE_ID	INTEGER	Index into the DevelopmentType table
GROUP_ID	INTEGER	Index into the development_type_groups table

Notes:

- DEVELOPMENT_TYPE_ID must be a valid index into the development_types table
- GROUP_ID must be a valid index into the development_type_groups table
- The combination of DEVELOPMENT_TYPE_ID and GROUP_ID must be unique

primary_uses

PRIMARY_USE_ID	INTEGER	Unique identifier
NAME	VARCHAR	Case-insensitive name for this primary use, e.g. "commercial" or "mixed Use".

Notes:

- PRIMARY_USE_ID must be unique and greater than zero.
- There must be exactly one row with each of the following NAME values: "commercial", "industrial", and "government".

geographies

GEOGRAPHY_TYPE_ID	INTEGER	Unique identifier for this geography.
GEOGRAPHY_TYPE_TITLE	VARCHAR	A short title describing this geography, e.g. "district".
SHAPEFILE_PATH	VARCHAR	Relative path to this ESRI shapefile (relative to the UrbanSim workspace directory).
SHAPEFILE_JOIN_COLUMN	VARCHAR	The name of the join column in the shapefile for joining the shapes with the GEOGRAPHY_IDs in the databases.
COLUMN_NAME	VARCHAR	The name of the gridcell table's column containing the data to plot in a JUMP map. Jump gets it as the value of "org.urbansim.jump.Join Column".
POLYGON_ID_VALID_MIN	INTEGER	Minimum valid GEOGRAPHY_ID value for use by this geography.
POLYGON_ID_VALID_MAX	INTEGER	Maximum valid GEOGRAPHY_ID value for use by this geography.

Notes:

- GEOGRAPHY_TYPE_ID must be greater than zero.
- GEOGRAPHY_TYPE_ID = 1 must be the "region" geography.
- GEOGRAPHY_TYPE_ID = 2 must be the "grid" geography.
- GEOGRAPHY_TYPE_TITLE must be non-empty with no leading, trailing, or embedded spaces.
- SHAPEFILE_PATH must be non-empty for a map-based visualization to work with this geography.
- POLYGON_ID_VALID_MIN must be greater than zero.
- POLYGON_ID_VALID_MAX must be greater than or equal to POLYGON_ID_VALID_MIN.

geography_names

GEOGRAPHY_TYPE_ID	INTEGER	The geography
GEOGRAPHY_ID	INTEGER	A unique identifier for a particular polygon of the geography with this GEOGRAPHY_TYPE_ID
NAME	VARCHAR	The name of this polygon, e.g., "Seattle" or "PSRC" or "downtown"

gridcells_in_geography

GRID_ID	INTEGER	The grid cell
GEOGRAPHY_TYPE_ID	INTEGER	The geography containing this grid cell
GEOGRAPHY_ID	INTEGER	A unique identifier for a particular polygon of the geography with this GEOGRAPHY_TYPE_ID

Notes:

- GEOGRAPHY_ID must be between POLYGON_ID_VALID_MIN and POLYGON_ID_VALID_MAX, inclusive, of the geography with GEOGRAPHY_TYPE_ID as specified in the geographies table.

gridcells

GRID_ID	INTEGER	Unique identifier
COMMERCIAL_SQFT	INTEGER	The sum of the square footage of buildings that are classified as commercial (generally including retail and office land uses). This is not a measure of land area.
DEVELOPMENT_TYPE_ID	INTEGER	Index into the Development Types table
DISTANCE_TO_ARTERIAL	FLOAT	Units: urbansim_constants.UNITS
DISTANCE_TO_HIGHWAY	FLOAT	Units: urbansim_constants.UNITS
GOVERNMENTAL_SQFT	INTEGER	
INDUSTRIAL_SQFT	INTEGER	
COMMERCIAL_IMPROVEMENT_VALUE	INTEGER	See description, above
INDUSTRIAL_IMPROVEMENT_VALUE	INTEGER	See description, above
GOVERNMENTAL_IMPROVEMENT_VALUE	INTEGER	See description, above
NONRESIDENTIAL_LAND_VALUE	INTEGER	Units, e.g. dollars
RESIDENTIAL_IMPROVEMENT_VALUE	INTEGER	See description, above
RESIDENTIAL_LAND_VALUE	INTEGER	Units, e.g. dollars
RESIDENTIAL_UNITS	INTEGER	Number of residential units
RELATIVE_X	INTEGER	X coordinate in grid coordinate system
RELATIVE_Y	INTEGER	Y coordinate in grid coordinate system
YEAR_BUILT	INTEGER	e.g. 2002
PLAN_TYPE_ID	INTEGER	An id indicating the plan type of the grid cell
PERCENT_AGRICULTURAL_PROTECTED_LAND	INTEGER	*(optional)*
PERCENT_WATER	INTEGER	Percentage of this cell covered by water
PERCENT_STREAM_BUFFER	INTEGER	Percentage of this cell covered by stream

		buffer
PERCENT_FLOODPLAIN	INTEGER	Percentage of this cell covered by flood plain
PERCENT_WETLAND	INTEGER	Percentage of this cell covered by wetland
PERCENT_SLOPE	INTEGER	Percentage of this cell covered by slope
PERCENT_OPEN_SPACE	INTEGER	Percentage of this cell covered by open space
PERCENT_PUBLIC_SPACE	INTEGER	Percentage of this cell covered by public space
PERCENT_ROADS	INTEGER	Percentage of this cell covered by roads
PERCENT_UNDEVELOPABLE	INTEGER	*(optional)*
IS_OUTSIDE_URBAN_GROWTH_BOUNDARY	BOOLEAN	
IS_STATE_LAND	BOOLEAN	*(optional)*
IS_FEDERAL_LAND	BOOLEAN	*(optional)*
IS_INSIDE_MILITARY_BASE	BOOLEAN	*(optional)*
IS_INSIDE_NATIONAL_FOREST	BOOLEAN	*(optional)*
IS_INSIDE_TRIBAL_LAND	BOOLEAN	*(optional)*
ZONE_ID	INTEGER	Traffic analysis zone that contains this grid cell's centroid
CITY_ID	INTEGER	City this Grid Cell belongs to
COUNTY_ID	INTEGER	County this Grid Cell belongs to
FRACTION_RESIDENTIAL_LAND	FLOAT	Fraction of residential land in this cell
TOTAL_NONRES_SQFT	INTEGER	*(optional)*
TOTAL_UNDEVELOPABLE_SQFT	INTEGER	*(optional)*

Notes:

- FRACTION_RESIDENTIAL_LAND must be between 0 and 1
- COMMERCIAL_SQFT must be >= 0 and <= ABSOLUTE_MAX_CELL_SQFT
- DEVELOPMENT_TYPE_ID must be a valid index in the development_types table
- DISTANCE_TO_ARTERIAL must be >= 0 and <= ABSOLUTE_MAX_DISTANCE
- DISTANCE_TO_HIGHWAY must be >= 0 and <= ABSOLUTE_MAX_DISTANCE
- INDUSTRIAL_SQFT must be >= 0 and <= ABSOLUTE_MAX_CELL_SQFT
- GOVERNMENTAL_SQFT must be >= 0 and <= ABSOLUTE_MAX_CELL_SQFT
- GRID_ID must be unique and > 0
- INDUSTRIAL_SQFT must be >= 0
- COMMERCIAL_IMPROVEMENT_VALUE must be >= 0 and <= ABSOLUTE_MAX_CELL_DOLLARS
- INDUSTRIAL_IMPROVEMENT_VALUE must be >= 0 and <= ABSOLUTE_MAX_CELL_DOLLARS
- GOVERNMENTAL_IMPROVEMENT_VALUE must be >= 0 and <= ABSOLUTE_MAX_CELL_DOLLARS
- NONRESIDENTIAL_LAND_VALUE must be >= 0 and <= ABSOLUTE_MAX_CELL_DOLLARS
- RESIDENTIAL_IMPROVEMENT_VALUE must be >= 0 and <= ABSOLUTE_MAX_CELL_DOLLARS
- RESIDENTIAL_LAND_VALUE must be >= 0 and <= ABSOLUTE_MAX_CELL_DOLLARS
- RESIDENTIAL_UNITS must be must be >= 0 and <= ABSOLUTE_MAX_CELL_RESIDENTIAL_UNITS
- RELATIVE_X, RELATIVE_Y coordinate pairs must be unique, and >= 1.

- There must be at least one grid cell whose RELATIVE_X is 1 and at least one grid cell whose RELATIVE_Y is 1, but grid cell (1,1) is, itself, not required.
- The RELATIVE_X and RELATIVE_Y columns are measured in grid cell units. They are specifically **not** latitude/longitude or any other universal measurement system.
- YEAR_BUILT must be less than or equal to the start date of the scenario, and must be between ABSOLUTE_MIN_YEAR and ABSOLUTE_MAX_YEAR
- PLAN_TYPE must be a valid index in the plan_types table
- PERCENT_WATER must be between 0 and 100
- PERCENT_STREAM_BUFFER must be between 0 and 100
- PERCENT_FLOODPLAIN must be between 0 and 100
- PERCENT_WETLAND must be between 0 and 100
- PERCENT_SLOPE must be between 0 and 100
- PERCENT_OPEN_SPACE must be between 0 and 100
- PERCENT_PUBLIC_SPACE must be between 0 and 100
- PERCENT_ROADS must be between 0 and 100
- ZONE_ID must be a valid id in the zones table
- CITY_ID must be a valid index into the cities table or zero if there is no city
- COUNTY_ID must be a valid index into the counties table or zero if there is no county
- gridcells with any households on them (i.e., households.GRID_ID = gridcell.GRID_ID), then the gridcell.RESIDENTIAL_UNITS must be greater than 0

plan_types

PLAN_TYPE_ID	INTEGER	Unique identifier
NAME	VARCHAR	Unique name of the Plan Type

Notes:

- PLAN_TYPE_ID must be unique, greater than zero, and less than or equal to 9999.
- We recommend that PLAN_TYPE_IDs start at 1 and be sequential.
- NAME must be unique. We recommend that NAMEs follow the style guide.

selected_pumas

STATE_ID	VARCHAR	
PUMA_ID	VARCHAR	
FAMILY_TREE	VARCHAR	"Family", "NonFamily", or "GroupQuarters"

Notes:

- For each row, the appropriate FAMILY_TYPE_* entry in the joint_distribution_tables must be non-null.
- Similarly, is it expected that the corresponding PUMS table will have at least one row with the PUMA_ID.

gridcell_block_group_mapping

GRID_ID	INTEGER	A unique identifier for an UrbanSim gridcell
STATE_ID	VARCHAR	First of four key fields that, together, uniquely identify the block group that contains this gridcell's centroid
COUNTY_ID	VARCHAR	Second of four key fields...
CENSUS_TRACT	VARCHAR	Third of four key fields...
BLOCK_GROUP	VARCHAR	Fourth of four key fields...

Notes:
- This table must contain all the GRID_IDs that are in the gridcells table.
- The four key fields must contain entries from the block_group_summary table.

joint_distribution_tables

STATE_ID	VARCHAR	PUMS tables differ by state so this is the key to the table.
FAMILY_TYPE_FAMILY	VARCHAR	Name of the PUMS table for FAMILY_TYPE=Family; e.g., PUMS905H
FAMILY_TYPE_NONFAMILY	VARCHAR	Name of the PUMS table for FAMILY_TYPE=NonFamily; e.g., PUMS905H
FAMILY_TYPE_GROUPQUARTERS	VARCHAR	Name of the PUMS table for FAMILY_TYPE=GroupQuarters; e.g., PUMS905P
CENSUS_YEAR	INTEGER	Description only, e.g., 1990
SAMPLE_TYPE	VARCHAR	Description only, e.g., 5%
WEIGHT_COLUMN_NAME	VARCHAR	Name of column in PUMS that has household weight

Notes:
- For each of FAMILY_TYPE_FAMILY, FAMILY_TYPE_NONFAMILY, and FAMILY_TYPE_GROUPQUARTERS, if the cell is not null, then the named table must exist in the database.
- Similarly, those PUMS tables must have a column specified by the WEIGHT_COLUMN_NAME.
- STATE_IDs must be unique.
- There is one row in this table for each STATE_ID in the selected_pumas table.

classification_variables

FAMILY_TYPE	VARCHAR	
VARIABLE_NAME	VARCHAR	e.g., "Age"

Notes:
- The VARIABLE_NAMEs must be unique for a given FAMILY_TYPE, although different FAMILY_TYPEs may share VARIABLE_NAMEs.
- Each VARIABLE_NAME must have at least one row in the classification_categories table.
- There is at least one row (one variable) for each FAMILY_TYPE that is not null in the joint_distribution_tables for the STATE_IDs of all rows in the selected_pumas table.

classification_categories

VARIABLE_NAME	VARCHAR	e.g., "Age"
CATEGORY	INTEGER	the index number of the category; effectively the array index if the variable categories were an array. Must be contiguous integers starting at 1.
MIN	INTEGER	minimum non-negative value (inclusive) for a value to be in this category, or -1 for negative infinity.
MAX	INTEGER	Maximum non-negative value (inclusive) for a value to be in this category, or -1 for positive infinity.

Notes:

- If the variable category is categorical, i.e. one number, list that number as both the min and the max.
- If the variable category is formed from several distinct categories, i.e. one range cannot cover them since other categories come in between, then that variable name and category are listed repeatedly with several min-max ranges to grab all of the ranges contained by the category.
- MIN <= MAX

joint_distribution_details

FAMILY_TYPE	VARCHAR	
VARIABLE_NAME	VARCHAR	e.g., "Age"
COLUMN_NAME	VARCHAR	Name of the column in the joint distribution table that contains this variable

Notes:

- The VARIABLE_NAMEs must be specified in the classification_variables table.
- The COLUMN_NAMEs must exist in the corresponding PUMS tables.

block_group_summary

PUMA_ID	VARCHAR	
STATE_ID	VARCHAR	First of four key fields that, together, uniquely identify a block
COUNTY_ID	VARCHAR	Second of four key fields...
CENSUS_TRACT	VARCHAR	Third of four key fields...
BLOCK_GROUP	VARCHAR	Fourth of four key fields...
FAMILY_HOUSEHOLDS	INTEGER	Number of family households in this block group
NONFAMILY_HOUSEHOLDS	INTEGER	Number of non-family households in this block group
GROUPQUARTERS_HOUSEHOLDS	INTEGER	Number of persons in group quarters in this block group
POPULATION	INTEGER	Total population in this block group
RESIDENTIAL_VACANCY_RATE	DOUBLE	The residential vacancy rate in this block group (from STF-1)

Notes:

- FAMILY_HOUSEHOLDS >= 0, NONFAMILY_HOUSEHOLDS >= 0, GROUPQUARTERS_HOUSEHOLDS >= 0, and the sum of those three > 0
- POPULATION > 0
- RESIDENTIAL_VACANCY_RATE >= 0

selected_marginal_distribution_tables

FAMILY_TYPE	VARCHAR	
MARGINAL_TABLE_NAME	VARCHAR	Name of the STF-3A table, e.g., P112NumberOfWorkers

Notes:

- The FAMILY_TYPE must be valid.
- The MARGINAL_TABLE_NAME table for each row must exist.

- At least one row for each MARGINAL_TABLE_NAME must exist in the marginal_distribution_details table.

marginal_distribution_details

MARGINAL_TABLE_NAME	VARCHAR	Name of the marginal distribution table, for example the name of an STF-3A table
COLUMN_NAME	VARCHAR	Name of the column in the table
VARIABLE_NAME_1	VARCHAR	*(Not checked by Schema Checker)* Name of the first classification variable for which this column defines a marginal
CATEGORY_1	INTEGER	*(Not checked by Schema Checker)* Category in the first classification variable for which this column defines a marginal
VARIABLE_NAME_2	VARCHAR	*(Not checked by Schema Checker)* Name of the second classification variable for which this column defines a marginal
CATEGORY_2	INTEGER	*(Not checked by Schema Checker)* Category in the second classification variable for which this column defines a marginal
VARIABLE_NAME_3	VARCHAR	*(Not checked by Schema Checker)* Name of the third classification variable ...
CATEGORY_3	INTEGER	*(Not checked by Schema Checker)* Category in the second classification variable ...
...	...	*(Not checked by Schema Checker)* Any of number of consecutive integer VARIABLE_NAME_N and CATEGORY_N columns can be defined.

Notes:

- The contents of the VARIABLE_NAME_* columns must exist in the classification_categories table.
- The contents of the CATEGORY_* columns must exist in the classification_categories table.
- The COLUMN_NAME must exist in the table named in MARGINAL_TABLE_NAME

household_variable_mapping_direct

FAMILY_TYPE	VARCHAR	Key to joint_distribution_tables
JOINT_DISTRIBUTION_COLUMN_NAME	VARCHAR	Name of column in the joint distribution table for this family type
URBANSIM_VARIABLE_NAME	VARCHAR	The name of this variable in UrbanSim

Notes:

- URBANSIM_VARIABLE_NAME must be from the set: PERSONS, WORKERS, AGE_OF_HEAD, INCOME, CHILDREN, RACE_ID
- URBANSIM_VARIABLE_NAME cannot be repeated in two rows in household_variable_mapping_direct
- URBANSIM_VARIABLE_NAME must not be contained in both household_variable_mapping_direct and household_variable_mapping_indirect
- Each URBANSIM_VARIABLE_NAME must match a unique JOINT_DISTRIBUTION_COLUMN_NAME for that FAMILY_TYPE

household_variable_mapping_indirect

FAMILY_TYPE	VARCHAR	Key to joint_distribution_tables
URBANSIM_VARIABLE_NAME	VARCHAR	The name of a variable in UrbanSim
URBANSIM_VALUE	INTEGER	The value that the UrbanSim variable should take for this joint distribution category.
JOINT_DISTRIBUTION_COLUMN_NAME	VARCHAR	Name of column in the joint distribution table for this family type
JOINT_DISTRIBUTION_CATEGORY_MIN	INTEGER	Minimum positive value (inclusive) for a value to be in this category, or -1 for negative infinity.
JOINT_DISTRIBUTION_CATEGORY_MAX	INTEGER	Maximum positive value (inclusive) for a value to be in this category, or -1 for positive infinity.

Notes:

- If the joint distribution category is not a range but a single numeric code, then repeat that code for the min and max.
- URBANSIM_VARIABLE_NAME can be repeated in several rows in household_variable_mapping_indirect to capture several disjoint ranges.
- See household_variable_mapping_direct for other restrictions.

vehicle_ownership_distribution

DISTRIBUTION_CATEGORY	INTEGER	A unique identifier for this vehicle ownership distribution category
ZONE_ID	INTEGER	A unique identifier for a traffic analysis zone or -1 if not used
MINIMUM_INCOME	INTEGER	Minimum household income for this category or -1 for don't care
MAXIMUM_INCOME	INTEGER	Maximum household income for this category or -1 for don't care
PERSONS	INTEGER	The number of persons in a household for this category
PROBABILITY_0_CARS	DOUBLE	*(Not checked by Schema Checker)* The probability of a household in this category having 0 cars
PROBABILITY_1_CAR	DOUBLE	*(Not checked by Schema Checker)* The probability of a household in this category having 1 car
PROBABILITY_2_CARS	DOUBLE	*(Not checked by Schema Checker)* The probability of a household in this category having 2 cars
...	DOUBLE	*(Not checked by Schema Checker)* Add additional rows for 3, 4, etc. vehicles. The last row indicates that number of vehicles or more.

Appendix IV. UrbanSim Variables and Descriptions

The following is a list of variables that are available from UrbanSim and names in short form are in the parentheses.

- average_income (AVINC): Average household income in grid cell
- average_land_value_per_acre_within_walking_distance (LVA_W): Average land value per acre within walking distance
- basic_sector_employment_within_walking_distance (E_BW): Quantity of basic sector employment within walking distance
- building_age (AGE_BL): Aggregate measure of building ages in the grid cell
- commercial_sqft_recently_added_within_walking_distance (SFCWRT): Commercial square feet. recently added within walking distance
- constant (ONE): Constant
- cost_to_income_ratio (COS_IN): A ratio of rents in the grid cell to income of the household making the decision
- devtype_? (DT_??): Indicator for development type “?”. There is one variable corresponding to each defined development type, where “?” is the development type ID (e.g. devtype_1, devtype_2, etc.).
- devtype_dynamic_land_use_variables_? (DTC_??): Indicator for development type group “?”. There is exactly one variable corresponding to each defined development type group, where “?” is the name of the development type group (e.g. residential, commercial).
- devtype_industrial_or_governmental (DT_I_G): Indicator for industrial or governmental development type
- distance_to_development (D_DEV): Distance to development
- distance_to_highway (DIST_H): Distance to the nearest highway
- employment_in_sector_?_within_walking_distance (E_??W): Indicator for employment sector “?”. There is one variable corresponding to each defined Employment Sector, where “?” is the SECTOR_ID (e.g. employment_in_sector_1_within_walking_distance).
- household_size_and_ln_residential_units (S_LDU): Interaction term between household size and log of density of housing within walking distance
- household_size_and_ln_residential_units_within_walking_distance (S_LDUW): Interaction term between household size and log of density within the grid cell
- household_with_no_children_in_high_density_residential (NC_HDR): Indicator for households with no children in a high density residential cell
- household_with_no_children_in_mixed_use (NC_M): Indicator for households with no children in a mixed use cell
- income_and_ln_commercial_sqft_within_walking_distance (INCLCW): Interaction term between income and log of commercial space within walking distance
- income_and_ln_improvement_value_per_unit (incival): Interaction between income and log of improvement value per residential unit
- income_and_ln_industrial_sqft_within_walking_distance (INCLIW): Interaction term between income and log of industrial space within walking distance
- income_and_ln_residential_units (INCLR): Interaction term between income and log of residential units

- income_and_ln_total_residential_value_per_residential_unit_within_walking_distance (INCLRVU): Interaction term between income and log of the average residential unit value across the walking radius.
- income_and_year_built (INCYRB): Interaction term between income and year built
- income_times_percent_residential (INCPRE): Household income interacted with percent of the gridcell that is residential land
- is_developed (DVLPD): Indicator for cells developed
- is_in_floodplain (FLOOD): Indicator for cells in a floodplain
- is_in_stream_buffer (STRBUF): Indicator for cells in a stream buffer
- is_in_wetland (WTLND): Indicator for cells in wetland
- is_near_arterial (ART): Indicator for cells near an arterial
- is_near_highway (HWY): Indicator for cells near a highway
- is_on_steep_slope (SLOPE): Indicator for cells on a steep slope
- is_outside_urban_growth_boundary (O_UGB): Indicator for cells outside the urban growth boundary
- ln_available_job_locations (LAJL): Log of available job spaces in the cell
- ln_available_residential_units (LDUA): Log of the number of residential units in the cell
- ln_average_land_value_per_acre_within_walking_distance (LALVAW): Log of the average land value per acre within walking distance
- ln_average_total_value_per_residential_unit_within_walking_distance (LAVURW): Log of the average total value per residential unit within walking distance
- ln_commercial_sqft (LSFC): Log of commercial sq. ft. in the grid cell
- ln_commercial_sqft_recently_added_within_walking_distance (LSFCWR): Log of commercial sq. ft. recently added within walking distance
- ln_commercial_sqft_within_walking_distance (LSFCW): Log of commercial sq. ft within walking distance
- ln_distance_to_highway (LD_HY): Log of the distance to nearest highway
- ln_home_access_to_employment_0_cars (LHAE00): Log of accessibility to employment if the decision-making household chooses this cell TAZ, given that it is a zero-vehicle household
- ln_home_access_to_employment_1 (LHAE1): Log of accessibility to employment for one-vehicle households in the cells TAZ
- ln_home_access_to_employment_1_car (LHAE11): Log of accessibility to employment if the decision-making household chooses this cells TAZ, given that it is a one-vehicle household
- ln_home_access_to_employment_2 (LHAE2): Log of accessibility to employment for two-vehicle households in the cells TAZ
- ln_home_access_to_employment_2_cars (LHAE22): Log of accessibility to employment if the decision-making household chooses this cell TAZ, given that it is a two-vehicle household
- ln_home_access_to_employment_3plus (LHAE3): Log of accessibility to employment for three-vehicle households in the cells TAZ
- ln_home_access_to_employment_3plus_cars (LHAE33): Log of accessibility to employment if the decision-making household chooses this cell TAZ, given that it is a three-or-more-vehicle household

- ln_home_access_to_employment_no_workers (LHAENW): Log of accessibility to employment given that the decision-making household has no workers
- ln_home_access_to_employment_transit (LHAE0): Log of accessibility to employment for zero-vehicle households in the cells TAZ
- ln_home_access_to_employment_workers (LHAEWW): Log of accessibility to employment given that the decision-making household has workers
- ln_home_access_to_population_1 (LHAP_1): Log of accessibility to population for one-vehicle households in the cells TAZ
- ln_housing_affordability (LOGIP): When the quantity (household income minus one tenth of the gridcells average price per residential unit) is negative, this is a very low negative number, otherwise it is the log of that quantity
- ln_improvement_value_per_residential_unit_within_walking_distance (LIVUW): Log of improvement value per residential unit within walking distance
- ln_income_times_home_access_to_employment_1 (LINCAE1): Log of the quantity household income times accessibility to employment for one-car households
- ln_industrial_sqft_within_walking_distance (LSFIW): Log of industrial sq. ft. within walking distance
- ln_land_value (LLV): Log of the total land value in the grid cell
- ln_n_residential_units_recently_added_within_walking_distance (LDUWRT): Log of the number of residential units recently added in cells within walking distance
- ln_percent_?_within_walking_distance (LP_W??): Log of the percent of development type group "?" within walking distance.
- ln_residential_improvement_value_per_residential_unit (LIVU): Log of the residential improvement value per residential unit
- ln_residential_units (LDU): Log of the number of residential units in the grid cell
- ln_residential_units_within_walking_distance (LDUW): Log of the number of residential units within walking distance
- ln_residential_units_within_walking_distance_0_cars (LDUW_0): Log of the number of residential units within walking distance, given that the decision-making household has no vehicles
- ln_residential_units_within_walking_distance_1_car (LDUW_1): Log of the number of residential units within walking distance, given that the decision-making household has one vehicle
- ln_residential_units_within_walking_distance_2_cars (LDUW_2): Log of the number of residential units within walking distance, given that the decision-making household has two vehicles
- ln_residential_units_within_walking_distance_3plus_cars (LDUW_3): Log of the number of residential units within walking distance, given that the decision-making household has three or more vehicles
- ln_residential_units_within_walking_distance_fewer_cars_than_workers (LDUW_F): Log of the number of residential units within walking distance, given that the decision-making household has fewer cars than workers
- ln_retail_within_walking_distance (LSFREW): Log of quantity of retail within walking distance

- ln_retail_within_walking_distance_fewer_cars_than_workers (LSFREF): Log of quantity of retail within walking distance, given that the decision-making household has fewer cars than workers
- ln_total_employment_within_walking_distance (LE_W): Log of total employment within walking distance
- ln_total_improvement_value (LIV): Log of total improvement value in the cell
- ln_total_job_locations (LTJL): Log of total number of job spaces in the cell
- ln_total_land_value_per_acre_within_walking_distance (LLVA_W): Log of total land value per acre within the walking radius
- ln_total_nonresidential_sqft_within_walking_distance (LNRSFW): Log of non-residential sq. ft. within walking distance
- ln_total_residential_value_per_residential_unit (LVU_R): Log of total residential value per residential unit in the grid cell
- ln_total_residential_value_per_residential_unit_within_walking_distance (LVU_RW): Log of total residential value per residential unit within walking distance
- ln_total_value (LV): Log of total value of the cell
- ln_work_access_to_employment_1 (LWAE_1): Log of work accessibility to employment for one vehicle households in the cells TAZ
- ln_work_access_to_population_1 (LWAP_1): Log of work accessibility to population for one vehicle households in the cells TAZ
- n_recent_transitions_to_?_within_walking_distance (TR??WR): Number of recent transitions to development type "?" within walking distance
- n_recent_transitions_to_developed_within_walking_distance (TRDWRT): Number of recent transitions to development type group 'developed' within walking distance
- n_recent_transitions_to_same_type_within_walking_distance (TRSWRT): Number of recent transitions to this cell's development type within walking distance
- n_residential_units_recently_added_within_walking_distance (DURWRT): Number of residential units recently added within walking distance
- non_residential_sqft (NR_SF): Non residential square feet in cells
- percent_?_within_walking_distance (P??W): Percent of development type group "?" within walking distance
- percent_developed_within_walking_distance (P_DEV): Percent of cell covered by developed area
- percent_floodplain (PFL): Percent of cell covered by floodplain
- percent_high_income_households_within_walking_distance (PHIW): Percent of households within walking distance that are designated as high-income
- percent_high_income_households_within_walking_distance_if_high_income (PHIW_H): Percent of households within walking distance that are designated as high-income, given that the decision-making household is high-income
- percent_high_income_households_within_walking_distance_if_low_income (PHIW_L): Percent of households within walking distance that are designated as high-income, given that the decision-making household is low-income
- percent_high_income_households_within_walking_distance_if_mid_income (PHIW_M): Percent of households within walking distance that are designated as high-income, given that the decision-making household is mid-income

- percent_low_income_households_within_walking_distance (PLIW): Percent of households within walking distance that are designated as low-income
- percent_low_income_households_within_walking_distance_if_high_income (PLIW_H): Percent of households within walking distance that are designated as low-income, given that the decision-making household is high-income
- percent_low_income_households_within_walking_distance_if_low_income (PLIW_L): Percent of households within walking distance that are designated as low-income, given that the decision-making household is low-income
- percent_low_income_households_within_walking_distance_if_mid_income (PLIW_M): Percent of households within walking distance that are designated as low-income, given that the decision-making household is mid-income
- percent_mid_income_households_within_walking_distance (PMIW): Percent of households within walking distance that are designated as mid-income
- percent_mid_income_households_within_walking_distance_if_high_income (PMIW_H): Percent of households within walking distance that are designated as mid-income, given that the decision-making household is high-income
- percent_mid_income_households_within_walking_distance_if_low_income (PMIW_L): Percent of households within walking distance that are designated as mid-income, given that the decision-making household is low-income
- percent_mid_income_households_within_walking_distance_if_mid_income (PMIW_M): Percent of households within walking distance that are designated as mid-income, given that the decision-making household is mid-income
- percent_minority_households_within_walking_distance (PMNW): Percent of heads of households within walking distance that are in a minority race
- percent_minority_households_within_walking_distance_if_minority (PMNWMJ): Percent of heads of households within walking distance that are in a minority race, given that the decision-making households head is in a minority race
- percent_minority_households_within_walking_distance_if_not_minority (PMNWMN): Percent of heads of households within walking distance that are in a minority race, given that the decision-making households head is in a minority race
- percent_open_space (POPEN): Percent of cell covered by open space
- percent_public_space (PPUB): Percent of cell covered by public space
- percent_roads (PROAD): Percent of cell covered by roads
- percent_same_type_cells_within_walking_distance (PSTCW): Percent of cells of this cells development type within walking distance
- percent_slope (PSLOPE): Percent of cell covered by slope
- percent_stream_buffer (PSTBUF): Percent of cell covered by stream buffers
- percent_water (PWATER): Percent of cell covered by water
- percent_wetland (PWETLA): Percent of cell covered by wetland
- plantype_? (PT????): Indicator for plantype "?". There is exactly one variable corresponding to each defined plan type, where "?" is the plan type (e.g. plantype_1, plantype_2).
- proximity_to_development (PRXDEV): A measure of proximity to development
- residential_units (UNITS): Residential units in cells
- residential_units_when_household_has_children (DUR_C): Number of residential units in the cell, given that the decision-making household has children.

- residential_units_within_walking_distance (DURW): Number of residential units within walking distance
- retail_sector_employment_within_walking_distance (E_REW): Quantity of retail sector employment within walking distance
- same_household_age_in_faz (SAGEFAZ): Number of households of age category (i) in faz gridcell is located in if household is in category (i)
- same_household_child_in_faz (SCHDFAZ): Number of households of child category (i) in faz gridcell is located in if household is in category (i)
- same_household_income_in_faz (SINCFAZ): Number of households of income category (i) in faz gridcell is located in if household is in category (i)
- same_household_race_in_faz (SRACFAZ): Number of households of race category (i) in faz gridcell is located in if household is in category (i)
- same_household_size_in_faz (SSIZFAZ): Number of households of size category (i) in faz gridcell is located in if household is in category (i)
- same_household_workers_in_faz (SWRKFAZ): Number of households of workers category (i) in faz gridcell is located in if household is in category (i)
- same_sector_employment_within_walking_distance (E_SAW): Quantity of same sector employment within walking distance
- same_sector_jobs_in_faz (SJOBFAZ): Number of jobs of sector (i)in faz gridcell is located in if job is in sector (i)
- service_sector_employment_within_walking_distance (E_SEW): Quantity of service sector employment within walking distance
- travel_time_to_airport (TT_AIR): AM peak hour travel time by single-occupancy vehicle from the cells TAZ to the airport TAZ.
- travel_time_to_CBD (TT_CBD): AM peak hour travel time by single-occupancy vehicle from the cells TAZ to the CBDs TAZ (or are representative TAZ for the CBD).
- trip_weighted_travel_time_for_transit_walk (TT_TW): Travel time from zone by transit and walking
- trip_weighted_travel_time_from_zone_for_SOV (TT_SOV): Travel time from zone by single occupancy vehicle
- vacancy_rate (VAC): Overall vacancy rate in cell
- vacancy_rate_for_?_within_walking_distance (VAC_??): Overall vacancy rate in cells in development type group "?" within walking distance
- young_household_in_high_density_residential (YH_HDR): Indicator for a young head of household in a high density residential cell
- young_household_in_mixed_use (YH_M): Indicator for a young head of household in a mixed use cell