Technical Report Documentation Page

1. Report No. FHWA/TX-09/0-5667-1	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle An Examination of Land Use Models, Emphasizing UrbanSim, TELUM, and Suitability Analysis		5. Report Date October 31, 2008
		6. Performing Organization Code
7. Author(s) Kara Kockelman, Jennifer Duthie, Siva Karthik Kakaraparthi, Bin (Brenda) Zhou, Ardeshir Anjomani, Sruti Marepally, and Krishna Priya Kunapareddychinna		8. Performing Organization Report No. 0-5667-1
9. Performing Organization Name and Address Center for Transportation Research The University of Texas at Austin 3208 Red River, Suite 200 Austin, TX 78705-2650		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. 0-5667
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Implementation Office P.O. Box 5080 Austin, TX 78763-5080		13. Type of Report and Period Covered Technical Report September 2006-August 2008
		14. Sponsoring Agency Code

15. Supplementary Notes
Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.

16. Abstract
This work provides integrated transportation land use modeling guidance to practitioners in Texas regions of all sizes. The research team synthesized existing land use modeling experiences from MPOs across the country, examined the compatibility of TELUM's and UrbanSim's requirements with existing data sets, integrated both UrbanSim and a gravity-based land use modeling with an existing travel demand model (TDM), and examined the results of Waco and Austin land use model runs (with and without transportation and land use policies in place). A simpler, GIS-based land use suitability analysis was also performed, for Waco, to demonstrate the potential value of such approaches. Finally, the knowledge gained from this research project was disseminated to practicing planners through a Resource Manual/Guidebook and two workshops.

17. Key Words	18. Distribution Statement
Land use modeling, Integrated transportation-land use modeling, Suitability analysis, Land use planning and policy	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161; www.ntis.gov.

19. Security Classif. (of report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	224	

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

An Examination of Land Use Models, Emphasizing UrbanSim, TELUM, and Suitability Analysis

Kara Kockelman
Jennifer Duthie
Siva Karthik Kakaraparthi
Bin (Brenda) Zhou
Ardeshir Anjomani
Sruti Marepally
Krishna Priya Kunapareddychinna

CTR Technical Report:	0-5667-1
Report Date:	August 31, 2008
Project:	0-5667
Project Title:	Analysis and Guidelines for Establishing Unified Urban Land-Use and Transportation System Planning Framework and Procedures
Sponsoring Agency:	Texas Department of Transportation
Performing Agency:	Center for Transportation Research at The University of Texas at Austin

Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.

Center for Transportation Research
The University of Texas at Austin
3208 Red River
Austin, TX 78705

www.utexas.edu/research/ctr

Disclaimers

Author's Disclaimer: The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of the Federal Highway Administration or the Texas Department of Transportation (TxDOT). This report does not constitute a standard, specification, or regulation.

Patent Disclaimer: There was no invention or discovery conceived or first actually reduced to practice in the course of or under this contract, including any art, method, process, machine manufacture, design or composition of matter, or any new useful improvement thereof, or any variety of plant, which is or may be patentable under the patent laws of the United States of America or any foreign country.

Notice: The United States Government and the State of Texas do not endorse products or manufacturers. If trade or manufacturers' names appear herein, it is solely because they are considered essential to the object of this report.

Engineering Disclaimer

NOT INTENDED FOR CONSTRUCTION, BIDDING, OR PERMIT PURPOSES.

Project Engineer: Kara M. Kockelman
Professional Engineer License State and Number: Texas # 93443
P. E. Designation: Research Supervisor

Acknowledgments

The authors express appreciation to our Project Director, TxDOT's Janie Temple, Project Coordinator Jack Foster, and Project Advisor Eduardo Calvo. Thanks also go to those at TxDOT's Research and Technology Implementation office: Duncan Stewart (RMC 2 Research Engineer), and Sylvia Medina and Loretta Brown (RMC 2 Contract Managers for this project). The research team appreciates the contributions of Caliper Corporation's Dr. Howard Slavin, Margaret Holmes, and Jian Zhang, for the use of TransCAD software and technical support. Finally, we appreciate Houston-Galveston Area Council's Dmitri Messen's and the University of Washington's Hana Ševčíková's assistance in running UrbanSim, via helpful email correspondence.

Products

Products associated with this project include the provision of land use modeling workshops for TxDOT engineers and MPO planners and a Guidebook/ Resource handbook for guidance in the land use forecasting process.

The Part I case-study workshop was held May 27, 2008 in Houston (in conjunction with TxDOT's Transportation Planning Conference) with TxDOT staff, consultants, and MPOs from various regions as participants (47 persons total). The workshop described the value of land use models (LUMs) in transportation planning and the different methodologies that can be used to implement such LUMs. Project PMC members Janie Temple and Duncan Stewart were in attendance.

The Part II hands-on workshop was held on Tuesday, July 29, 2008 in Austin, at TxDOT's Riverside Drive Annex. This workshop concentrated on the demonstration and practical application of three land use models. Attendees worked with data in a TransCAD tutorial, in order to forecast land use change using gravity-based land use modeling commands. Demonstrations of Suitability Analysis using ArcGIS tools (with Waco data) and UrbanSim model operations and results (with Austin data) were also provided. The workshop was attended by 20 participants, including TxDOT district personnel, MPO staff members, and transportation consultants. Duncan Stewart and Janie Temple were in attendance.

The Guidebook is not contained in this report, and was delivered separately to RTI on August 31, 2008.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

Rates of population and job growth, location choices of various regional actors, and demographic shifts, along with transportation system investments and other policies, impact each region's future land use and traffic patterns—and, by extension, its air quality, congestion, energy use, housing affordability, and various other important qualities. This project developed land use modeling tools for transportation planning staff working in regions of different sizes, in order to ascertain the long-run land use and traffic impacts of transportation investments and policies.

The integration of land use and travel demand models is essential to recognizing and allowing for the feedbacks that can occur across these twin systems. While much of the project work emphasized data acquisition for and application of the UrbanSim land use model for Austin, in concert with a relatively standard travel demand model, more traditional gravity-based land use models and suitability analyses were also developed and applied. The research team also synthesized existing land use modeling experiences from MPOs across the country, examined the compatibility of various land use model requirements with existing data sets and planning staff capabilities, examined the policy-making implications of uncertainty in model inputs, and performed GIS-based land use suitability analysis. In the final months of the project, key recommendations and modeling details were disseminated to practicing planners through a Resource Manual/Guidebook and two workshops. (Note: Brief workshop details are included in the Products description section of this report.) The following chapters discuss the results of key project tasks. The first of these (and its associated Appendix A) describes a variety of land use forecasting approaches, and existing experiences for planning purposes.

Chapter 2. Synthesis of Existing Land Use Modeling Experiences

In practice, three main approaches to anticipating land use futures exist. These are distinguished here through the use of the following terms: (1) "mathematical land-use model" refers to a tool that forecasts and allocates population and employment to small areas/zones in a way that is well suited for input into a travel demand model (TDM); (2) "suitability process" refers to the use of either a visioning process or a suitability model (i.e., a model that has not been developed for forecasting, but rather for ascertaining where land development of all types may best occur); and (3) "Delphi process" refers to the use of only personal, expert opinions, often supplemented by readily available data from sources such as the United States Census.

2.1 Advantages and Disadvantages of Each Approach for Anticipating Land Use Future

Each of the three approaches for forecasting land use (mathematical modeling, suitability analysis, and the Delphi process) has advantages and disadvantages, noted briefly here. For more details on individual models, readers may consult Appendix A, which summarizes the research team's review of ten land-use and land-suitability models (UrbanSim, TELUM, PECAS, TRANUS, *What if?*, SLEUTH, UPlan, LUSDR, STEP 3, and MEPLAN). Mathematical models are predictive in nature, using historical data and, ideally, behavioral and economic theory to forecast future developments. Such models allow users to test alternative policies in order to better understand their implications. A major strength of these models is their ability to integrate with a TDM. Not only can the results of the mathematical land use model feed into a TDM, but the results of the TDM can then be fed into the land use model, and so on until the solution becomes stable (if a stable solution exists). The main disadvantage cited with these types of models is the relatively large amount of resources required for implementation. Agencies may have to devote to the task at least one full-time staff person who has the expertise to understand the underlying theory, or hire a consultant. Also, the data requirements are such that it may take months to collect the necessary data, especially if the data must be geo-coded at a small level of aggregation. A significant amount of time should also be devoted to calibrating the model to a base year.

Suitability processes have a very different underlying philosophy than the mathematical models. The mathematical models aim to predict what *will* happen, and the suitability models aim to forecast what a community *wants* to happen. Suitability processes are typically very open, allowing for interactive participation from the public in developing multiple scenarios for a region's future. The results of these processes can be formatted for input in a TDM; however, feedbacks between the land use and travel systems are typically not used.

Delphi processes require nothing more than the combined opinions of a group of experts. No knowledge of modeling theory is necessary. Many planning organizations choose to inform their Delphi process with the data that is available from sources such as the United States Census, state workforce commission, and aerial photographs, creating future year predictions for large areas based on trends, then allocating the forecasts to smaller areas using an expert panel. While the results of a Delphi process can be used as inputs to a TDM, they really cannot be part of an iterative feedback model between land use and transport systems. A Delphi process therefore has limited capabilities for testing policy implications.

2.2 Summary of Survey Responses

Staff members from 13 of Texas's 25 MPOs and 5 non-Texas MPOs were interviewed under this task, in order to determine what methods currently are used to predict population and employment, and what methods are sought for future use. Detailed responses are provided in Appendix B. The responses varied widely, but demonstrated a clear correlation between MPO size and sophistication of forecasting methods. The state's three largest MPOs (Dallas-Fort Worth, Houston-Galveston, and San Antonio-Bexar County) already use land use models to allocate forecasted regional growth to smaller areas, such as traffic analysis zones (TAZs). The other ten Texas MPOs surveyed use simpler methods, combining Delphi process results with State Data Center control totals, census information, Texas Workforce Commission data, and aerial photos, in order to predict population and employment at the level of TAZs. Texas's fourth and fifth largest MPOs (Austin-Capital Area and El Paso) expect to join the top three in using mathematical models in the near future, with UrbanSim being a primary candidate (Griffin, 2007; Tinajero, 2007). Three other Texas MPOs expressed interest in using scenario planning or visioning exercises (see, e.g., Lemp et al. 2007) to aid their forecasting process. Waco is the furthest along, with consultants already in place (Evilia, 2007). Amarillo is considering the use of a suitability model, such as *What if?* (Shaw, 2007), and Brownsville will begin a visioning process this summer (Lund, 2007).

The remaining five Texas MPOs that were interviewed are not considering such approaches, for several reasons. Most have very limited staff, often only two people, and they do not feel equipped to conduct time-intensive analyses. These smaller MPOs were more likely to comment that land use modeling is unnecessary for their situation, and a few of the interviewees were not aware that such models exist. Despite small staff sizes, many of these MPOs do have available to them Geographic Information System (GIS)-encoded data sets, particularly parcel level information (i.e., lot value, land use type) used for tax assessment purposes. However, these may need major adjustment/cleaning before being used in a mathematical model such as UrbanSim.

Figures 2.1 through 2.3 offer a summary of survey results. Figure 2.1 shows the positive correlation between the region population and staff size. (A logarithmic scale is used here, to temper the great range in populations). Figures 2.2 and 2.3 illustrate the positive relationship between modeling method sophistication (current and future) and population and staff size. These correlations are intuitive, as MPOs with greater resources are better able to devote staff time to implementing complex models and methods.

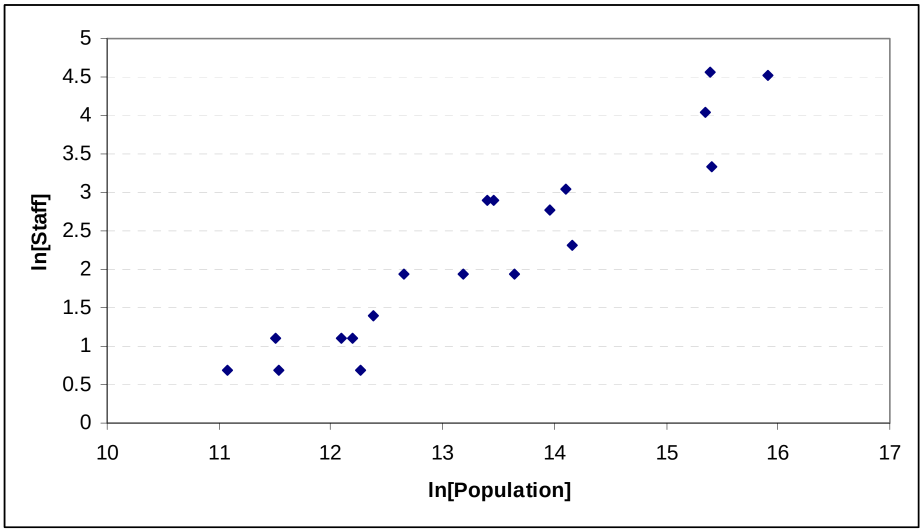


Figure 2.1: MPO Staff Size v. Population

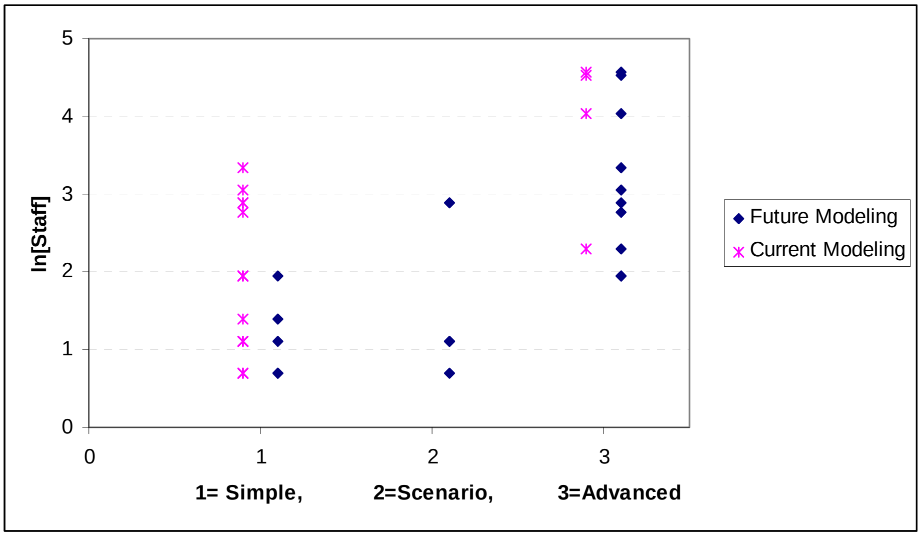


Figure 2.2: Staff Size v. Choice of Current and Future Modeling Method

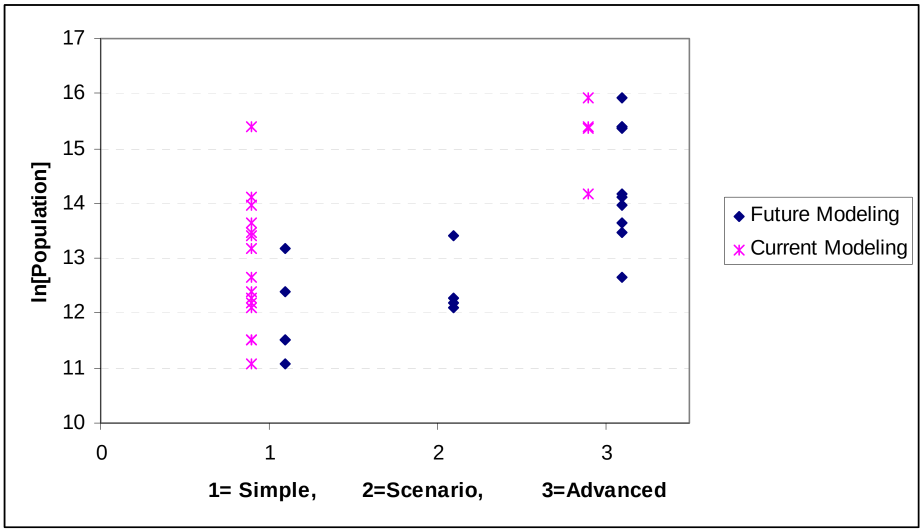


Figure 2.3: Population versus Choice of Current and Future Modeling Method

As this project progressed, the research team further explored the benefits and challenges of creating a unified urban land use transportation framework for forecasting the future of Texas regions. It is clear from the synthesis presented here that, while advanced land-use models can provide detailed outputs and test the implications of several policy decisions, the effort needed to implement such models may exceed the abilities of most MPOs. TxDOT's Transportation Planning and Programming staff expertise can provide excellent support in this regard.

Chapter 3. Investigation of Land Use Model Requirements and Compatibility with Existing TxDOT Data and Methods

For a variety of reasons, it is important that Texas's many planning organizations move towards anticipating future land use and transportation patterns in an integrated fashion. A description of existing data and methods used by TxDOT's Transportation Planning and Programming (TP&P) Division and major Texas MPOs to predict future land use and travel is provided here, along with an overview of the software, hardware, data, and statistical requirements for four eminent land use models: UrbanSim (Waddell, 2002), PECAS (Hunt and Abraham, 2007), TELUM/ITLUP (Putman, 2005), and TRANUS (de la Barra, 2007). The following discussion examines how each of these four land use models could dovetail with current TxDOT practice and what steps are needed to integrate each with existing travel model processes. Originally, this task was to focus solely on UrbanSim, but the research team decided to broaden the scope of the project due to the quantity and quality of available land use models, and the very serious data (and calibration) challenges that UrbanSim poses for TxDOT's and other agencies' current operations.

Figure 3.1 depicts the integration of a land use model with a travel model. Typically, the land use model is run every 1 to 5 years, starting from the base (most recent) year until the forecast year, and the travel model is run at that time (or less often, in the case of one-year land use modeling intervals). For example, UrbanSim suggests integration similar to that shown in Figure 3.1, while the FHWA's TELUM (based on Putman's ITLUP model) typically runs every 5 years, with a travel model run every 5 years (or less frequently). Essentially, the argument is that travel reacts quickly to land use changes, while land use changes require more time and therefore lag traveler behavior (by 1 to 5 or more years)[1]. In addition, travel models require relatively long running time (usually a couple of hours), but their outputs (needed in the next land use model run) do not change significantly across relatively short time intervals (e.g., just 1 year in UrbanSim).

Of course, all models are abstractions of reality, and such assumptions permit straightforward application of the integrated model system: a travel demand model is run for the base year's land use distribution, and then the next time step's land use distribution is forecast, permitting a contemporaneous forecast of travel demand (for that forecast year). To predict land use conditions a decade or more into the future, analysts must have an estimate of the transportation system's travel conditions in the prior period, to provide the accessibility indices, interzonal impedances, and other inputs that the land use models require. Thus, for a 20-year planning horizon, the entire travel demand model may be run five times (current/base year and then every 5 years), and the land use model four times (starting 5 years out from the current/base year). Like most "integrated" model systems, the land use and transportation models discussed here are run in sequence, feeding forward, step by step, rather than simultaneously. This approach often is referred to as a "dynamic disequilibrium," recognizing the dynamic and

[1] Of course, land use development can anticipate planned infrastructure investments, and occur in greater synch with the resulting traffic shifts. To better mimic such a process, analysts may wish to pre-locate the investment and anticipate its traffic changes to feed into contemporaneous (or prior) land use model applications.

unequilibrated nature of land use markets (where not all households or firms are in their most preferred locations at any point in time)[2].

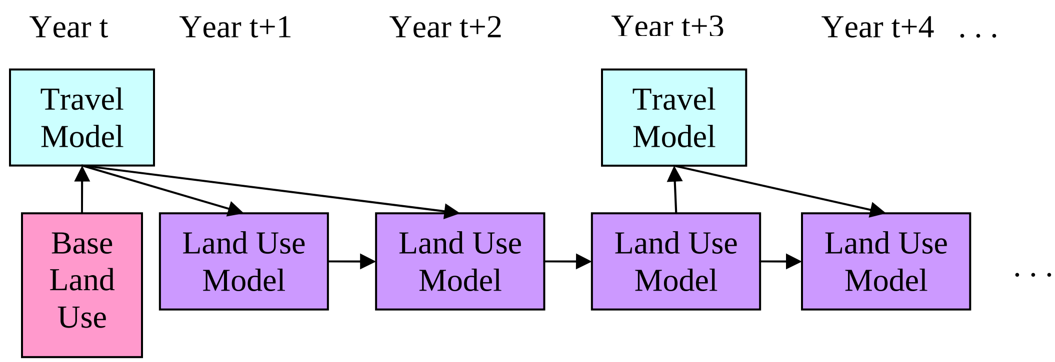


Figure 3.1: Example of Land Use Model and Travel Model Integration Process

UrbanSim, TELUM, PECAS, and TRANUS could all be integrated with existing TxDOT travel demand models (discussed in more detail in the next section) with varying degrees of effort. TELUM is the easiest of the four to implement, as the input data are more readily available, it requires no external software (ArcGIS for mapping is optional) other than Microsoft Excel, and it has an internal parameter calibration tool. However, automatic feedbacks between TELUM and a travel demand model are not readily available, due to differences in computing language/software. TELUM's code can be adjusted to run back and forth with a travel model every 5 years (or as specified by the user). Such feedback is critical to capturing the transportation-land use interaction.

UrbanSim and PECAS are much more data intensive than TELUM. PECAS requires data on the economic interactions (e.g., trade and worker flows) between zones, in order to inform its input-output model, making it a strong candidate for statewide models. (See, e.g., Kockelman et al.'s RUBMRIO model for Texas's 254 counties [in Ruiz-Juri and Kockelman, 2006].) UrbanSim emphasizes disaggregation and seeks greater behavioral realism in location and land development decisions, making it more appropriate for intra-regional analyses. Both UrbanSim and PECAS require external software, although PECAS' reliance on Microsoft Office products is more user-friendly than UrbanSim's reliance on the database software MySQL. Both models can be fully integrated with a travel model by creating linkages via coding. The UrbanSim development team and Caliper Corporation are working together to develop a seamless integration between UrbanSim and TransCAD transportation planning software.

TRANUS has a relatively simple built-in travel model, eliminating the need for an external travel model such as TransCAD. If desired, its land use model could be run stand-alone and integrated within a TxDOT travel model. TRANUS' land use model works at the aggregate level, similar to TELUM. In addition to easily obtainable inputs such as employment by sector, TRANUS requires economic data similar to PECAS, typically obtained from national databases and commodity flow surveys. For each sector, the user must specify the initial quantities of each

[2] Equilibration of such complex, multiple-agent systems requires time, and "shocks" occur along the way. These shocks include in-migration of population and jobs, changes to the transport networks, new transport and land use policies, and other events, all of which help ensure that a stable, unchanging equilibrium condition is never reached.

sector in each zone (typically bundled by value, in any industry, and by acreage, in the case of land), and in the case of land, the price per acre of land.

Significant effort is required to move outputs of one model into the next model, for example, in a series of 9 applications for a 20-year horizon with 5-year intervals. However, Caliper may soon hard-code the linkages for application of its ITLUP-based procedures, which is very useful for those already familiar with TransCAD software (like TP&P's modelers). TRANUS uses a basic travel demand model setup, hard-coded into its system. But the others discussed here (like almost all other competing land use models) require a back-and-forth between land use and transport models. Some semi-expert coding could facilitate the continuous running of such models, via Python[3], Java, or other programming languages.

The remainder of this chapter covers existing land use and travel model practices at TxDOT and major Texas MPOs, an overview of the requirements of each of the four land use models, and some concluding remarks about how each of these four major land use models could be integrated with TxDOT's travel model. A summary of the legal foundation for modeling land use can be found in Appendix C.

3.1 Existing Agency Data and Methods

Staff at the Texas Transportation Institute (TTI) provide support to Texas's MPOs in developing base year and forecast year estimates of population and employment across each region's TAZs. MPOs are encouraged to use the latest census data to develop population, households, and median household income at the TAZ level for a base year and to use the population estimates developed by the State Data Center to obtain control total estimates of population, households, and median household income. Data from the American Community Survey may be used if it is available for the base year. Allocation of the change in population and households from the last census year to the base year is based on permits from the area (i.e., building permits, septic permits). Employment for the base year is available from the Texas Workforce Commission through an agreement with TxDOT. This data includes name, address, NAICS code, X and Y coordinates, and number of workers (Ellis 2007).

TTI recommends that MPOs use forecasts from the State Data Center to obtain forecast year estimates of population, household, and income, and that these forecasts are analyzed based on historic trends, local economic conditions, and other local factors. Forecasts of employment are based on historic and current trends by type of employment. Some MPOs use the Delphi process to allocate forecast year data to zones, but others do the allocation in-house (Ellis, 2007).

In addition to providing guidance to MPOs, TTI creates the regional distributions of households by size and income range that are required as inputs into the travel model. TTI has held workshops in the past to teach MPOs how to develop these distributions; however, none of the smaller MPOs have attempted to do this (Ellis, 2007).

TxDOT's TP&P Division assists MPOs in running regional travel demand models using TransCAD with custom add-ins for trip generation and trip distribution, both implemented in TransCAD. TRIPCAL5 (Pearson et al., 1990) is designed to be a flexible trip generation model, with up to ten trip purposes, three trip production models, and five trip attraction models.

[3] Detroit's SEMCOG recently had Paul Waddell's team code Python to link its travel demand model to their UrbanSim. These files were specific to SEMCOG's TDM specification and thus were not as helpful to this project as the research team had hoped.

ATOM2, TxDOT's trip distribution model, differs from a traditional gravity model in its "direct use of the desired trip length frequency as an objective function in the iterative process" (TTI, 2001). ATOM2 has been updated since its conception in 1970 to interface with TransCAD, to allow the user to input F-factors instead of trip length frequency if desired, and to use terminal times in the trip distribution process (TTI, 2001). Several of the largest MPOs in Texas (NCTCOG, HGAC, CAMPO, El Paso MPO, and AACOG) have an in-house mode choice model (often calibrated and applied by outside consultants). The following inputs to the travel model are required for each TAZ: total developable acres, population, households, median income, total employment, basic employment, retail employment, service employment, educational employment, special generator employment, and special generator population[4]. Any land use model that is used must provide at least these inputs. The travel model provides the outputs necessary for feedback into a land use model, such as trip length distribution and zonal travel time (Hodges, 2007).

Patty Ellis, a TTI research associate who provides demographic assistance to MPOs, does not feel that the effort to create an advanced land use model is worth the effort for slow-growing areas. She expressed concern that most small Texas MPOs do not have the money or staff required to develop such a model (Ellis, 2007). Mark Hodges, with TxDOT's TP&P Division, agrees, adding that TP&P does not have the staff required to be responsible for running an advanced land use model for regions that need assistance (Hodges, 2007). Of course, the opportunity for relatively simple model implementations in such regions still exists (e.g., TELUM), allowing for more interesting policy analyses than manual forecasting methods currently in place. Few travel demand modelers are, as yet, well aware of such opportunities. Before discussing the merits of such a model, this report investigates UrbanSim, which lies at "the other end" of the land use model spectrum, in terms of data challenges and system complexity.

The larger MPOs in Texas have more resources for data collection and modeling than the smaller MPOs and they do not have to rely on TP&P or TTI as heavily for assistance. As discussed in the Chapter 2 of the report, all of the largest MPOs in the state use or are considering the use of an advanced land use model such as UrbanSim or TELUM. While implementing an advanced land use model is still challenging, these MPOs are better equipped to handle the challenge due to their larger staff size. For example, NCTCOG has over 100 staff members in its transportation department, which works closely with its research and information services department, where parcel data is collected and land use model inputs are gathered (NCTCOG, 2007). Its budget for modeling is also considerably larger as planning funds are distributed based on the transportation needs of the region. The following sub-sections describe key details of major land use model options, most of which are readily available to TxDOT, Texas MPOs, and others, but offer challenges implementation and/or limitations in specification.

[4] "Special generators are sites in the study area that do not exhibit 'typical' trip making characteristics that can be replicated using trip generation rates that are a product of the travel surveys. Examples of special generators include but are not limited to regional malls, airports, colleges, and hospitals. To properly estimate the total number of trips by trip purpose generated by a special generator, it is necessary to have additional information regarding the special generator so an appropriate trip rate can be used" (Hodges, 2007).

3.1.1 UrbanSim

UrbanSim is a disaggregate land use model that seeks to provide greater behavioral realism in land use markets than past LUMs. UrbanSim keeps track of real estate and households so that it can link households to individual dwelling units, and individual jobs to work spaces. It does not, however, link actual workers (and their associated households) to these jobs. These work spaces may be nonresidential or residential in nature, to account for telecommuters and home-based employment. When the households or jobs move, the units they occupy are considered vacant until another household or job occupies the space, or the unit is removed (Waddell, 2002).

UrbanSim is freely available, open-source software that can be run on Windows, Macintosh, or Linux operating systems. UrbanSim 4.0 requires Python 2.4.3 or 2.3.5, and MySQL 5.0 or 4.1.18, all of which are open-source. Previous versions of UrbanSim were written in Java. The data requirements and statistical methods for parameter estimation are described below.

UrbanSim requires the following attributes for each parcel in the region for a single (base year) point in time: land use, lot size, housing units, square footage of building space, year built, zoning, land use plan, assessed land value, assessed improvement value, and tax exempt status. The land use codes should be standardized across the region. If some data are missing, scripts are provided to impute land use codes, year built, and number of housing units. To impute housing units for a parcel, the housing units for the census block must be available.

Employment locations must be geocoded with the industrial classification and number of employees. The following environmental layers are optional: water, wetlands, floodplains, parks and open space, national forests, steep slopes, and stream buffers. The following planning and political layers are optional: cities, counties, urban growth boundaries, military, major public lands, and tribal lands. The traffic analysis zone layer is required and is used to connect UrbanSim with the travel demand model.

Data from the U.S. Census Bureau are required for synthesizing the household database: Public Use Microdata Sample (PUMS) Household and Person records and the Summary Tape File 3 (STF3) tables. Family households, non-family households, group quarters households, residential vacancy rate, and population are required for each block group. Other STF3 attributes required are vehicle ownership, income, number of workers, age of head of household, and race.

UrbanSim has built in parameter estimation modules, one to estimate parameters that are part of a multinomial logit (MNL) model, and another to estimate parameters from linear regression. MNL parameters are estimated using a maximum likelihood method. Linear regression parameters are estimated using the least squares method. The number of parameters to be estimated will differ for each region depending on such factors as the number of development types and employment sectors. The Eugene, Oregon application for the 1980 base year has 1,865 estimated parameters. For example, Table 3.1 shows the estimated parameters for the residential land share equation:

$$\mathrm{R}_i = \frac{\exp\left(\sum_{j=1}^{\text{all } j} \beta_j \mathrm{V}_j\right)}{1+\exp\left(\sum_{j=1}^{\text{all } j} \beta_j \mathrm{V}_j\right)} \tag{3.1}$$

where R_i is the fraction of residential land for cell *i;* β_j is an estimated parameter for development type, *j;* and V_j is an indicator variable for a development type, *j*. Essentially, Equation 3.1 is a binary logistic function ensuring that the residential share lies between zero and one.

Table 3.1: UrbanSim Estimated Parameters for Residential Land Share Equation

Coefficient	Estimate	Standard Error	T-Statistic	P-Value
development type 1	0.284	0.058	4.911	0.000
development type 2	0.329	0.061	5.422	0.000
development type 3	0.554	0.064	8.672	0.000
development type 4	1.338	0.066	20.434	0.000
development type 5	1.852	0.066	28.265	0.000
development type 6	2.055	0.096	21.368	0.000
development type 7	1.495	0.141	10.635	0.000
development type 8	1.130	0.366	3.092	0.002
development type 9	-0.671	0.087	-7.762	0.000
development type 10	0.887	0.162	5.486	0.000
development type 11	0.111	0.110	1.008	0.314
development type 12	-0.523	0.221	-2.367	0.018
development type 13	-1.172	0.309	-3.788	0.000
development type 14	0.127	0.219	0.580	0.562
development type 15	-0.366	0.277	-1.323	0.186
development type 16	-0.313	0.359	-0.871	0.384
development type 17	-1.010	0.088	-11.487	0.000
development type 18	-1.670	0.178	-9.398	0.000
development type 19	-1.863	0.244	-7.649	0.000
development type 20	-1.543	0.267	-5.780	0.000
development type 21	-2.689	0.473	-5.690	0.000
development type 22	-2.533	0.639	-3.964	0.000
development type 23	-1.649	0.057	-28.730	0.000
development type 24	0.000	0.000	0.000	0.000
development type 25	0.000	0.000	0.000	0.000
non_residential_sqft	0.000	0.000	-3.276	0.001
residential_units	0.008	0.002	4.781	0.000

Source: Table 2 of Waddell, 2002

The research team ultimately implemented UrbanSim for the 3-county CAMPO region, as described in Chapters 8 and 9. From this process, it became very clear that collecting reliable land use, building data, and many other types of information relevant to the LUM process is a challenge. For example, the team encountered many discrepancies between the parcel-level land use data created by the City of Austin and the same data set created by the Capital Area Council of Governments (CAPCOG). Also, CAPCOG's datasets did not include reliable information on non-taxable parcels (as the data come from county appraisal districts, which keep track of taxable/private parcels). Moreover, the geocoded employment data was really only reliable at the county-level, rather than the TAZ-level or grid-cell level required by UrbanSim (Kitten, 2007). The cleaning required of these datasets was beyond the scope of this project, but should be possible over the longer term, assuming adequate cooperation between public agencies and sufficient resources. To get a sense of how much time should be set aside to implement UrbanSim, the Puget Sound Regional Council spent 2 years on database development and refinement and 1 year on model estimation and testing (Borning et al., 2006). This was similar to the team's own experience, which included about 8 to 10 months in database development and 3 to 4 months estimating models and running simulations. Even after significant allocation of time and resources, several inconsistencies in the data remain.

3.1.2 PECAS

PECAS ("Production, Exchange and Consumption Allocation System") consists of two submodels, the activity allocation submodel and the space development submodel. PECAS runs the two submodels through discrete time-steps, recommended to be 1 year, and through spatial land use zones, recommended to not exceed 750. These recommendations allow for quick response between land development and land prices, while minimizing the burden on the operating system running the model (Hunt and Abraham, 2003). While PECAS is still evolving, the following discussion is based on the most recent documentation available, as directly obtained from the model developers. The activity allocation module represents how activities locate within the space and how these activities interact with each other at a given point in time. It finds the prices (at exchange locations) that clear all markets, as part of its allocation process. This module allocates the flow of commodities from production locations to exchange locations and then to consumption locations, using a three-level nested logit model: the highest level allocates the study-area total quantity of each activity among land use zones, the middle level allocates the quantity of each activity in each land use zones among available technology options, and the lowest level allocates the produced and the consumed quantities among exchange locations using two logit models. The space development module represents actions of developers in the provision of different types of developed space (where activities can locate) from one point in time to the next. The activity allocation module is aggregate in the sense that quantities of activities, flow of commodities and exchange prices are modeled at the level of land use zones, while the space development module has two forms: a disaggregate version and an aggregate version (Hunt and Abraham, 2007).

PECAS is a Java-based model, requiring Microsoft Office and ArcGIS. The required inputs described below are categorized as follows: input-output tables, coefficient elasticities, employment, labor force participation, trip length distributions, and commodity pricing.

Input-output tables for the base year are required to show the amount produced and consumed of each commodity in each zone, including import and export quantities. The tables contain information on three activities by three entities: industry, government, and households. Industry can be broken up by sector and floorspace type, (e.g.,"light industry in office floorspace"). Government can also be broken up by floorspace type. Households should be categorized by income level.

The elasticities of technical coefficients in the input-output tables are required for the base year, including floorspace consumption rates. The spatial distribution of employment by industry and occupation is required in the base year, and is optional for future years. Labor force participation rates by household type and occupation is required for the base year. Trip length distributions are required for each commodity in the base year. Commodities can be goods, services, labor, or education. Indications of the average price for each commodity, the overall variation in price across the region, and some indication of how the prices vary across the region is required.

Most parameters in the activity allocation model can be estimated by adjusting them to match a series of targets using the base year data described in the previous section as well as spatial distribution of employment for one or more years after the base year. Some parameters are estimated using conditional choice data and all parameters are adjusted during the overall calibration. Data required for calibrating the space development model include change in quantity of space by category for at least 200 zones from the base year to another year, zoning regulations, construction costs, and amount of vacant land. Parameters are adjusted in order to best match the calibration year development patterns. A statistical software with multinomial logit capabilities, such as Stata or LimDep, is recommended for PECAS sub-models' calibration.

3.1.3 TELUM

TELUM ("Transportation Economic and Land Use Model") is an FHWA-sponsored and Putman-provided ITLUP-based model that can be used to predict the short-term as well as long-term impact of a planned transportation policy on the land use pattern of a region. TELUM forecasts can also be used to examine transportation and land use consistency, which is a necessary input to air quality estimates. TELUM uses the current and lag year residential, employment, and land cover data of a region to forecast the future land use patterns of the region. TELUM consists of three sub-models, namely TELUM-EMP, TELUM-RES, and LANCON, designed on the basis of Putman's DRAM®, EMPAL® and LANCON models. Land use forecasting is generally done in increments of 5 years, with each increment beginning with the execution of the TELUM-EMP sub-model, which forecasts the future location of the region's employment. Next, TELUM-RES produces a residential location forecast for all households in the region. Finally, LANCON computes the land consumption in all zones based on the demand for land for employment and residential purpose in each zone and the supply of land in that zone.

TELUM requires Windows 2000 or later with Service Pack 5 and the Microsoft Office package. To use TELUM's mapping function, MAP IT, ArcGIS is required. Ten types of data are required to run TELUM: employment, household, land use, employment to household conversion ratios, employee per household by income, unemployment rate, net commutation rate, total reported jobs per region, projected zonal totals, and travel impedance. Each type is discussed in more detail below. Lagged year data are required for calibrating the model. A five-year lag is recommended by the model's developers as best insuring model reliability.

The zonal employment data must be divided into a minimum of four employment classes and a maximum of eight employment classes. The zonal employment data (by the employment class) is required for both the base year and lagged year. The base year zonal household data must be organized into a minimum of four categories and a maximum of eight categories (preferably based on income). The total number of lag year households in a zone is also required.

The total land area of each zone must be classified into useable land and unusable land. The useable land must be divided into the land used for basic employment, land used for commercial employment, residential land, and vacant developable land.

An employment-to-household conversion ratio must be specified for each one of employment and household classes. TELUM uses the conversion ratio for allocating the residential location of employees. If the data is not available then an even distribution is assumed. The average number of employees living in each of the regions household by income group must be specified. A default value of 1.00 can be used if the data is unavailable. Specifying the unemployment rate helps TELUM account for persons currently unemployed. An assumption is that unemployed people made their location decision while they were employed. If data is unavailable, a default value of 0% is used.

Also required is the net commutation rate, a measure of work trips commuting in and out of region. A value of less than 1.00 indicates net inbound and greater than 1.00 indicates net outbound. Total reported jobs in the region should be specified if available. If this data is unavailable then the number of reported jobs is assumed to be equal to the number of employees. The total projected population of the region and projected strength of employment classes of the region must be known for the forecast years. The average population per zone must lie between 3,000 and 10,000. The zone to zone travel time or travel cost data are also required.

In the TELUM model there are certain parameters whose values depend upon the regional data and have to be determined by the calibration process. In TELUM-EMP, five parameters have to be estimated for each employment type. In TELUM-RES, 6 parameters (plus the number of household types) have to be estimated for each household type and 19 parameters have to be empirically determined in LANCON model. The calibration process involves fitting the TELUM equation to the data of the region. The better the fit of the model the more reliable are the forecasts. The standard multiple regression techniques cannot be used to estimate the parameters of the model as the equations of TELUM are non-linear and the regional data may not be normally distributed.

CALIBEL, a computer program within the TELUM software, is used to estimate the parameters. Gradient search is used for the estimation of the parameters. The partial derivatives of goodness of fit criterion with respect to each parameter are calculated and these derivates determine the search direction. The goodness of fit measure used in CALIBEL is the likelihood function $L = \sum_i \left(N_i \ln \hat{N}_i\right)$, where N_i is the observed value of the i^{th} parameter and $\hat{N}_i$ is the parameter's estimated value. The "best fit" occurs if $N_i = \hat{N}_i$. L_b denotes the "best fit" likelihood value. The "worst fit" occurs when all values of dependent variable are estimated by the mean of that variable. L_w denotes the "worst fit" likelihood value.

$$\text{Relative goodness of fit } (\varphi) = \frac{L - L_w}{L_b - L_w}$$

Typical results obtained when fitting CALIBTEL give φ in the range of 0.70 to 0.90.

The research team's implementation of TELUM for the Austin region has found TELUM to be simple to use—once some initial challenges are overcome. TELUM is inflexible when it comes to formatting the input data and allowing data to fall outside of hard-coded bounds.

3.1.4 TRANUS

TRANUS (TRANUS 2007) runs on Microsoft Windows and provides a user shell for input and model running, although it still uses a command line for report generation. Data can be imported from spreadsheets, and transportation network information can be imported from a GIS program. The resulting reports are in a tab delimited format that can be exported to a spreadsheet or GIS program.

TRANUS is essentially two integrated models—a transportation model and a land-use model, each of which could, theoretically, be run stand-alone. Both models require the study area be divided into zones, the size and number of which are completely user-defined. The centroid for each zone must be specified, along with transportation links between zones. TRANUS automatically calculates the usage and disutilities of the transportation links to find an equilibrium distribution of route flows and to determine the transportation cost inputs into the land use location model.

Transportation demand is created by a trip generation model (either based on the land use model, or from origin-destination pairs directly created by the user). If using a land use model, the user must specify trip categories (e.g., low income commuting trips, or service trips, or freight trips). Trip categories include information regarding value of travel time, value of waiting time, transportation elasticity parameters (i.e., minimum and maximum usage if disutilities were 0), and mode choice and path choice elasticity and scaling factors. As with most parameters in the TRANUS model, the user must specify appropriate desired numbers, with little reference to standard values.

The transportation model also requires information regarding the transportation links and transportation modes. Transportation links can be used by multiple operators (private vehicles, public transportation, pedestrians, etc.), and links can be combined into public transportation routes. Each link requires data regarding speed, capacity, and speed reduction percentage as a link reaches capacity. Mode information includes fares, capacity, energy costs, target occupancy, operator costs, transfer times, and user desirability (divided by trip category given above).

The land use model requires less information than the transportation model, but the information and parameters required are even more open to interpretation. The land use model requires that the user create economic “sectors.” The number of these is completely up to the user, but at a minimum should include sectors of land, basic (exogenous) employment, service employment, and workers. A more complete model might include division of workers into high, middle and low income, the inclusion of multiple industrial sectors, and the inclusion of non-worker dependents. Each sector must include data for elasticity, price scale, attractor factor, and a logit scale. A value of 1 is appropriate for each of these parameters, at least as an initial estimator although basic employment and land should have an elasticity parameter of 0, reflecting the fact that no new land or basic industry can be induced.

Sectors interact through a sector interaction table that lists the minimum and maximum demand from one sector to another. For example, basic industry will demand a certain number of workers per unit of output, and will use between a minimum and maximum amount of land per unit of output depending on the price of land. Substitutions are also allowed if the user specifies the substitution along with an elasticity and logit scaling parameter, and a penalty value. Sector placement decisions can also be influenced by user-specified attractor weights (e.g., service employers may be most attracted to locations near other service employers, and somewhat attracted to basic industry employers).

Sectors interact with transportation categories via a sector/trip generation table. This table is relatively straightforward, except that it requires some mathematical manipulation in case there is a difference in terms of units of time or volume. For instance, freight may be measured in yearly truckloads, while commuters are measured in terms of daily travel.

Finally, the user must specify the initial base case. For each sector, the user must specify the initial quantities of each sector in each zone, along with economic data representing the price of the sector and the value added by the sector. The user must also specify exogenous production (production not induced by the model, and so a growth rate must be specified manually) and induced production factors. Note that it may also be the case that the user must specify the initial quantities and inter-sector demands such that the system is already in equilibrium between quantities and demands (for example, the total number of service employees given should reflect the number of service employers multiplied by the demand of service employers for service employees).

Like any meta-model, TRANUS is modular, which reduces the amount of data elements required for a simple model. However this comes at the price of requiring the user to specify proper parameters for the interactions between the data elements. Many parameters may be correct at certain standard values, but true settings would require a calibration exercise, where the model is run using different parameters on a known base year, and the results are compared with a known final year.

Table 3.2 shows the number of parameters that are required or optional by category. Parameters listed as “Easy” are retrievable (e.g., transit fares) or directly calculable (e.g., value of travel time). “Medium” refers to something that is calculable, but the calculations may take significant work (e.g., model calibrations). “Default” values are given and do not need to be changed. “Hard” refers to parameters for which a default value may be given, but likely should be changed. “Optional” parameters need not be given. Depending on the characteristics of the model (i.e., number of modes), the number of parameters will differ. For example, if there are 5 sectors, then the number of “land use inter sector” parameters will be 250 or 10×5×5. If there are 100 zones and 8 economic sectors, approximately 7,500 parameter values are needed.

Table 3.2: TRANUS Parameters

	Number of Parameters (by ease of estimation)				
Category	Easy	Medium	Defaulted	Hard	Optional
Transport Categories (per category)		2	5	3	
Transport Link Defaults (per link type)			4		1
Transport Link (per linktype&vehicle) (also must do for individual links if individuality is desired)	3	1	2		1
Transport Modes (per mode)			1	2	
Transport Operators (per mode)	13	4	1	1	
(per mode&category combo)			2	1	
Transport Transfers (per mode/mode)	1				
Transport Routes (per pub transit route)	2		2		
Land Use Sectors (per sector)			4		
Land Use Inter Sector (per sector/sector)		2	1	6	1
Land Use Sector/Transport Category (per sector/category)	5				
Land Use Economic Data (per sector/zone)	4	2	1		1

Source: TRANUS, Modelistica

3.1.5 Linking to a Model of Travel Demand

The three-step travel model used by TxDOT is relatively simple, and any of the land use models discussed in this review (UrbanSim, PECAS, TELUM, and TRANUS) can provide the necessary inputs. Of course, the relative simplicity and spatial coarseness of the existing travel model means that many estimates provided by an advanced land use model may not be fully exploited. For example, parcel or grid cell land use details will have to be aggregated to traffic analysis zones and variations in person and household characteristics (such as income) will be reduced when running the travel model. Nevertheless, the land use models are designed with such linkages in mind, and should pose no problem, in this regard.

All four land use models examined in this report can be run on a Windows platform. TELUM is the easiest of the four to implement, because the input data is the most readily available, it requires no external software other than Microsoft Excel (and ArcGIS for mapping is optional), and it has an internal parameter calibration tool. However, iterative feedback between TELUM and a travel demand model is not possible. Such feedback is critical to capturing the transportation-land use interaction.

In addition to common software tools like Microsoft Excel and Access, PECAS requires a Java platform. Several such platforms are freely available. UrbanSim requires ArcGIS and several free software packages (namely MySQL and a Python platform). Both UrbanSim and PECAS require that someone on staff be familiar with the relevant programming language (Python or Java, respectively); otherwise, a consultant should be hired for assistance. TRANUS has a built-in travel model, eliminating the need for external travel models like TransCAD. If desired, its land use model could be run stand-alone and integrated within a TxDOT travel model.

Among all land use models felt to be relevant for transportation planning, UrbanSim and PECAS are the most data intensive. PECAS requires data on economic interactions (e.g., goods trade and worker movements) across zones, thanks to its spatial input-output model framework. UrbanSim requires geocoded employment data assigned to parcels. Full model integration that captures the transportation-land use interaction can be achieved with both UrbanSim and PECAS. Both models run the land use model every year from the base year to the forecast year, and run the travel model every 1 to 5 years, allowing the land use patterns to be inputs into the travel model and the accessibility measures to be input into the land use model. TELUM can also be integrated with a travel model through some manipulation (i.e., run TELUM with a short-term forecast year, then run the travel model, and iterate until the desired forecast year is reached).

With adequate budgetary and personnel resources, calibration and application of an advanced land use model should be possible for any region. If a parcel-level or small grid-cell-level model is used, at least one year should be set aside for data development and a second year for model estimation (with significantly less time needed for TELUM or similar applications). For regions that contain multiple county and local jurisdictions, inter-jurisdictional cooperation (via a Council of Governments or MPO, for example) and standardized data sets will be critical. Sufficient funds should be set aside for hiring support personnel, even in regions where staff is technically competent. UrbanSim has a growing community of users, offering alternative forms of support through a list-serv and “UrbanSim Commons,” a website that allows users to share experiences. While TELUM and TransCAD’s new land use model (based on the same key equations and concepts as TELUM and thus ITLUP) are the easiest to implement, they offer less behavioral detail. UrbanSim, PECAS, and TRANUS each require a significant investment of time and resources for implementation; however, each can be modified (typically by reducing the number of modules and parameters used, and very possibly at the expense of behavior realism).

Among all these, only TransCAD’s land use model is thoroughly documented, and can be reproduced by the research team. However, TransCAD’s land use model does not yet include a calibration component. As part of the implementation portion of this research project, the team will convert its MATLAB-based gravity-land use model (very similar to TransCAD’s tool, and thoroughly documented, as described at http://www.ce.utexas.edu/prof/kockelman/G-LUM_Website/homepage.htm) into a Visual Basic program. This will allow open-source access, complete documentation, a relatively simple modeling structure, parameter calibration, and application on all computers running Microsoft Windows operating system.

Chapter 4. Review of Suitability Models and their Potential Contribution

Along with modeling land use and transportation in an integrated fashion, consideration of environmental factors in land use-transportation decisions is becoming more and more important. Also, there has been growing interest in and need for spatial disaggregation in land use transportation modeling efforts, which makes the use of geographic information systems (GIS) highly attractive—and practically necessary. These models require a great deal of data and generate many spatial results which need to be communicated to a broad range of experts, city leaders, and the general public, via visually appealing and understandable presentations (Ducrot et al., 2004). Land suitability analysis is a technique that has long been used for the ecologically sensitive allocation of land resources and planning (McHarg, 1969). With the advent of GIS, this technique has evolved and been integrated into existing GIS platforms; it may become a major contributor to the process of integrated land use-transportation modeling. This chapter presents relevant aspects of the suitability analysis literature, including applied processes and available software tools. (These include ArcGIS, ArcView spatial analyst tool, *What if?*, UPlan, and Anjomani's method.) After briefly discussing meaningful opportunities for suitability analysis in modeling land use and transportation, the chapter introduces land use suitability analysis and offers a review of these tools and their usefulness for integrated land use/transportation planning. Section 4.1 reviews the literature on related software, and relies on default data. Section 4.2 further evaluates SA software, for ease of use and applicability, while applying Waco's McLennan County data sets in Appendix D provides additional, complementary information, separately for both sections.

4.1 The Role of Land Suitability Analysis in Land Use/Transportation Modeling

There are at least two distinct contributions suitability analysis (SA) can make to integrated modeling of land use-transport systems. First, such analysis can readily identify environmentally sensitive areas meriting preservation and/or other special treatments. Provisions can then be made for restricting their use in the modeling process, in order to avoid (excessive) development of sensitive areas. Second, suitability analysis illuminates opportunities for targeting locations to accommodate regional growth (Harshorn 1993 and O'Sullivan 2003 & 2007). In this way, it provides a meaningful marriage of mathematical modeling and community visioning, two very distinctive yet complementary processes that focus on a region's future (Lemp et al. 2007).

For the purpose of integrated land use transportation modeling, SA can be used in two ways—either integrated with land use transportation models or as a land use modeling tool for smaller and mid-sized metropolitan areas. Integration of land use transportation models with suitability analyses promises great payoffs for MPO-level land use and transportation planning. Most land use models have to begin with a variety of assumptions, such as the distribution of certain employment types. Locations of future employment centers for the study areas identified by SA techniques can be utilized in the land modeling process. SA ultimately is a type of land use model (LUM), and as such requires demographic and economic projections (e.g., population and employment) for allocation to the study area.

4.1.1 Suitability Analysis

In more recent years, local and regional efforts have begun to recognize the effect of land use on the quality of the environment. This recognition coincided with an attempt to improve land use, and to reduce environmental hazards and land use conflicts. Reducing these conflicts can be achieved by land use suitability analysis, which determines the best land use suitable for a land area. Several suitability analysis techniques have been used in the past. Hopkins (1977) reviewed several SA methods, including the Gestalt method, generating suitability maps by mathematical operations such as the ordinal, linear and non-linear combination methods, as well as overlay map techniques.

Suitability analysis is the traditional method for determining land use allocation, and is a major stage in the Chapin and Kaiser (1979) seven-stage and Kaiser et al. (1995) five-step land-use planning process. It is used to identify the most suitable places for locating future developments for each land use under consideration. This process is done for one land use at a time. It relies upon overlaying maps of physical, locational, or institutional attributes, such as soil type, vegetation, slope, or distance from highway, etc., to calculate suitability. The resulting suitability score is expressed on a single scale that can be simply high, medium, or low; or a more continuous numerical scale. SA's key output is a land suitability map that shows vacant and under-utilized land well suited for future development, while protecting environmentally sensitive areas, mitigating natural hazards, and avoiding, for example, reduction in impervious cover. For regional land use/transportation planning, suitability analysis also assists in allocating new growth to areas scoring highest in suitability across various land use types.

Let us briefly discuss how land suitability analysis works and describe its main elements before reviewing some of the models. If the suitability of an area for a particular use (light industrial) is assumed to be a linear function of several factors (such as soil condition, slope, accessibility, etc.), then the suitability of a location *i* and for land use *j*, S_{ij}, can be represented as (Anjomani, 1984):

$$S_{ij} = b_1 F_{1ij} + b_2 F_{2ij} + b_3 F_{3ij} + ... + b_k F_{kij} \tag{4.1}$$

or

$$S_{ij} = \Sigma_k b_k F_{kij} \tag{4.2}$$

where *F*'s are ratings of suitability of each factor according to the degree of its effects—positive or negative—on each of the selected land uses, *j*, in locations *i*, and *b*'s are coefficients measuring importance (weight) of the *k* selected factors. Suitability scores for each land use and location, S_{ij}'s, can then be calculated by multiplying coefficients and their associated factor scores. They are summed up as shown in equations 4.1 and 4.2.

Suitability factors (*F*s) are the relevant characteristics for determining the relative suitability of different locations or parcels for different land uses. The factors to be used for land use are specified based on theoretical knowledge of variables affecting land use locations. Soil quality and slope, access to water/sewage and other utilities, proximity to other land uses and highway intersections, and quality of vegetation are good examples of such variables. Factor weights essentially are numerical scores indicating the relative importance of the different factors; these weights are used to determine the relative suitability of different locations for different land uses. Thus, for example, if the slope has a factor weight of 2 and the soil has 1, the indication is that slope is twice as important as soils in determining a site's suitability for certain land uses, such as parks and playgrounds (Kaiser et al., 1995).

In order to determine factor weights, one needs to analyze factor types and factor ratings. Factor types are the set of possible values for a particular factor. For example, there could be three different slope categories defined for a particular location within the study area: less than 10%, 10% to 15%, and more than 15%. Factor ratings are numerical values indicating the relative suitability of locations for locating particular land uses, such as open space or commercial buildings. Thus, for example, factor ratings can be specified for each of the three slope types listed above to indicate their relative suitability for each land use type. The rating for each factor could be on five-point or other scale. Zero values can also be used to exclude development from areas corresponding to a particular factor value (e.g., all areas with slopes over 15%). When regression techniques cannot be used, Anjomani (1984, 1992, and 2005) has suggested the Delphi technique or a multi-criteria evaluation technique, such as Analytical Hierarchy Process (AHP) (Saaty 1982, 2007), for determining factor weights.

The suitability score is a numerical value indicating a location's overall suitability for a land use type when all factors are considered. The suitability score for a particular zone is determined by multiplying the zone's factor rating for each factor by the corresponding factor weight and summing up these products. Later, suitability scores are computed for all zones and all land uses.

4.1.2 Review of Available Software and Applied Processes

Technological advances in computerized mapping and geographic information systems have given land use planners a more efficient and effective way of handling and analyzing large amounts of spatial data. Similarly, there have been advances in the process and application of suitability analysis and software developed for use of such techniques. SA models such as ArcGIS, ArcView Spatial Analyst tool, *What if?*, UPlan, and Anjomani's method are relevant to this work and thus are briefly reviewed in this section. A more detailed review of the software through a series of actual applications (using Waco's McLennan County) is provided in Section 4.2, including a description of software strengths and weaknesses.

ArcGIS

A GIS is a system for management, analysis, and display of geographic knowledge, which is represented using a series of information sets. It also is a spatial database containing datasets that represent geographic information in terms of features, rasters, topologies, networks, and so forth. In the context of land suitability, GIS helps the user determine what locations are most/least suitable for development. GIS's spatial function such as overlays, buffers, measures of contiguity, and the like, along with the facilitation of data manipulation, analysis, and display make GIS a necessary tool for every land use analysis (Goodchild, 2000 and Anjomani, 2005a). GIS in general and ArcGIS in particular could be used for a variety of spatial analysis, including land suitability analysis. However, because the general GIS software such as ArcGIS is not particularly developed solely to perform suitability analysis, they will require more time and more advanced levels of software knowledge to apply it to this specialized use. ArcGIS is commercially available GIS software through Environmental Scientific Research Institute (ESRI), which is widely used in a variety of public agencies and private corporations. This GIS software helps in representing, managing, organizing, storing, displaying, analyzing, and reporting spatial information.

ArcMap is the central tool in the ArcGIS software suite. It is used to perform the final procedure of representing the best land use pattern for the area. The process of land suitability analysis in ArcMap using ArcGIS involves the following tasks (also described in Appendix D):

1) Identify the land uses to be studied.
2) Define the boundary for the study area.
3) Select a grid size (e.g., 150X150 meter) and prepare a grid/fishnet for the study area identified by the cells using ArcInfo.
4) Convert or reclassify data into the required form for further GIS analysis (e.g., converting raster data sets into shapefiles for vector analysis).
5) Determine the suitability factors to be considered for the analysis (e.g., soil, slope, vegetation, distance to expressways and airports, etc.).
6) Identify the land uses to be considered for the study such as open space, single-family, multi-family, and low commercial.
7) Specify factor ratings.
8) Specify factor weights for each suitability factor.
9) Buffer and overlay "intersection" commands are used at this stage to assign the factor ratings for the study area.
10) Once the weights are obtained, each of the factor ratings are multiplied with these weights to obtain reliable final figures for each factor and all the land uses.
11) Suitability for each land use is the sum of the adjusted ratings for all the factors.
12) Transform the outcome into a suitability map by choosing a set of patterns to represent the different degrees of suitability by assigning best land use percent for each cell based on the demand (e.g., open space 30%-87 cells, low residential 25%-71 cells, such that it totals 100% land use).
13) The final output map is a composite map symbolizing land uses in a color scheme.

ArcView's Spatial Analyst Tool

The research team reviewed both ArcGIS and ArcView Spatial Analyst tools separately. ArcGIS software helps in developing a suitability map with vector data as an input, whereas ArcView Spatial Analyst helps analyze spatial relationships, build spatial models, and perform complex raster operations only with raster data as input.

There are many similarities between the two, with one exception being that Spatial Analyst is an ArcGIS extension for use in ArcMap. (ArcMap is an ArcView function, and ArcView is one of ArcGIS's three products.) ArcGIS does not include spatial analyst by default, which has to be purchased as an extension.

As opposed to ArcGIS's spatial analysis tool, ArcView's tool is more specialized, and designed to perform various forms of spatial analysis including suitability analysis. One of the important tools in ArcView's version is Model Builder, which currently is being widely used as an industry standard. Model Builder processes wizards and diagramming tools to construct various types of spatial models, including suitability analysis. The models are represented as process flow diagrams. This technology provides both beginning and advanced users with a set of easy-to-use tools for building various types of spatial models.

ArcView's Spatial Analyst provides a broad range of powerful spatial modeling and analysis features. Similar to ArcGIS's tool, ArcView's Spatial Analyst Extension enables the user to create, query, map, and analyze cell-based raster data, and to perform integrated vector-raster analysis, derive new information from existing data, query information across multiple data layers, and fully integrate cell-based raster data with traditional vector data sources. These tools form the foundation for all spatial modeling and geoprocessing. Of the three main types of GIS data (raster, vector, and tin), the raster data structure provides the richest modeling environment and operations for spatial analysis. GIS capabilities for spatial analysis overcome the drawbacks of a paper–map-based overlay approach. This system enables planners to create and modify SA results, making the best use of available data.

In addition to storing, retrieving, and displaying spatial data, a geographic information system enables the user to create buffers, overlays, and intersections, and perform proximity analysis, spatial joins, map algebra, and other analytical operations. By using this tool, we can derive new information from the existing data, analyze spatial relationships, build spatial models, and perform complex raster operations. In the context of land suitability analysis, GIS spatial analyst helps the user determine what locations are most/least suitable for development. In this way, the results of GIS spatial analyst tool can provide support for decision-making.

The nine steps in ArcView Spatial Analyst are as follows (CGIA, 2005):

1. Conversion of data into raster format.
2. Ratings are assigned using buffer and reclassify commands.
3. Weighted overlay command is used to assign weightings for each factor.
4. The above steps are repeated for all the land uses separately.
5. To avoid repetition of the above steps, a model is created using model builder.
5. Create a model.
6. Run the model.
7. Analyze results.
8. Refine the model as needed.
9. The result will produce separate suitability maps for each land use.

(More details of this process can be found in the Appendix D1 of this report.)

While ArcView's program can be somewhat inflexible, it offers a rather quick way for ArcGIS users to examine a variety of land use scenarios, quickly changing criteria to view results.

What if?

What if? is an interactive GIS-based planning support system (PSS) (http://www.what_if-pss.com/w_intro.htm). It uses existing GIS data to be used in conducting land suitability analysis, projecting future land use demand, and allocating the projected demand to suitable locations. This is a software program derived from GIS utilities in which certain scenarios were customized with pre-selected parameters and criteria of suitability features in a compact package that does not require the running of GIS in the background. *What if?* was developed with Microsoft's Visual Basic and the Environmental Sciences Research Institute's (ESRI) Map Objects GIS component software. Richard Klosterman (1999), *What if?*'s primary developer, describes the software as follows:

> "As its name suggests, *What if?* is an explicitly policy-oriented planning tool that can be used to determine *what* would happen *if* clearly defined policy choices are made and assumptions concerning the future prove to be correct. Policy choices that can be considered in the model include the staged expansion of public infrastructure and the implementation of alternative land use plans or zoning ordinances. Assumptions for the future that can be considered in the model include different population and employment trends, assumed household characteristics, and anticipated development densities. *What if?* projects future land use patterns by balancing the supply of, and demand for, land suitable for different uses at different locations. It does this by providing three integrated model components: (1) a suitability option for developing land suitability scenarios which determine the supply of land; (2) a growth option for creating growth scenarios which determine the demand for land; and (3) an allocation option that projects future land use patterns by allocating the projected land use demands to the most suitable sites. Alternative visions for an area's future can be explored by defining alternative suitability, growth, and allocation scenarios."

What if? provides four options that allow the users to customize the program to the kinds of data that are available for the study area. These options include suitability analysis option, land use analysis option, land use/population analysis option, and land use/population/employment analysis option. The suitability analysis option allows the user to produce suitability maps showing the relative suitability of different areas for accommodating future land uses. This can be used only with GIS information. The land use analysis option allows the user to produce maps showing the projected land uses for up to five projection years. The land use/population analysis option provides information regarding land use and population for the sub-areas. The land use/population/employment analysis option allows the user to produce all of the projection information for the population and land use data option provided by the land use analysis option (Klosterman, 2007).

As elaborated on in Appendix D of the report, *What if?* follows the following steps:

1) Define the suitability scenario option (e.g., conservation or suburbanization)
2) Specify importance weights
3) Specify suitability ratings
4) Specify land use conversions
5) Compute suitability scores
6) Create new suitability scenarios
7) View suitability analysis results
8) Compare suitability maps
9) View suitability report

10) Evaluate suitability scenarios

While *What if?* may seem fairly easy to use and requires no additional software, like other SA methods, it neglects market clearing behavior, profit maximization, and other behaviors that are likely to be at play in the land development process.

UPlan

The UPlan Urban Growth Model developed by Johnston et al. (2003) (http://www.des.ucdavis.edu/faculty/johnston/Uplan_Demo_files/frame.htm) is a rather simple rule-based model developed by based on a platform of ArcView GIS. UPlan provides land use evaluations and change analysis based on general land use plan, population and employment, projections, characteristics of housing, and other user defined conditions. It allocates increments of urban growth in user specified discrete categories consumed in future years. UPlan can help decision-makers protect important open spaces by allowing them to consider the consequences of land-use planning on natural habitats, agricultural lands, reducing densities, and securing conservation easements.

UPlan was designed to simplify the urban growth planning process by creating a simplified user interface. It uses GIS shape files as a basis for data input. UPlan users begin by assembling information about the subject area. This includes parcel size, dimensions, and location in the form of shape files (database files developed to hold spatial information based on any number of coordinate systems).

With UPlan, users collect additional shape files for the location of roads, infrastructure such as utility lines, and topographic information. A number of software tools are available to build shape files including some freeware applications such as Christine or Grass GIS. The user would then create data bases or data tables that have relevant information about the objects in the shape files. This would include the land use type for the parcels, the population by parcel, traffic counts on the roads, capacity of the utilities, and elevation of the topography lines. These tables would be readily interchangeable between computer platforms and would likely be tied into other departments of the local planning organization. They can be built using Microsoft Excel, Microsoft Access, or a text editor program.

UPlan allows the user to load these data files into the application. Once loaded the user defines the parameters regarding expectations on the changes in the supply of land given the local legal restrictions and cost associated with developing new land or redeveloping land for a different use. The model allows the user to input expectations on the demand. In addition to using general plans to determine areas that will develop, UPlan uses population and employment projections, information on other important features (e.g., presence of highways and soil slope), and user-defined constraints for development (e.g., floodplains and publicly held land). UPlan is written in the Avenue programming language to run in ArcView (ESRI, Redlands, CA), and converts user-specified data inputs into grids that are then used to form new grids which forecast patterns of future land use. Each grid cell is referenced by its x and y coordinates. The resulting layers are based on attraction and exclusion grids, the general plan, and areas of existing urban development. Attraction grids are sites that are preferentially developed (e.g., near roads and freeway ramps), while exclusion grids are comprised of areas where development is restricted (e.g., parklands and water ways etc.).

Some of UPlan's input data requirements are as follows (Johnston et al. 2006) (and as described in Appendix D1):

1) Choosing model type (single-county or sub-area model)
2) Defining relevant variables
3) Setting up resolution and units
4) Residential inputs (% Res High, % Res Low)
5) Employment inputs (% Empl High, % Empl Low)
6) Choose attractions (e.g., freeways, airports, and boundary layers)
7) Choose discouragements (e.g., parks and watershed management areas).
8) Definition of mask layers (i.e., lakes and streams), which are not developable.
9) General plans and farmland are reclassified in the UPlan model, and are represented by codes (e.g., Unclassified = 0, agriculture = 1, industry = 2).
10) Finally, the model is run, generating a recommended allocation map for each land use group, an allocation map for each land use type, and an allocation map with all seven land use types.

UPlan provides an assortment of prepackaged reports to be used for later analysis, as well as GIS maps displaying the model-suggested changes in land use. Additionally, UPlan's flexibility allows the user to define as many future time periods as needed, for purposes of suitable land use pattern forecasting. UPlan also has the ability to incorporate more application tools. Once a scenario is loaded into UPlan, the user can take advantage of ArcView tools, such as Spatial Analyst, Geo Statistical Analyst and 3D Analyst. UPlan requires fairly extensive input parameters and data. Its flexibility and customization advantages could become liabilities for planning departments that are limited in labor and financial resources. Its consideration of the demand side of land use is based on growth scenarios, which limits its applications. UPlan gives the user more options, yet requires more training and resources to manage. It may be appropriate for larger municipalities and government organizations with more resources. UPlan can help decision-makers protect important open space by allowing them to consider the consequences of land-use planning on natural habitats and agricultural lands.

Overall, UPlan is relatively easy to use, customizable, and useful in addressing a variety of land use and planning questions. However, the rules that guide it—as in nearly all SA modeling approaches—are quite simplistic, relying on subjective weights and scores, rather than data-calibrated models. And it is somewhat less complicated to run than *What if?*.

Anjomani's Method

The Anjomani Method (AM) is a planning support system (PSS) that integrates an optional optimization model with GIS, and the knowledge and expertise of the planner-analyst. The planner-analyst also incorporates the community's wants and values into the planning process. The method attempts to provide a comprehensive approach to environmental and land-use planning on a large scale. This method aids decision makers in overcoming the negative effects, costs, and uncertainties in the planning process for relatively large areas.

In contrast to the four models reviewed above, the Anjomani method provides a composite result (a land use plan). Additionally, while the other models require that the user define values for ratings and weights, Anjomani's method provides more rigorous and defensible guidance for deriving these.

Anjomani (1984) has argued that the overlaying map technique deals only with the supply side (i.e., what land is good for what use) without considering the demand side (i.e., how much of each use is needed). As a result, the analysis normally shows oversupply of some uses and undersupply, or even a total lack of, other uses. In order to overcome this problem he has suggested adding analytical methods, such as optimization tools. In particular, an optimization model that recognizes demand for land uses and maximizes total net benefits (or minimizes total negative effects) is quite appealing when compared to a rule-based or other, conventional method. Furthermore, the total cost figure of the optimization model provides a measure for planner guidance in selecting alternatives or evaluating suggested changes at the later stages in the development process, based on observed cost increases or decreases. Land-use planning is a special *process* that a planner/analyst goes through using tools, intuition, and knowledge. Deriving one-shot results from any particular tool or model for land-use/environmental planning is unrealistic. A key part of the PSS is a planning process that employs the suggested analytical tools (e.g., constrained optimization) in different planning stages, repeatedly and iteratively (Figure 4.1). In this way, the results from the model are not final by themselves and gradually improve through the application of the process.

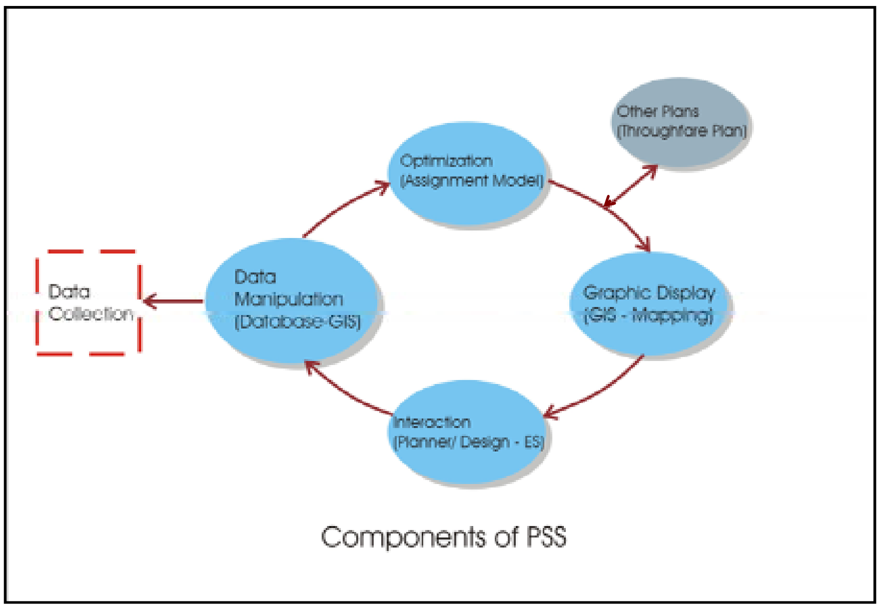


Figure 4.1: Components of PSS

The process for use of the model in large-scale planning endeavors, after following general steps, similar to those described in any of the above models, is summarized below and offered in greater depth in the Appendix D of this report (along with illustrated examples of the other modeling processes).

- Stage 1: Data manipulation and preparation of final ratings. Application of the optimization model to derive the first optimum solution.
- Stage 2: Preparation of a general thoroughfare plan.

- Stage 3: Overcoming the problems in the first optimal solution by considering the thoroughfare plan, design, and community's input.
- Stage 4: Checking the plan with respect to the problems discussed above.
- Stage 5: Overcoming the problems by design inputs and by incorporating the intuition and expert knowledge of the planner.
- Stage 6: Manipulating the data based on improvements (removal of the related data for cells that do not need improvements) and running the optimization model again.
- Stage 7: Repeating Stages 3 through 6 until the results become satisfactory.
- Stage 8: Further refining the results and applying the process to major sub-regions and different jurisdictions of the study area.

In this method, it is important to point out that the conceptual planning and design process begins at Stage 2 of the process and goes hand-in-hand with subsequent stages. The conceptual planning and design process starts by identifying the most important planning elements and then deciding their approximate locations. Simultaneously, the thoroughfare plan and plans for other infrastructure are developed. The process is then extended by stepping down gradually to less important elements.

In Stage 2 of the process, the planner-analyst develops a rough thoroughfare plan (the future alignment of highway network) for the whole region. The planner starts with a preliminary consideration of land use and transportation. GIS features, like buffer zones and the analysis of scenarios, can be helpful in this step; however, the planners' intuition, design principles, and community's wishes are the driving forces behind both selection of overall location of important land use elements and the thoroughfare plan.

The method's overall net benefit measure serves as an index in the planning process that allows planners or decision makers to know the relative cost savings magnitude as a result of planning decisions. As such, this method can be a useful tool for planners and decision makers in examining the multitude of land development or spatial planning options that exist. The planner can objectively analyze the options and determine the ones that best resolve conflicts or make the most of available opportunities. Options that avoid heavy environmental costs can also be identified. This information can be combined with other relevant information to make informed and intelligent planning decisions which not only respect short term economic concerns, but also long term environmental conservation and community goals as well.

While the Anjomani methods seeks to optimize in allocation, rather than simply allocating demand for each use type in turn, it is a more involved process, requiring more time for derivation of results.

4.2 Results of Waco, Texas Comparisons using SA Methods

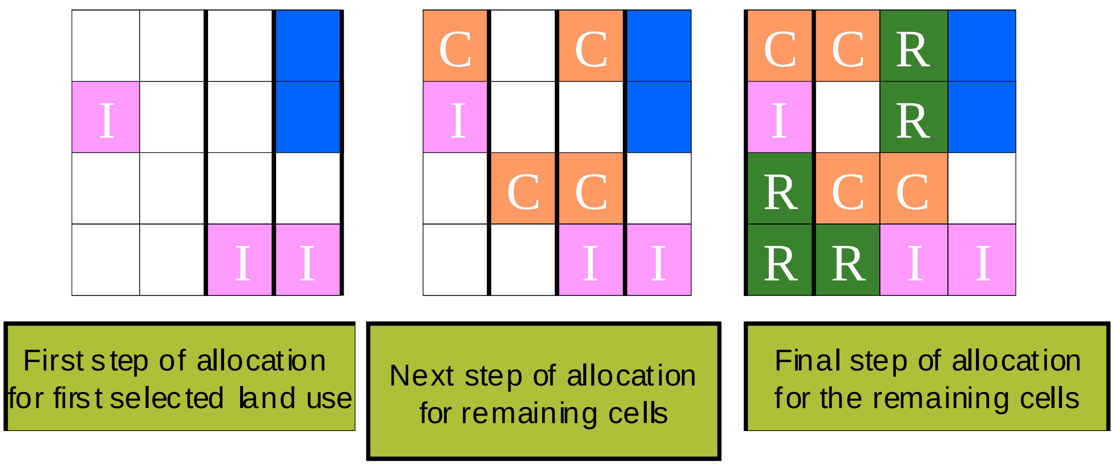
As mentioned above, the goal of this chapter is to review all relevant aspects of suitability analysis literature including applied processes and available software. As described above, the team reviewed various software packages through an examination of related literature and by working with the default data provided in each package.

In order to review the packages in a more detailed manner, the team selected Waco's McLennan County—a small Texas region—and used existing data sets for that region to more carefully evaluate performance of ArcGIS, UPlan and *What if?*. The following review describes their respective strengths and weaknesses.

4.2.1 Suitability Analysis

Suitability analysis can be used in two ways: either as a complement to land use-transportation models or as a type of land use modeling tool for smaller and mid-size metropolitan areas. Integration of land use-transportation models with suitability analyses is useful for evaluating the impacts of transportation improvements and establishing land use policies that actively promote the "best" patterns of development (according to SA scores). SA techniques are particularly well suited to issues of environmental impact and protection.

For the purpose of allocation, a rule-based method is used that allocates projected growth starting with the highest/lowest score for attraction and continues to work down until it has either allocated all needed space or has run out of available space.



I-Industrial, C-Commercial, R-Residential

Figure 4.2: The SA Allocation Process

Berke et al. (2006) proposed allocating new land uses using the following order of importance: open space, industrial, commercial, and finally residential. Open space comes first because it recognizes (and removes) sensitive lands as well as hazardous areas.

4.2.2 Use of ArcGIS

The most popular GIS software, in widespread use today, is ESRI's ArcGIS. However, it is not particularly designed to perform suitability analysis and thus requires more advanced levels of software knowledge to apply. For this reason, this study provides a step-by-step guide to SA methods for existing ArcGIS users that should be useful for staff at smaller and mid-size MPOs (as well as cities, counties and other jurisdictions).

Part (a) presents the use of SA for an entire study area before development, and Part (b) presents the use of SA as a land use model to allocate new development in an existing region.

4.2.3 Part (a): Suitability Analysis as a Complement to Integrated Modeling

The team was able to obtain a composite suitability map for the selected study area (Waco, Texas). We used ArcGIS version 9.2 for this analysis. ArcGIS commands such as buffer, overlay, intersection, and query were used for the analysis. To obtain the ratings and weightings, we used Anjomani's method as described earlier and the Delphi technique. Finally, a composite map was derived through the application of a rule-based combination method. However, such derivations require that the planner select the order of land allocation. As a result of changing the order, many alternative allocations can be produced. Further research on the implications of such analyses is warranted.

The various steps involved in this application process are:

- Selecting land uses & suitability factors.
- Creating a grid/fishnet.
- Rating selected factors for all the land uses (e.g., -10 to +10).
- Assigning factor weights (e.g., through Delphi Technique).
- Multiplying both weights & rates for all the factors & sum them up for each land use.
- Obtaining a suitability map for each land use.
- Deriving the composite map (e.g., through rule-based method using query command).

The process of land suitability analysis in ArcMap using ArcGIS involves the following (also see the Appendix D2 section on ArcGIS):

1. The selected land uses are open space, single-family residential, multi-family residential, low-density commercial, high-density commercial, and industrial. To create a grid/fishnet for the study area, a 150x150 meter grid was selected.
2. The boundary for the study area has to be defined next. In the case of Waco, the metropolitan planning organization (MPO) boundary was used for the analysis. For further GIS analysis the data was converted into the required form (e.g., converting raster data sets into shapefiles for vector analysis).
3. Suitability factors to be considered for the analysis need to be determined. As our simplified example, we have selected these factors in two categories of proximities or accessibilities and environmental factors. Accessibility factors considered are distances to airports, expressways, and urban centers. And the environmental factors are soil, slope, and land cover.
4. Each of the selected suitability factors is rated for all the land uses based on a scale of -10 to +10, with +10 for highly suitable conditions for a particular land use and -10 for highly unfavorable conditions for a particular land use.
5. Factor weights for each suitability factor were specified using the Delphi Technique.
    a) Each factor is weighted in terms of its relative importance for suitability for the use under study.

6. *Buffer* and *intersection* commands are used at this stage to assign the factor ratings and weightings for the study area.
7. Distances to airports were buffered for distances of 1, 3, 5, 7, 10, and 11 miles.
    a) *Select by location* command was used for each buffer to attribute ratings for land uses.
8. Distance buffers for highways were placed at 50, 200, 500, 1000, 2000, and 10,000 meters.
9. Distances to urban centers were buffered for the distances of 2000, 4000, 6000, 8000, 10,000, and 50,000 meters.
10. Vegetation/land cover was buffered using categories including lake, no trees, trees, and urban.
11. Slope was buffered using slopes greater than 0%, 3%, 6%, 9%, 12%, 15%, and 16%.
12. The weighted rating table that was created in database format (dbf) was attached to the buffer shapefile. These buffer shapefiles were then intersected with the fishnet to identify the cells within these buffers.
13. For the soil shapefile, a buffer was not carried out, so the weighted rating table was attached to the original shapefile and then joined to the grid shapefile.
14. Once the weights are obtained, each of the factor ratings are multiplied with these weights to obtain reliable final figures for each factor and all the land uses.
15. Suitability for each land use is the sum of the adjusted ratings for all the factors.
16. The outcome was transformed into a suitability map with the help of a rule-based method. The land uses and their corresponding percentage values given are as follows: Open space- 27%, Single-family- 22%, Multi-family- 20%, Low commercial- 15%, High commercial- 7%, and Industrial- 9%.
17. Query command *Select by Attributes* was used to derive a composite suitability map for the study area (rule-based combination method).
18. The final output is a composite map symbolizing land uses in a color scheme.

Strengths:

- A rule-based method helps in allocating projected growth starting with the highest/lowest score for attraction and continue to work down until it has either allocated all needed space or run out of available space.
- Select by Attributes query command was used in deriving the composite map. This command helps in allocating the land uses for the grid cells in the order of priority.

Limitations:

- ArcGIS is quite capable of performing suitability analysis, but because it is not designed specifically for this analysis, it requires more time and effort by the user.
- ArcGIS is capable and helps in producing a composite map but it doesn’t have a feature that derives a composite map.

- The calculation of ratings and weightings is a cumbersome process because it requires Delphi technique, which involves several rounds of discussions with the expert groups.

4.2.4 Part (b): SA as for Allocation of Projected Population and Employment

In the case of integrated transportation-land use modeling, SA's identification of optimal locations of future employment centers can be used in the land use modeling process. As part of this modeling process, demographic and economic changes need to be projected. "Population and economic indicators are fundamental to the demand side input to land use planning" (Berke et al. 2006, p. 117). As an LUM, SA techniques need demographic and economic projections similar to other LUMs.

Many mathematical LUMs, such as DRAM®/EMPAL® (Putnam 1991, 1992) and TRANUS (de la Barra 1989, 1998), are predictive in nature, using historical data, behavioral and economic theory to make forecasts of future land use patterns. To use suitability analysis through GIS application as a new approach for this modeling purpose will require that the projection and allocation process of demographic and economic changes be brought into the process.

The projections will be allocated based on accessibility or proximity and environmental variables in SA. The accessibility or proximity variables are those that can be measured by distance or time such as distance to major highways, proximities to major urban centers, distance to airports, and railways etc. The environmental variables are considered as ecological features such as the soil strength, slope, flood plain, environmentally sensitive areas, land cover etc. Spatial allocations will be converted to land uses based on density assumptions.

There are different methods that can be used to allocate the growth to the study area zones. One of the popular techniques especially with professional planners as already discussed is the rule-based method which helps in allocating the projected growth. It starts with allocation based on highest/lowest composite score (for accessibility and environmental variables) and continue to work down until it has either allocated all needed space or run out of available space.

The major steps are:

- Data collection and classification,
- Deriving data sets, and
- Final allocation.

4.2.5 Data Collection

The inputs for data collection had been observations of different mathematical and suitability models like UPlan and DRAM®/EMPAL®. Berke et al. (2006, p. 119) also have identified the three important dimensions of population and economy relevant for land use planning. They are the "size, composition, and spatial distribution." The required data for this analysis includes (1) total number of households and total number of employees for the base year (2000) and the projected year (5-year period), (2) the number of persons and employees per household, (3) the average parcel size for each household income, and (4) the average area per worker.

For the purpose of this study, the base year and the projected year data had been compared for McLennan County from the U.S. Census Bureau and the Texas State Data Center and Office of the State Demographer. The Texas State Data Center followed a popular technique known as cohort-component projection to develop population projections (2006 methodology for Texas Population Projections). In this technique, separate cohorts- persons with one or more common characteristic are used. After selecting different cohorts, by way of applying migration, mortality, and fertility rates for each separate set of cohort, the analysis is performed (Klosterman,1990).

The employment data, such as the number of persons working by industry, were collected from the Texas State Data Center for the base year. For the projected data in this example, we used the project report of the Master Development Plan: TTC-35 High-Priority Trans-Texas Corridor.

4.2.6 Data Classification

Household data

With the help of the collected data described above, both households (HHs) and employment had been divided into different classes for further analysis. The total number of HHs was classified into 16 income categories, starting from less than $10,000 to $200,000 or more for the base year 2000 (Texas State Data Center and Office of the State Demographer). The total number of HHs for the year 2000 is 78,926 and for the year 2005 is 90,951. The difference between the total number of HHs in 2000 and 2005 is the estimated total number of households required for the projected year.

By way of observing other related mathematical models, income levels, and past trends, different categories for residential land uses were derived. For purposes of this study the 2000 year household numbers were combined to have 3 basic categories: high income under Residential Low (RL) land use, medium income under Residential Medium (RM) land use, and low income under Residential High (RH) land use.

Planners rely on various methods to derive the future projections for each category (Berke et al. 2006), depending on the data and time availability. Due to the lack of required data for the year 2005, for the purpose of this study the 2000 year data was used as the basis. As per the 2000 year data, depending on the percentage share of households among the three categories (RL, RM, and RH in 2000), the projections for the year 2005 were estimated.

Table 4.1: Total Number of Households and Percentage Share for Each Category Year 2000

Category	Total number of Households	% share
Low Income (RH)	5,840	7.4
Medium Income (RM)	64,877	82.2
High Income (RL)	8,208	10.4
Total	78,926	100

For the year 2005 the total number of households in each residential density class was developed by assuming that the three shares in 2000 also existed in 2005, as follows:

7.4 % of 12,025 = 890 HH.
82.2 % of 12,025 = 9,885 HH.
10.4 % of 12,025 = 1,250 HH.

Table 4.2: Total Projected Number of Households and the Percentage Share Used for the Projected Year 2005

Category	Total projected number of HH	% share
Low Income (RH)	890	7.4
Medium Income (RM)	9,885	82.2
High Income (RL)	1,250	10.4
Total	12,025	100

Employment Data

Employment types were divided into 13 industry categories, which were later collapsed into 3 categories for further analysis.

Table 4.3: Percent of Employed Persons by Industry of Employment for the State of Texas and Counties in Texas, 2000

Sector	Category	Total #	%
1	Agriculture, forestry, fishing, hunting, and mining	1,005	1.1
2	Construction	6,269	6.6
3	Manufacturing	14,034	14.9
4	Wholesale trade	2,973	3.2
5	Retail trade	11,694	12.4
6	Transportation, and warehousing, and utilities	3,864	4.1
7	Information	2,176	2.4
8	Finance, insurance, real estate, and rental and leasing	6,293	6.6
9	Professional, scientific, management, administrative, and waste management services	6,167	6.6
10	Educational, health, and social services	22,631	24.1
11	Arts, entertainment, recreation, accommodation, and food services	7,136	7.5
12	other services (except public administration)	5,770	6.2
13	Public administration	4,064	4.3

The employment data were then classified into three key categories, as industrial (IND), which constitutes manufacturing, transportation, and warehousing; high commercial (HC), which constitutes high-density office and high-density retail; and low commercial (LC), which constitutes low-density office and low-density retail. The basic idea was to translate the different employment categories into different land uses. So, we selected categories that would match well with land use types.

Because, the total number of projected employees for the year 2005 (13,564 workers) is known, the year 2000 percentage shares by category were used to determine year 2005 values. For example, as per 2000 data, 33% percent of employees are in industrial employment.

33% of 13,564 = 4,508 employees
22% of 13,564 = 2,951 employees
45% of 13,564 = 6,105 employees

Table 4.4: Projected Growth of Employees by Industry of Employment

Type	Projected Growth	%
Industrial (IND)	4,508	33
High Commercial (HC)	2,951	22
Low Commercial (LC)	6,105	45
Total	13,564	100

Further, the values of average area per worker, percentage of workers, and the floor area ratio (FAR), which acts as a zoning control for each employment class, are required. FAR is calculated by dividing the total square footage of a building by the square footage of its parcel lot. The average area per worker, the FAR, and the percent of employees are, respectively, 0.0115 acres, 0.23, and 33% for IND; 0.0046 acres, 0.35, and 22% for HC; and 0.0069 acres, 0.15, and 45% for LC. Berke et al. (2006) explained in detail the entire process of projecting, distributing, and allocating future employment (pp. 361–381).

4.2.7 Deriving Data Sets

This includes the derivation of proximity or accessibility and environmental parameters for both population and employment data. The different variables suggested by Berke et al. (2006) for employment land uses are accessibility to a nearby expressway and proximity to a major railroad or harbor, and environmental parameters that include such things as terrain—e.g., outside floodplain, etc. The proximity or accessibility variables for employment centers considered for this study are distances to airports, distances to major urban centers, access to highways for commercial land uses, and access to a nearby railroad for industrial land uses. Only slope is being considered for the employment-related land uses as an environmental variable. Iacono et al. presumed that households would like to locate where there is higher accessibility to employment or shopping (p. 324). The variables considered for this study are access to highways, distances to urban centers, and slope.

4.2.8 Final Allocation

To allocate the resulting employment and household numbers per selected types, we used the rule-based method, which was discussed earlier in this part. It will allocate projected growth starting with the highest/lowest score for attraction and continue to work down until it has either allocated all needed space or run out of available space.

For this prototype example, the open spaces were allocated first, and then those grid cells were removed for further allocation procedures. Next came the commercial, industrial, and finally the residential for the allocation of employment and households. As an example of application, the steps taken to allocate high commercial (HC) are discussed. The allocation process of the other remaining employment and residential land uses—low commercial (LC), industrial (IND), residential high (RH), residential medium (RM), and residential low (RL) which follow similar steps are discussed in the appendix section.

Data sets considered:

a) Access to Highways—less than 600 meters
b) Distances to Major Airports—greater than 5 miles
c) Slope—less than 3%
d) Distances to Urban Centers—50 to 500 meters.

- The total number of grids selected after performing the query command is 38.
- Anticipated total growth is 2,951 employees.
- Average sq. footage is 200, or 0.0046 acres.
- The size of a single grid is 5.5 acres (150 meters by 150 meters).
- One grid can hold 1,195 employees (5.5/0.0046 = 1,195).
- The total number of grids required for allocating 2,951 employees is 2,951/1,195 = 3 cells.

The steps followed in this allocation process are as described below (and in the appendix to this report):

1) Add the required variables into ArcMap. In the case of high commercial (HC) land use, access to Highways, Distances to Major Airports, Distances to urban centers (Buffer), and Slope are added.
2) *Select by Attributes* query command is used to select the variables within a set of given values. For the variable access to Highways buffer, *Select by Attributes* query command with a value less than or equal to 600 meters was used.
3) *Select by Attributes* query command is used for the remaining variables- distances to major airports with a value greater than 5 miles, distances to urban centers with value between 50 to 500 meters, and slope with a value less than 3%.
4) *Intersect* command is used to intersect these selected attributes in order to identify the needed cells for this particular high commercial (HC) land use.
5) The weighted rating values of high commercial are considered in order to identify the cells needed for allocation.

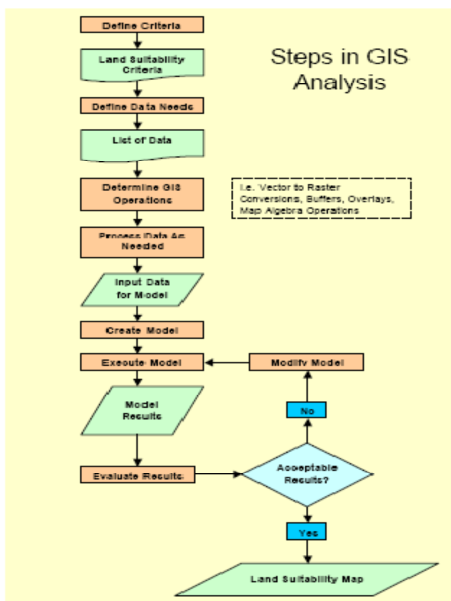
4.2.9 ArcView Spatial Analyst:[5]

ArcGIS's ArcView Spatial Analyst provides a broad range of powerful spatial modeling and analysis features previously not available to desktop users. ArcView Spatial Analyst allows to create, query, map, and analyze cell-based raster data and to perform integrated vector-raster analysis. ArcView Spatial Analyst includes ESRI's ModelBuilder technology for building and sharing spatial models.

In the context of land suitability analysis, spatial analyst helps the user determine what locations are most/least suitable for development. ModelBuilder lets users save models and rerun them using different input data. ModelBuilder helps in changing the datasets for each and every run of the model and generates the map with corresponding changes to the model.

We have successfully completed the suitability analysis application and were able to obtain a composite suitability map using ArcView Spatial Analyst- ModelBuilder for our selected area: McLennan County, Texas. We used version 9.2 for this analysis. This analysis application is intended for a more detailed review of the software and its applicability for suitability analysis with real land use data. However, this effort is merely an exercise rather than a thorough application with a complete and meaningful data. As such this prototype provides a glance into real application, although with fewer land uses and factors than a real application. It is hoped that a real application will be part of the future extension of this research.

Figure 4.3 shows the steps that the user needs to follow in order to perform land suitability analysis using ModelBuilder.



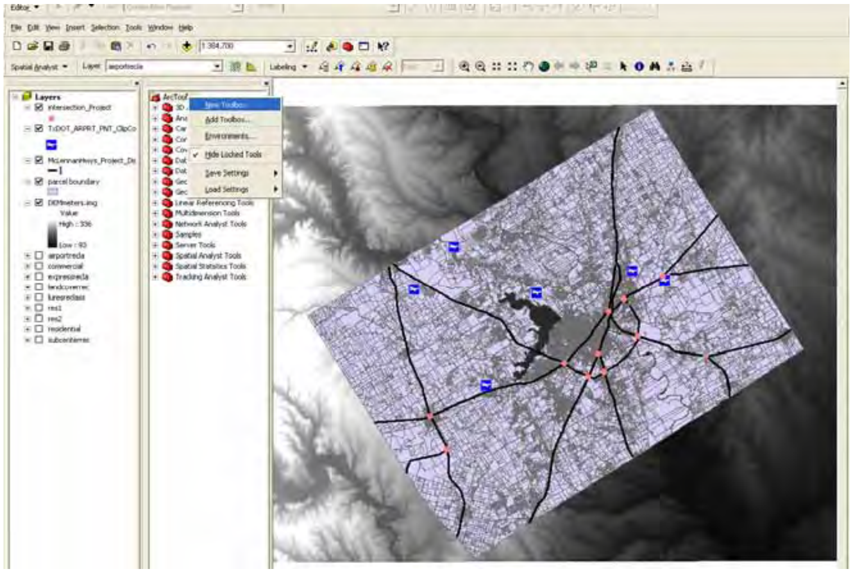
Source: Land Suitability Analysis User Guide. NC Center for Geographic Information and Analysis. December 2005. http://dcm2.enr.state.nc.us/Planning/user_guide_lsa2005.pdf.

Figure 4.3: Steps in GIS Analysis

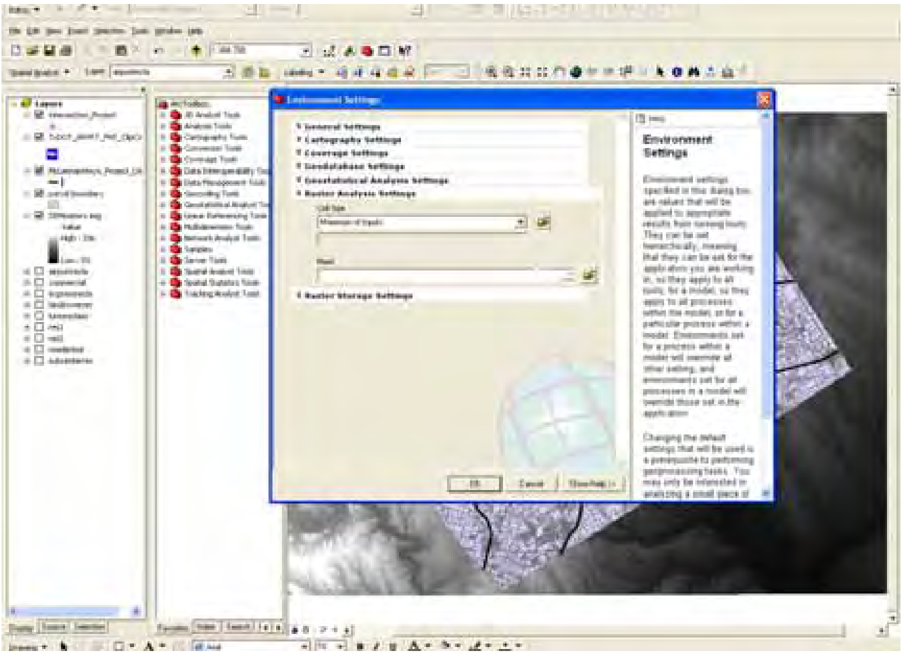
[5] The discussion in this section draws heavily on similar discussion in ESRI ArcView Spatial Analyst and Land Suitability Analysis User Guide (December 2005, pp.1-72).

The process of land suitability analysis in ModelBuilder involves the following:

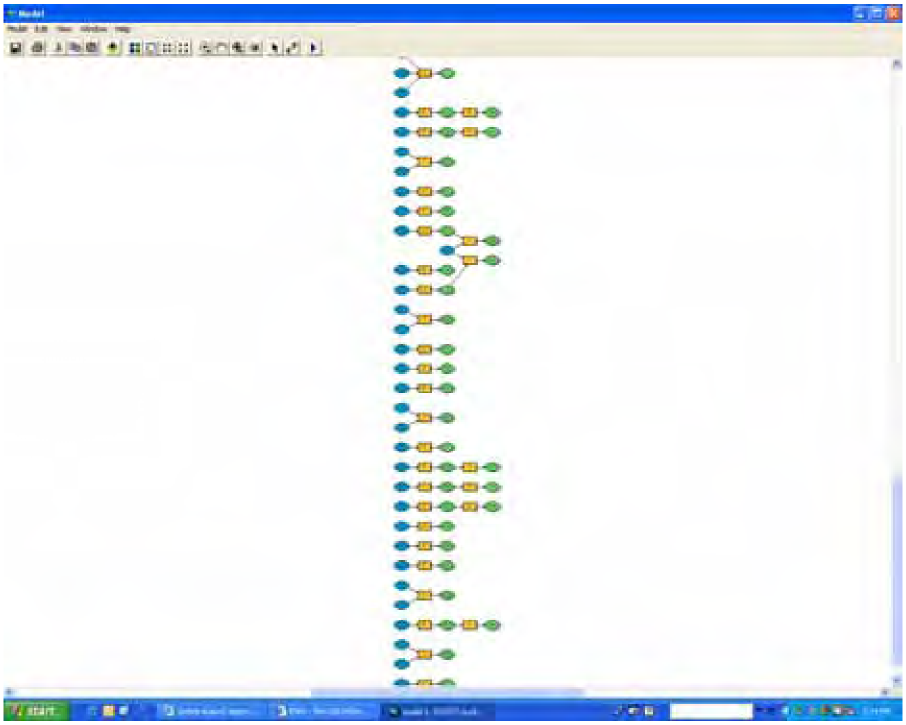
1) Selecting the land uses for the analysis. In our example, we selected four land uses: residential, commercial, industrial, and open space.
2) Selecting the factors for the analysis. Here, we use distances to airports, distances to urban centers, distances to expressways, land cover, and slope.
3) After the input data is ready, as can be seen in the following figure, a new tool is created by clicking on the Arc Toolbox.



4) In the new Toolbox we can create new model and then the 'general settings' and 'raster settings' are specified in order to run the model.



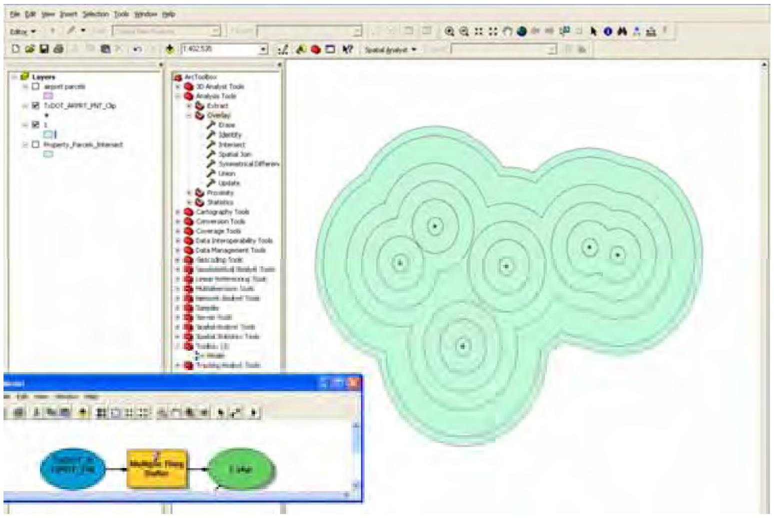
The following picture shows a part of the model used in the analysis.



Arc Toolbox is the common toolbox that also includes all the tools need to be used in the ModelBuilder. The different tools used in the model are buffer, convert to raster, reclassification, weighted overlay, and combine.

5) For all the selected distance factors (distance to airports, distance to urban centers, and distance to expressways), we need to create appropriate distance buffers using the tool ‘multiple ring buffer’.
6) Distances to airports are buffered for distances of 1, 3, 5, 10, and 15 miles. (For figures of the other distance buffers, see the Appendix.)

The created new model window can be seen in the following picture.



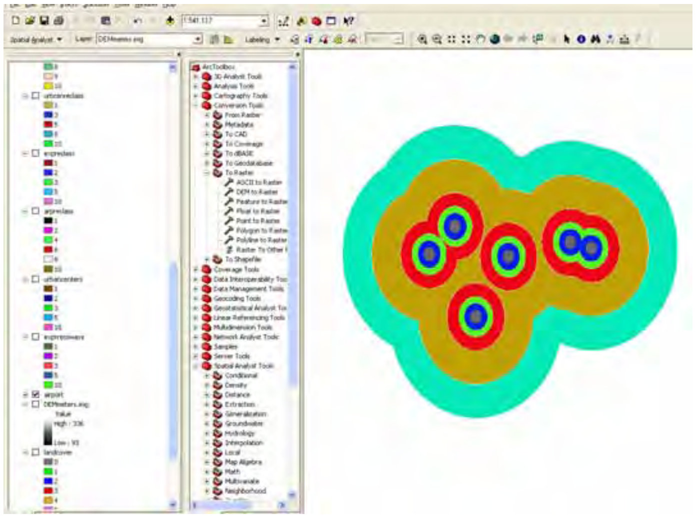
7) Distances to urban centers are buffered for distances 1, 2, 3, 5, and 10 miles.

8) Distances to expressways are buffered for distances 1, 2, 3, 5, and 10 miles.

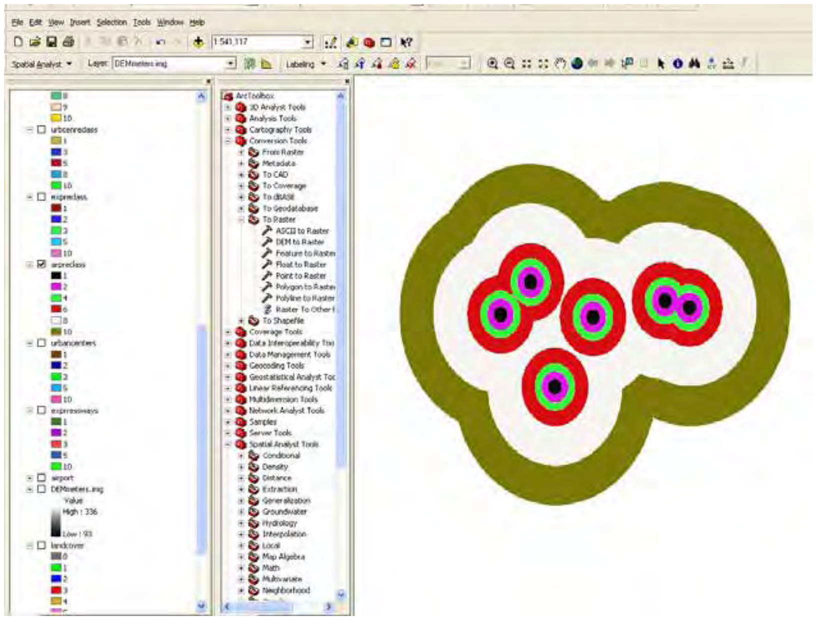
9) Land cover was buffered using the categories such as lake, no trees, trees, and urban.

10) Each of the suitability factors are rated based on a scale of 1 to 10, 1 being least suitable & 10 being most suitable[2].

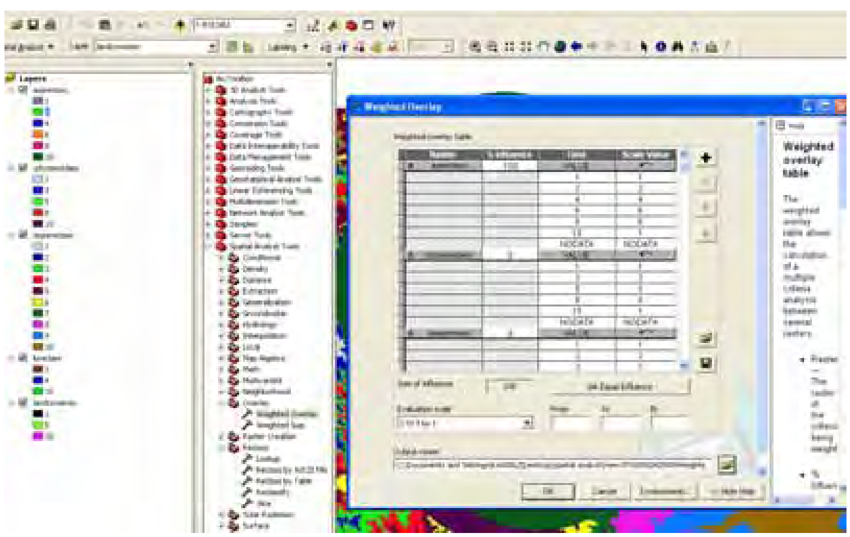
11) All the shapefiles are then converted to raster in order to perform further analysis. To do this all we need to use the tool ‘convert to raster’.



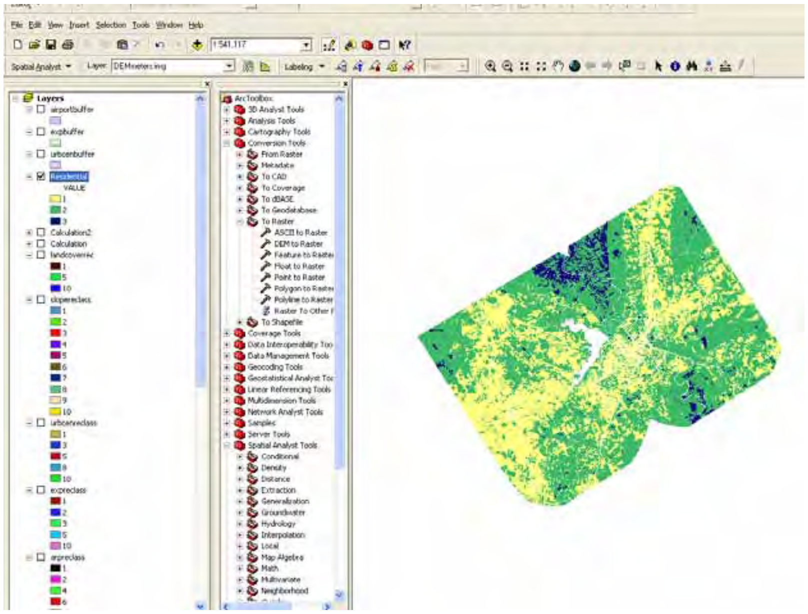
12) ‘Reclassify’ tool in spatial analyst toolbox is used to distribute or divide the range of values into an equal number of intervals. Examples of interval classes are equal interval, defined interval, manual, etc.



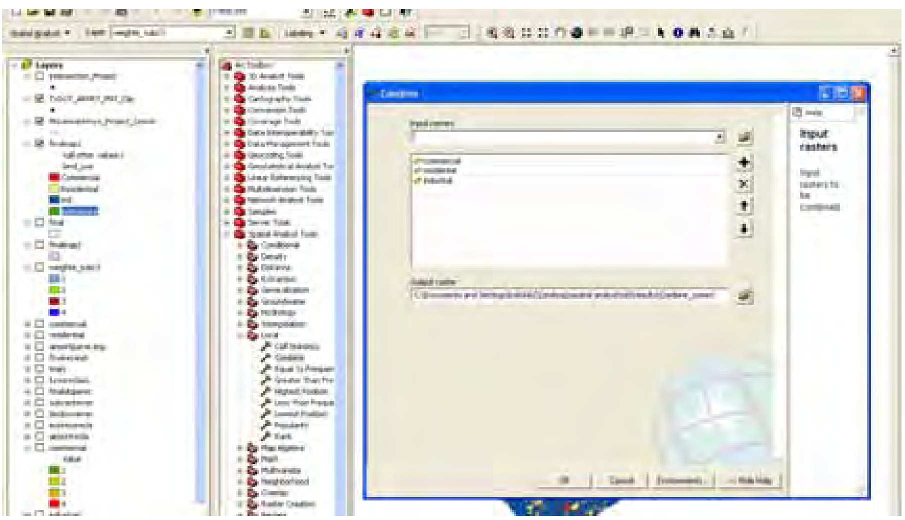
13) ‘Weighted overlay’ tool in spatial analyst toolbox is used to overlay the raster datasets based on their importance.



14) For example, for residential, the given ‘% influence’ values are distances to urban centers 10%, distances to airports 5%, distances to expressways 20%, land cover 15%, land use 25%, and slope 20%.

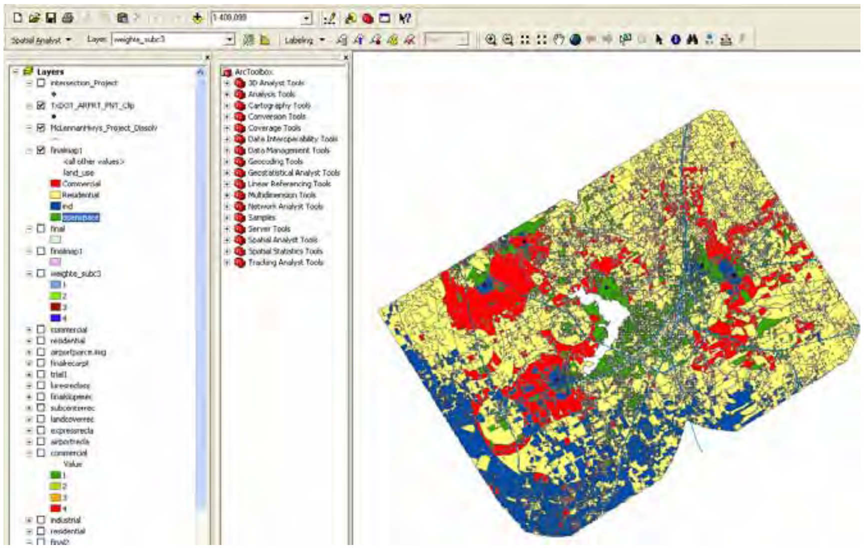


15) After developing the suitability maps for every land use, the ‘combine’ tool in spatial analyst toolbox is used to combine all the rasters (residential, commercial, industrial, and open space) to develop a composite map for the selected area.



16) ‘Combine’ tool “Combines multiple rasters so a unique output value is assigned to each unique combination of input values” (ArcGIS 9.2 Desktop Help).

17) Therefore, this final map is showing the result of suitability application through use of ModelBuilder using selected factor and land uses. This concludes the prototype application of ModelBuilder for suitability analysis.



Strengths:

- ModelBuilder allows the user to change the datasets and the corresponding values (layer weights) for each and every run of the model.
- 'Combine' tool helps in developing a composite map. The tool combines the rasters and gives an output.

Limitations:

- It is time consuming while working with raster datasets. Therefore, it takes time in changing the datasets and rerunning the model.
- Spatial analyst has a limitation in allocating negative values for the ratings.
- Due to larger size of the data, there is difficulty in converting some layers to raster format.

Evaluation

ArcView Spatial Analyst helps in performing the suitability analysis using raster data. The analysis with raster data occurs by combining the layers to create new layers with new cell values. Therefore, raster data is used in suitability analysis because analysis can be performed on several raster layers at once (e.g., weighted overlay function) and in a more efficient manner.

4.2.10 *What if?*

What if? is an interactive GIS-based planning support system (PSS) (http://www.what_if-pss.com/w_intro.htm). It uses existing GIS data to conduct land suitability analysis, project future land use demand, and allocate the projected demand to suitable locations. *What if?* specifically utilizes UAZs (Uniform Analysis Zones), which are GIS-generated polygons for the analysis. UAZs are supposed to be homogeneous in all respects considered in the model. These are created by using GIS overlay functions to combine all of the relevant layers of information to define the UAZs to be used in a study area (Klosterman, 2007).

The software manual in explanation of the software's features states that,

> "*What if?* projects future land use patterns by balancing the supply of, and demand for, land suitable for different uses at different locations. *What if?* provides four options which allow the users to customize the program to the kinds of data that are available for the study area. These options include suitability analysis option, land use analysis option, land use/population analysis option, and land use/population/employment analysis option."

The suitability analysis option allows the user to produce suitability maps showing the relative suitability of different areas for accommodating future land uses. This option can be used only with GIS information based on the current land uses and the suitability analysis layers of the study area.

We successfully ran the *What if?* software with the selected study area. However, it allowed the use of only two suitability factors from a set of required factors for the analysis. Furthermore, producing the existing land use map from the imported data was a cumbersome process. Perhaps the most important problem we encountered was that the software required the smaller polygons be removed before actually importing the data to *What if?* in order to reduce the size of the data. Because of this, a lot of data was eliminated during this process. To reduce the number of cells to be eliminated, many trial attempts were made to reduce the size of the zones to retain more of the data and to use them in the analysis. Because the Waco data is relatively large for this software, the analysis process slowed after the data was loaded. The review of the software is not yet complete. The team communicated several times with Dr. Klosterman, to try and overcome the problems encountered.

We used *What if?* 2.0 for the analysis. Because of the complexity of the application process we think it is worthwhile to briefly explain the steps taken to derive the results. The process of land suitability analysis we took in *What if?* involved the following (see also Appendix D2) (Klosterman, 2007):

1) Convert the GIS data into UAZs (Uniform Analysis Zones) by using the 'union' in ArcToolbox. UAZs are GIS-generated polygons that are homogeneous in all respects. These are considered as the smallest spatial unit of analysis for *What if?*
2) To avoid inconsistencies in the data, convert the coordinate system of the output shapefile (i.e., from geographic to projected coordinate system).
3) *What if?* has a limitation in computing the data; so, it requires the elimination of smaller polygons. In view of this, the following commands in ArcGIS are used:
   a) 'Compute Area' command in ArcToolbox helps in computing the area for all the polygons.
   b) 'Add Field' command in the attribute table helps in adding a new field to the attribute table.
   c) 'Calculate Values' command in the attribute table helps in converting the area of the polygons into a single measurable unit (hectares or acres).
   d) 'Select by Attributes' command from the main menu helps in specifying a size limit for eliminating the smaller polygons (e.g., 1 acre or 2 acres).

e) 'Delete Features' command in ArcToolbox helps in eliminating the selected polygons.

4) The final UNION shapefile should consist of all the important layers such as zoning, shape, soils, UAZ-area, and flood zone. At this stage, the UNION file is ready to be imported to *What if?* software for further analysis.
5) There are two options in *What if?* software. The first is *What if?* Setup Program, which creates the UAZ and other GIS and system files for further analysis. This Setup Program defines the land use categories and the names to be displayed in the selected scenario option. This option creates the project and saves the information. The second option in *What if?* is the Project Option, which is used to modify the land use categories. This option computes the suitability ratings and suitability scores. This is an option for changing the labels and correcting the spelling if any errors exist. Apart from this, the *What if?* Project option generates the required maps for the selected area.
6) After the UNION file is ready, it is first imported to the *What if?* Setup Program.
7) *What if?* Setup Program defines the Analysis options such as Suitability/ Land use, Land use/Population, and/or Land use/Population/Employment.
8) In the Setup Program, the land uses are defined. For example, in our analysis, existing and suitability land uses are to be defined.
9) In the Setup Program, the suitability factors are defined depending on the criteria selected for the suitability analysis. For example, in our analysis, slope, soil, flood zone, distance to airports, distance to expressways, and distances to urban centers are selected.
10) The display layers such as highways, water bodies, and political boundaries are defined.
11) Finally, the project is saved in order to be opened in the *What if?* Project option.
12) The *What if?* Project Option displays all the information already mentioned in the *What if?* Setup Program. The following options are available:
    a) Define the Suitability Scenario Option
    b) Specify Importance Weights
    c) Specify Suitability ratings
    d) Specify Land Use Conversions
    e) Compute Suitability Scores
    f) Create new Suitability Scenarios
    g) View Suitability Analysis Results
    h) Compare Suitability Maps
    i) View Suitability Report
    j) Evaluate Suitability Scenarios
13) These options assist the user in scanning through the input data already given in the Setup program.
14) The 'conversion' option in this Setup program is used to specify land uses that are available for conversion from their current land use to another land use. For example an area that is currently residential may be converted to commercial or other type of use.
15) After inputting the required data, the software computes the data and creates two maps for the study area—the current map and the suitability map.

16) The current map represents the existing land use map and the suitability map represents the map with the land uses and suitability layers.
17) Reports can also be generated that help in identifying the number of acres within each suitability class for all land uses for a specified suitability scenario.

Strengths:

- *What if?* has the advantage of making it possible to change the input data, i.e., the numerical values of the suitability layers, ratings, and suitability scores. For example, new layers can be added or deleted, the values of suitability ratings and suitability scores can be changed, and also new maps with the computed changes can be created. However, the data from the basic ArcGIS file cannot be changed after importing it to the software. Also, after importing the data from ArcGIS to *What if?* software, no new data such as suitability layers and display layers can be added or deleted.
- It provides an integrated software package that incorporates user-provided GIS and other data as a foundation.

Limitations:

- In the process of combining all the GIS layers into a single coverage using the UNION command, there is a possibility of creating more polygons than the software can handle due to the large number of smaller polygons created. These smaller polygons delay the analysis process in *What if?* Therefore, there is a need to remove some of the smallest polygons in ArcMap before importing the data to *What if?* software.
- Because the land uses are predefined, there was a difficulty in categorizing the land uses in ArcGIS before importing the GIS files to *What if?*
- There is a limitation in defining the number of land uses as they are predefined already. Basically, there are two options in the ‘define land uses’ section. They are the ‘existing’ and ‘suitability’ options. Each of these can accommodate a maximum of 50 different land uses. And in each of these different land uses, up to 10 land use categories are permitted.
- For example, in ‘existing’ land use, if ‘Residential’ is a predefined land use, further residential can be subdivided into 10 categories based on the analysis such as low density residential, medium density residential, high density residential etc.
- There is a limitation in the suitability layers too. It can accommodate a maximum of 20 suitability layers. And each suitability layer can be organized in up to five categories only. For example, slopes can be divided into five categories: 0 to <10%, 10% to <20%, and so on. There is a necessity to take care of the values given in the reclassification process in ArcGIS before importing the files to *What if?* Software.
- *What if?* can accommodate up to a maximum of 10 display layers only. These layers might be minor roads, major roads, highways, water bodies, political boundary etc. But these layers must be ‘line layers’.

- *What if?* software requires the data input to be in a more refined format, such as the land uses, suitability layers, and suitability factors.
- Finally, most processes are not discussed clearly in the User Manual.

4.2.11 UPlan

The team successfully implemented UPlan 2.6 for Waco's McLennan County. To complete the process, seven land uses (high density commercial, high-density residential, low density commercial, medium residential and low density residential) and seven factors (highways, airports, land cover, flood plain, urban area, water bodies and existing land use) were used to create buffers. The final output creates an allocation map and a mask map for each land use group, an allocation map for each land use type and an allocation map with all seven land uses types. The final allocation map with all land use types acts as a composite suitability map that depicts the distribution of all seven land uses within defined boundary.

The software manual in explanation of the software's features states that UPlan 2.6 consists of three models. UPlan creates a model that projects urban growth by using several land uses as inputs. This requires at least three densities, one industrial and two densities of commercial land use. The model is not calibrated on historical data as it is intended for long range scenario testing. The model relies on fine grained data to represent existing urban, general land use plans. This is a deterministic (knowable outcome) and rule-based model (as described in UPlan's User's Manual).

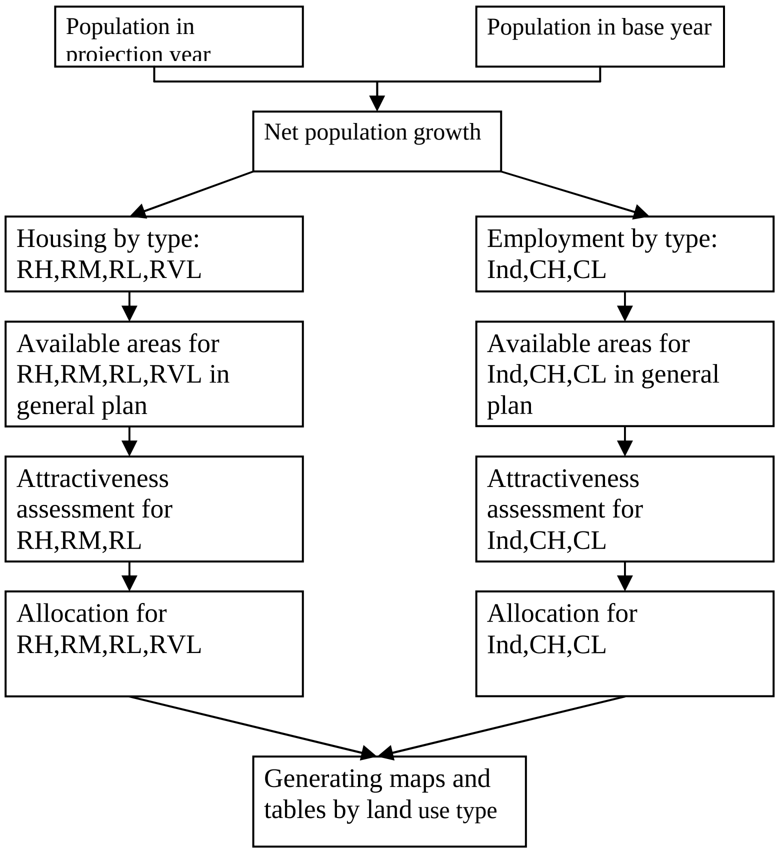
UPlan's assumptions, as per the User Manual can be summarized as follows: "The population growth can be converted into demand for land use by applying conversion factors to employment and households. Cells have different attraction weights because of accessibility to transportation and infrastructure. Some cells, such as lakes and streams, will not be developed. Other cells, such as sensitive habitats and floodplains, will discourage new development."

UPlan 2.6 consists of three sub-models. The first sub-model is a cluster model and designed to test the impacts of improvement of regional transportation infrastructure and land use policies. The second is a county model, designed to project the spatial allocation of residential and employment growth. The third is also a county sub-model. It is a share shift model and is designed to project the spatial allocation of residential and employment at the county sub-area level.

Summary of UPlan's Model Structure

1. UPlan uses any year as its base year and allocates the increment of additional land consumed in future areas. The increment of urban growth is in discrete land use categories: industrial, high-density commercial, low density residential.
2. Waco, McLennan County has been selected to review the UPlan model.
3. UPlan has a module that allows entering demographic and land use density factors which are converted to acres of land consumed for each land use.
4. To determine the acres needed for future housing, data for persons per household, percentage of households in each density class and average parcel size for each land use class is used.

5. To determine acres of land consumed for industry and commercial, data for workers per household, percent of workers in each employment class and average land per worker has been used.
6. The data allows the model to generate a land demand table, used to allocate land for each land use type.
7. According to UPlan, proximity to existing urban areas and transportation facilities has a better chance of development in the future. Therefore, each development attraction has been surrounded by buffers. These buffers vary in size and are assigned weightings based on the distance. Buffer specifications are applied to each of the attraction grids and then the grids are overlaid and added together to make a composite Attraction Grid. For analysis, highways and urban areas in Waco have been considered as attractions.
8. UPlan assumes that there are areas within the city where development cannot occur, such as lakes and rivers, public open space, existing built out urban cells, 100 year flood plains and farm lands. These are called "exclusions." Like the attraction grid, the mask grid is generated from the sum of individual exclusion grids. For the Waco study area, lakes, rivers, and urban areas have been selected as masking layers.
9. Some features such as flood plains and farm lands are costly to develop so these are discouraged from development and these are called discouragements. For the Waco study area, flood plain and airports has been considered as discouragement layers.
10. Similar to attractions, the discouragements are also buffered and weighted by land use groups.
11. To allocate future growth, a General Plan land use map for the region is loaded into the data loader.
12. This General Plan is overlaid with the Attraction grid and the Mask Grid. This enables the model to allocate the projected acres of land consumed in the future.
13. The model allocates future development starting with the highest-valued cells. By default, the model starts with industry, then proceeds to high density commercial, high density residential, low density commercial, medium density residential and low density residential (Johnston et. al). This essentially is a variation of the rule-based combination technique discussed in the ArcGIS section. Figure 4.4 illustrates the UPlan model structure.



Source: UPlan Land Use Allocation Model 2.6 User's Manual

Figure 4.4: UPlan Model Structure

Figure 4.5 explains the household and employment allocation process involved in UPlan. The model distributes the land for residential and employment using the number of persons per household and number of employees per household. Using the basic input data, UPlan tries to generate the new residential and employment areas.

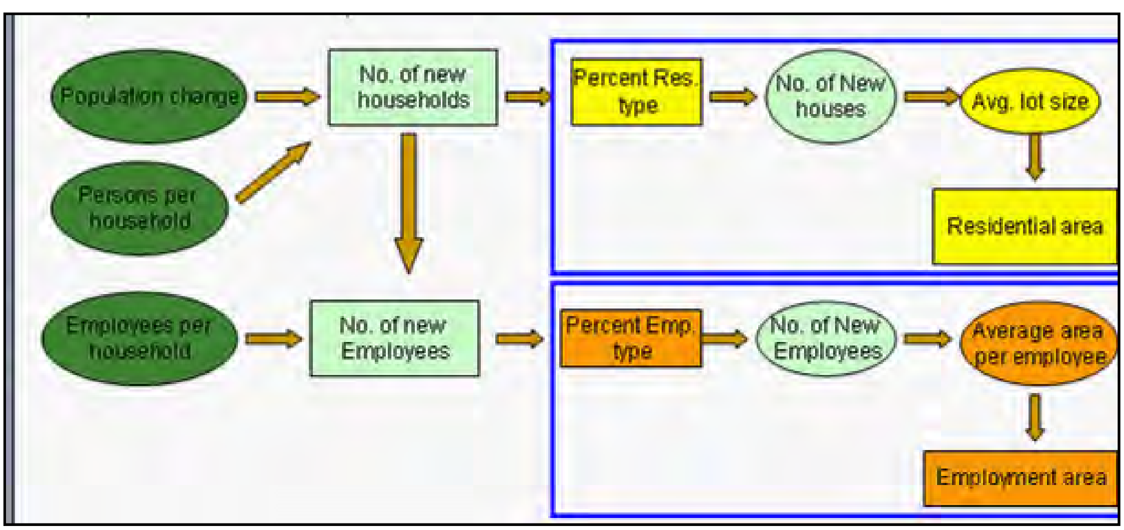


Figure 4.5: UPlan Household and Employment Allocation

Data Considerations

1) UPlan runs on raster (grid) data sets. Default grid size is 50m.
2) Grid data sets have only two fields, cell values and cell count of each of the values, so it becomes important to choose the field values that represent the study area.
3) All the vector shapefiles have been converted to raster data.
4) All files have been projected to the same coordinate system, while converting from vector to raster data.
5) All the converted raster files have been allocated the same grid size that matches with the UPlan grid size, i.e., 50m.
6) After the conversion, the data is reclassified for further analysis.
7) For analyzing the land use demand, UPlan requires area base population, future population, persons per household, and employees per household.
8) To calculate the demand for land, UPlan requires residential density percentage, like RH, RM, RL, and RVL. Data regarding average lot size of each of the above residential densities is also required. To calculate land demand for other land uses such as industrial and commercial, UPlan looks at employment percentage, square footage per employee and floor-to-area ratio. See Appendix D2 for additional details.

In UPlan, the final allocation map is based on rule-based method. Figure 4.6 explains the rule-based method followed in UPlan to allocate land uses.

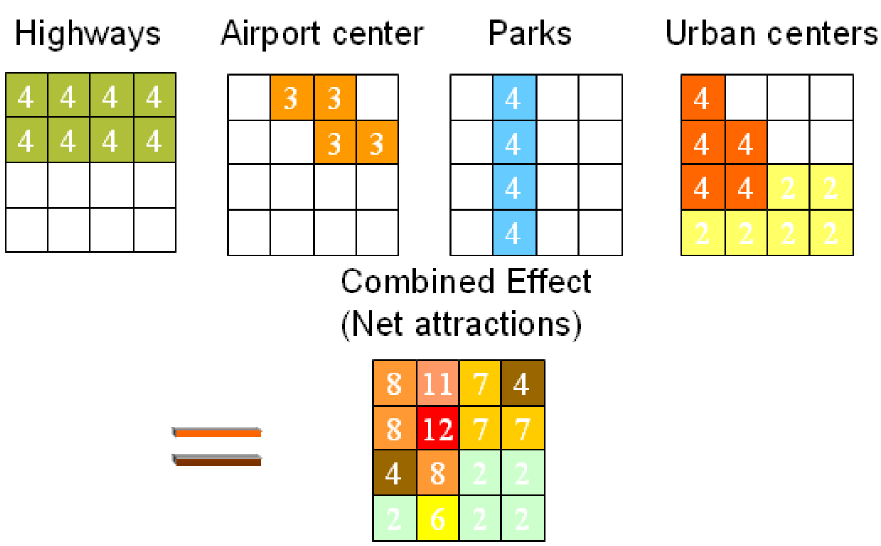


Figure 4.6: UPlan Rule Based Method

There are various factors rated based on their proximity to the land use under consideration. All the rates from the factors are added up to get the Net Attraction table. This net attraction combined with general plan and mask map together helps to derive the final composite map using rule-based method (Figure 4.7).

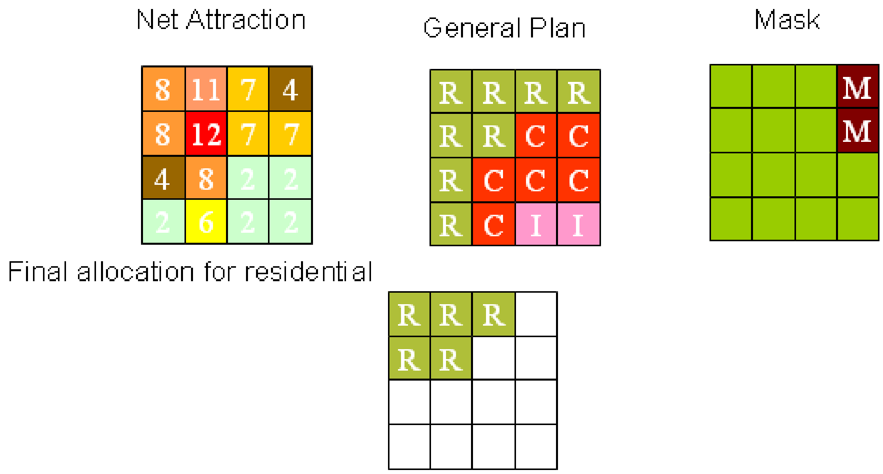


Figure 4.7: UPlan Final Allocation for Residential

These are then allocated considering the General plan, Mask map, and the net attraction from the previous image (see Figure 4.8).

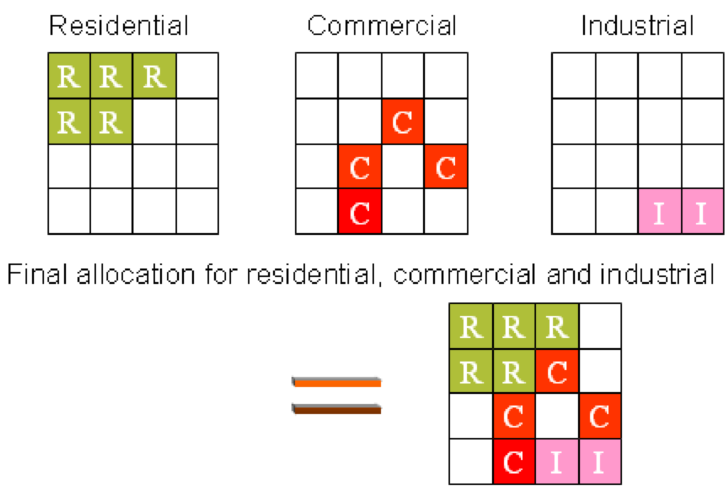


Figure 4.8: UPlan Final Allocation for Residential, Commercial, and Industrial

Finally, carrying the same process for all the land uses under consideration and using rule-based method; UPlan derives the final Composite map. This process was not discussed in the User Manual, but was discussed in various presentations by Johnston et.al which was available on their website.

UPlan provides an assortment of prepackaged reports to be used for later analysis, as well as GIS maps displaying the changes in land use. Additionally, UPlan's flexibility allows the user to define as many time periods as needed. UPlan also has the ability to incorporate more application tools. Once a scenario is loaded into UPlan, the user has the ability to take advantage of ArcView tools such as Spatial Analyst, Geo Statistical Analyst & 3D Analyst. UPlan requires fairly extensive input parameters and data. Its flexibility and customization advantages could instead become liabilities for planning departments that are limited in labor and financial resources. UPlan gives the user more options, yet requires more training and resources to manage. Perhaps UPlan could be appropriate for larger municipalities and government organizations with more resources. UPlan can help decision-makers protect important open space by allowing them to consider the consequences of land-use planning on natural habitats and agricultural lands.

Strengths:

- Easy to use: allows users to prepare and evaluate alternative suitability with specific prompts generated by the program.
- It allows the software to be customized to different geographic areas and conditions.
- Incorporates user-provided GIS and other data as a foundation and applies various tools to produce land use projection based on underlying data.
- The six default land-use types (industrial, commercial [high density and low density], and three residential densities) permit the evaluation of the impacts of future land uses and factors such as flooding and wildfires.
- Uses real time data (population, employment) to project future land demand.
- If proper data is available, the model is easy to run.
- It looks like it might be a good starter model for small MPOs and Counties.
- UPlan is interactive; a user can change the population growth rate, or other basic assumptions.

Limitations:

- Does not provide sophisticated modeling capability or theoretical basis to examine the interrelated factors.
- Requires extensive data reclassification and expertise in GIS.
- The model talks about land demand table but does not show it on screen while working with the software.
- The model encounters bugging problems, but does not provide a proper answer for the error, which is somewhat frustrating for the user.
- The user manual does not discuss the process for involving weights.
- The information regarding the GIS process is missing.
- Requires more training & resources to run than simpler models such as *What if?*

4.3 SA Summary

Suitability analysis is a useful technique that can be integrated into the land use-transportation modeling process. It helps identify and thereby preserve environmentally sensitive areas, while suggesting the locations and patterns of land use that permit allocation of projected employment and population to a region's most appropriate areas.

The reviewed models in general could show the best allocation of land uses if the study area is fully developed. As such, they can be complementary to the land use modeling part in the integrated models that intend to project and distribute the growth of population and employment in the study area. For smaller metropolitan areas, suitability analysis could be sufficient for use in the land use planning stage instead of other modeling techniques. In larger metropolitan areas, the technique could complement the land use models in the integrated transportation planning process. Depending on the size and complexity of the region being analyzed, the five models reviewed appear promising, depending on the context, need, and resource availability.

ArcGIS is an all encompassing and powerful GIS software. As such it is quite capable of performing suitability analysis; however, it is not designed specifically for such analysis and thus requires more time and effort by users. The Guidebook/Resource Manual presents the details of using suitability analysis and ArcGIS as a land use model which includes the allocation of projected changes in employment and population to a study area for a given time period. ArcGIS's Spatial Analyst tools rather specialized, and somewhat inflexible when converting data (e.g., it does not take negative values). Also, the derivation of the final composite map is defined by the user rather than by the data and theoretically problematic. *What if?* apparently has been developed for use by smaller communities, with most information (like factor weightings) being user specified. UPlan requires more training and resources than simpler models. Because it runs in ArcGIS, users still needs to have GIS knowledge. Finally, Anjomani's method relies on optimization and regression, making the application process more complex and somewhat tedious.

Though not without certain limitations, the five suitability models reviewed here may well be helpful in the LUM process. As mentioned above, all these models may assist in identifying environmentally sensitive areas for preservation and/or special treatments. Provisions then could be made to avoid allocating new development to these sensitive areas (other than open space and parkland). Of course, suitability models may also be somewhat competitive with predictive land use models. They may, to some extent, forecast market choices of future land uses, particularly if modeled with an emphasis on profit- and/or demand-maximizing land development behaviors.

SA models may assist in identifying environmentally sensitive areas for preservation or special treatments (to avoid allocating future activities on these sensitive areas) and overall land use planning through conventional planning practices & different scenarios. SA models may be helpful in the land use-transportation (LUT) modeling process, and in achieving TxDOT's objectives. To serve more as a LUM, SA techniques will need demographic and economic projections (e.g., population & employment) similar to other LUMs. Such projections need to be allocated spatially in the study area through use of different processes.

Chapter 5. Integration of Land Use Models with Travel Demand Modeling (TDM) Procedures

The basic procedures for integrating a land use model (LUM) with a travel demand model (TDM)—such as UrbanSim (Opus and UrbanSim User's Guide, 2007) or a gravity-based LUM system with a typical four-step TDM, are specified here, for a seamless integration of model systems.

5.1 Integrating Land Use and Travel Demand Models

Figure 5.1 illustrates the basic method for integrating a LUM and a TDM. The required base year data (discussed in detail in Chapter 3) is specific to the chosen LUM and typically includes demographic data (e.g., population and employment counts), levels of developable and undevelopable land, and accessibility measures for each forecast analysis zone (such as a traffic analysis zone or district). The LUM outputs population and jobs counts for a future year. Many LUMs simulate every year from the base year to the forecast year, and update the accessibility measures somewhat less frequently, when running the TDM. An example of this procedure is depicted in Figure 5.2, where the time between TDM runs is assumed to be just 3 years. One key argument for running the TDM less frequently than the LUM is that land uses (households and jobs) take several years to respond to changing travel conditions, whereas travel behaviors may change rather instantaneously, following location changes. A more practical reason is simply saving run time. Moreover, in practice a region's land use modelers often are unfamiliar with the TDM—and a region's travel demand modelers are unfamiliar with land use modeling. In any case, this iterative process, between the LUM and the TDM, continues until the forecast year is reached.

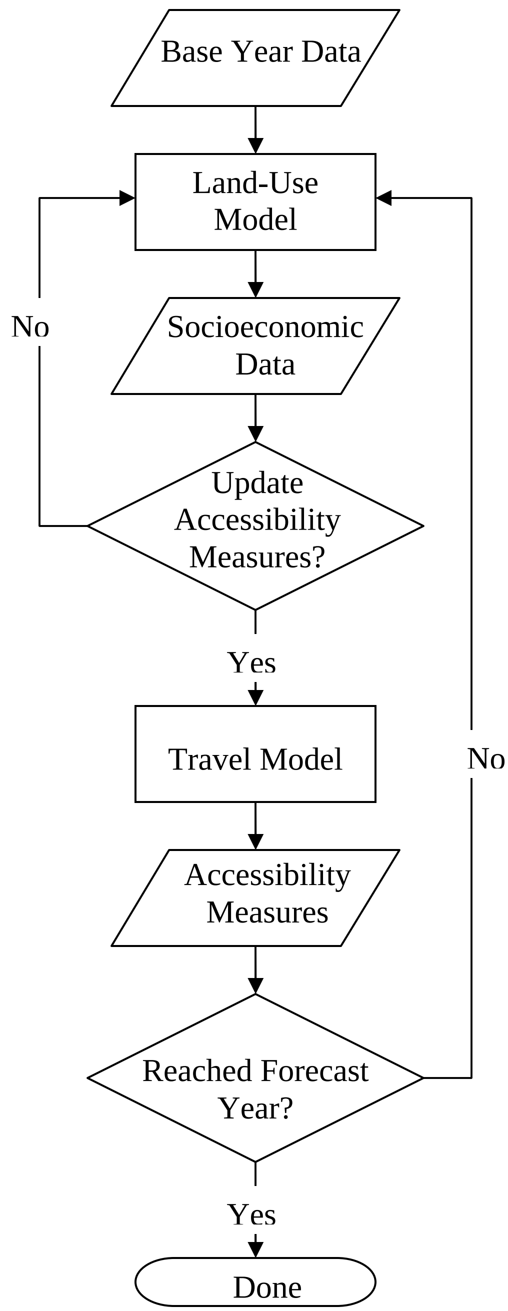


Figure 5.1: Integration of Land-Use and Travel Demand Models

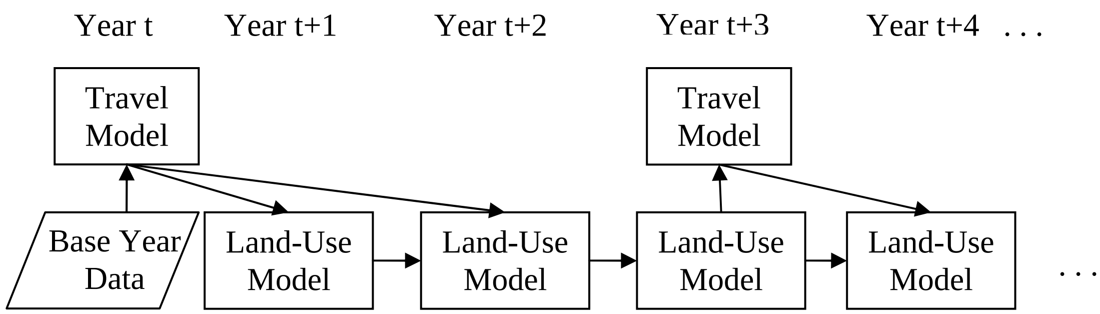


Figure 5.2: Example of Dynamic Nature of Integration

Figure 5.3 illustrates the sequential process of a traditional, four-step TDM. First, the numbers of trips generated by and attracted to each traffic analysis zone are estimated. Second, a gravity-based or multinomial logit model predicts the number of trips between each zonal pair. Third, a logit model for mode choice provides estimates of probabilities for trips by each mode, between each zone pair (e.g., drive alone versus transit). Fourth, these trips are assigned to the network such that each traveler cannot improve his travel time or cost (i.e., shortest-paths). This four-step procedure is iterated until convergence is achieved. The procedure used by TxDOT is a three-step procedure, omitting mode choice. A detailed description of the procedure used by TxDOT for transportation planning can be found in Chapter 2 of the report.

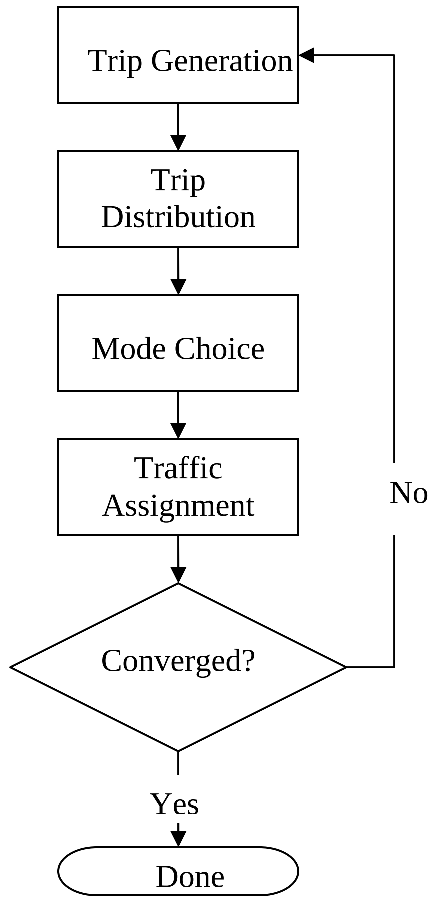


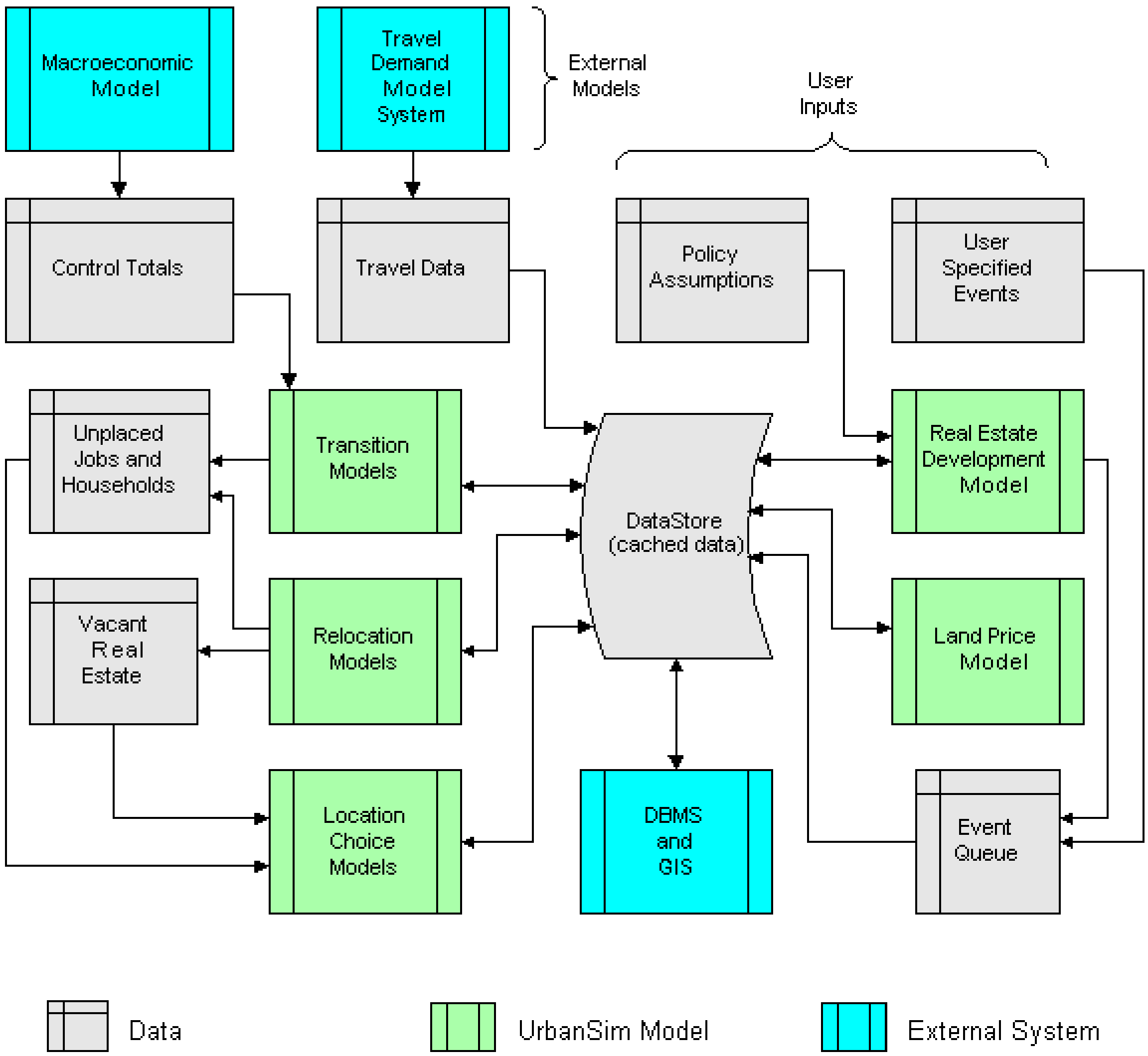
Figure 5.3: Standard Four-Step Travel Demand Model

While most LUMs rely, to some extent, on transportation network information, "the devil is in the details" in terms of actual model linkages. The following section describes the finer points for integrating UrbanSim and standard TDM practices in Texas. Details for the integration of TELUM and Texas TDMs are also provided, in the subsequent section.

5.2 Use of UrbanSim

As described in Chapters 2 and 3 of the report, UrbanSim is a microsimulation model developed at the University of Washington by Dr. Paul Waddell and teammates. The latest version of UrbanSim is designed to simulate land use at most any level of spatial aggregation, including small grid cells (e.g., 150m x 150m) or entire traffic analysis zones. The following paragraphs summarize UrbanSim's framework, the software required for TDM integration, appropriate accessibility measures, and associated module editing needs.

Figure 5.4 illustrates the interaction of UrbanSim's submodels. The TDM outputs are input into the DataStore, and then converted into accessibility measures for use in the location choice models. Regional control totals for households and employment are fed into the transition models, where the numbers of households and jobs to be added (or deleted) are calculated. The location choice models assign all unplaced households and jobs to grid cells. (Unplaced jobs and households consist of those that enter the region to meet a control total, and those that choose to relocate as specified by the relocation models.) The real estate development model predicts the probability that each location will be developed based on policy assumptions, construction costs, vacancy rates, and cost of land. The cost of each piece of land is calculated within the land price model (Opus and UrbanSim User's Guide, 2007).



Source: Opus and UrbanSim User's Guide [Figure 2], 2007

Figure 5.4: UrbanSim Model Structure

The following software tools are required to run UrbanSim: TDM software (e.g., TransCAD or Cube), OPUS (which stands for Open Platform for Urban Simulation) (UrbanSim, 2007), UrbanSim (UrbanSim, 2007), MySQL (MySQL, 2007), and Python (Python, 2007). Any TDM package that produces accessibility measures (based on travel costs/times and attractions/activity sites) is acceptable. Of course, Caliper Corp.'s TransCAD already is in regular use by TxDOT's Transportation Planning and Programming Division. OPUS and UrbanSim were developed by Paul Waddell's team at the University of Washington, and documentation can be found in the Opus and UrbanSim User's Guide (2007). MySQL is a powerful open-source database management system created and owned by MySQL AB and used here for handling the input and output tables. Microsoft Access can be used to view and edit MySQL data. Python is a high-level programming language managed by the non-profit Python Software Foundation. The user will need to be familiar with basic Python commands in order to run (and edit) the model because Python is the code used to create UrbanSim. Instructions for installing the appropriate versions of these software tools can be found at http://www.urbansim.org/docs/installation/.

Thanks to Python and MySQL programming languages, TransCAD can be rather seamlessly integrated with UrbanSim, and this was recently done for the South Eastern Michigan Council of Governments (SEMCOG). The integration of UrbanSim with a TDM can be facilitated through the use of pywin32, a Python extension for MSWindows that allows communication between the UrbanSim and TDM programming languages (e.g., GISDK for TransCAD and Python for UrbanSim).

Several different zonal accessibility measures can be chosen to link one's TDM to UrbanSim. Accessibility measures a zone's proximity to other zones' "activities" (where activities typically are proxied by the number of jobs and households). Accessibility measures can be as complicated as the double-logsum of destination and mode "utilities" (from a nested logit model of destination and mode choices) and as simple as a zone's travel time to the region's central business district. They can be the number of jobs within a given number of travel minutes from a zone's centroid or the sum of time-weighted job counts to all zones (resembling a gravity model's denominator term).

While the OPUS and UrbanSim User's Guide (2007) recommends using nested destination-mode choice logsums stratified by vehicle ownership level, other stratifications are possible (e.g., by household income). Such stratifications allow UrbanSim's household location choice models to be more finely tuned to each household class. Regardless, UrbanSim model development ideally should include testing of several different accessibility measures, to identify one that may work best. If an accessibility measure is chosen that differs from the default (nested destination-mode choice logsums stratified by vehicle ownership level), the user must edit the location choice models to ensure that the measure is calculated appropriately.

To date, UrbanSim has been successfully integrated with EMME/2 for an application by the Puget Sound Regional Council and with TransCAD for Detroit's SEMCOG. Presently, integration of UrbanSim with PTV's VISUM software is underway (Waddell, 2007). While these models have been integrated with a specific region's TDM and other needs in mind, altering such code for use in another region is a far less daunting undertaking than developing the integration code from scratch.

Chapter 6. Use of TELUM

Of course, TELUM (shown in Figure 6.1) is another, simpler LUM option for TxDOT and Texas MPOs. A gravity-based model developed by Dr. Steve Putnam at the University of Pennsylvania, TELUM predicts future zonal locations of jobs and households. TELUM requires a fairly high level of spatial aggregation, in what it terms “forecast analysis zones,” in order to meet minimum count requirements. These zones tend to be quite a bit larger than the traffic analysis zones used in Texas TDMs. The sub-model TELUM-Emp forecasts job locations (by industry sector), TELUM-Res forecasts household locations (by household type), and LANCON estimates land consumption (per job or household) in each zone (TELUM User’s Manual, 2005). Each of these sub-models has been discussed in more detail in previous sections of the report.

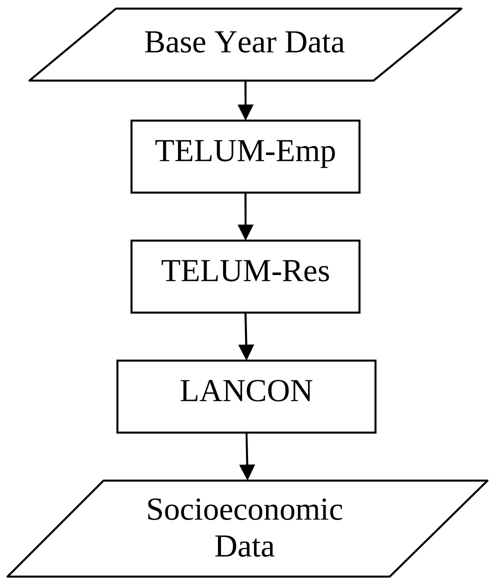


Figure 6.1: TELUM Model Structure

The inputs required for uses of TxDOT’s TDM process are as follows: total, population, households, median income, total employment, basic employment, retail employment, service employment, educational employment, special generator employment, and special generator population—all by zone (Hodges, 2007). These can be generated by TELUM. For example, TELUM-Emp can generate all the employment related inputs into the TDM.

In both TxDOT’s TDM and in TELUM, households (HH) typically are classified by income. The TELUM-Res (sub-model of TELUM) generates location forecasts for each of the HH categories. (A minimum of four such categories is required by TELUM). And median household income (by zone) then can be computed based on TELUM-Res outputs. If the average number of persons in a household is known, then the population of each zone also can be computed. LANCON outputs provide estimates of the land consumed and the developable land left in a zone.

Like any LUM, TELUM can forecast future household and employment locations over multiple time periods without updating zonal accessibility measures. TELUM users are given the option of entering future year travel times, essentially avoiding the use and application of a

TDM. The accuracy of TELUM forecasts, however, should be enhanced by updating accessibility terms and interzonal travel times, based on TDM runs that use TELUM outputs as inputs. Unlike in emerging code for UrbanSim, however, integrating TELUM with a TDM requires manual intervention at every time step. The user must be present to feed the output of TELUM to a TDM and to feed the output of a TDM back to TELUM (Figure 6.2).

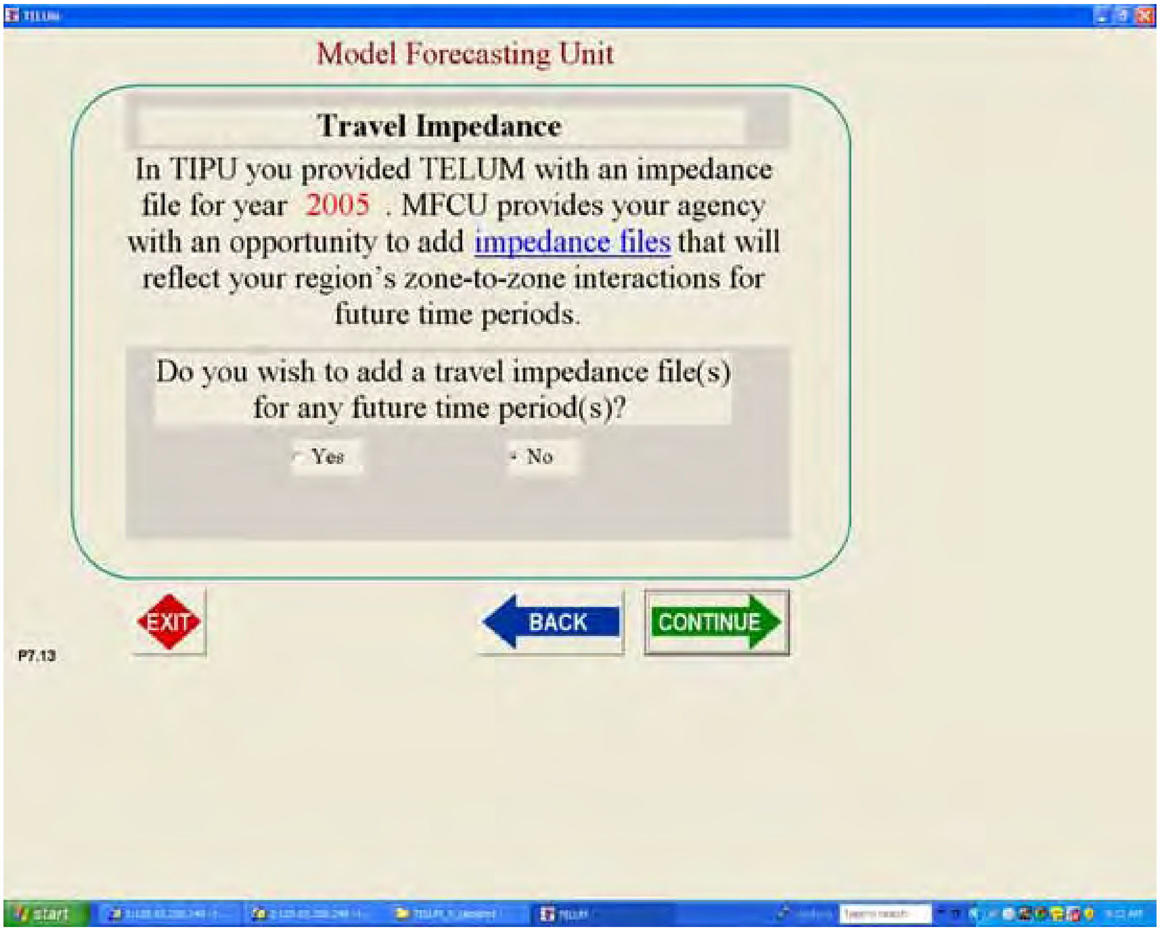


Figure 6.2: Updating the Travel Impedance File in TELUM

The procedure for integration of UrbanSim with a TDM is described in Chapter 5. Full integration is not possible with TELUM, because the code is privileged and not designed to accommodate such seamless transactions. Instead, TELUM requires direct user intervention to iterate it with a TDM. Emerging code for UrbanSim, however, is designed to automatically iterate with TransCAD software (or VISSUM). Both models output the data necessary to run most any TDM, including TxDOT's TDM. While the effort necessary to seamlessly integrate UrbanSim and a TDM is not insignificant, Texas should be able to build off of the work already done by other planning agencies, such as SEMCOG, who have recently developed integrated models.

Chapter 7. Establish Calibration and Validation Procedures for Integrated Models

This chapter begins by introducing the concepts behind land use model (LUM) calibration and validation and by outlining a basic framework for applying these concepts to integrated models. Sections describing specific procedures to be used with UrbanSim (Waddell, 2002) and the Transportation Economic Land use Model (TELUM) (Putnam, 2005) then follow.

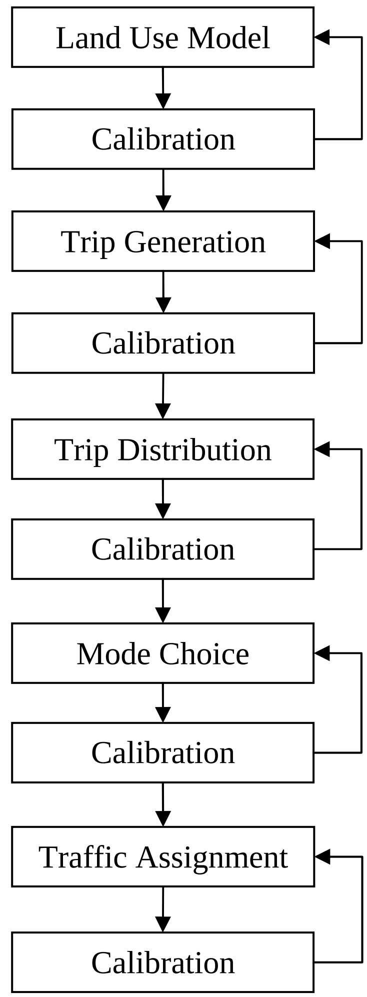
Three terms are fundamental to these discussions: estimation, calibration, and validation. These are sometimes used somewhat casually and interchangeably, but here they convey distinctive meanings, as described below.

Estimation is the process of finding model parameters that best fit the observed data. For example, household travel surveys and transit on-board surveys typically are used to estimate the coefficients in a mode choice model, using a multinomial logit (MNL) specification and maximum likelihood estimation techniques. When surveys are not available and estimation is not possible, such coefficients may be borrowed from other regions. The distinction between model estimation and model calibration is not always so clear, and TELUM actually refers to its estimation process as a calibration process. In practice, most modelers refer to *calibration* as the process of adjusting the estimated model parameters until model outputs match a second set of observed base year travel data, such as link flows, regional vehicle-miles-traveled, and total transit system ridership (Wegmann and Everett, 2004).

The third process, *validation*, examines a model's ability to predict future patterns of observed behavior. It thereby relies on at least two time periods worth of data. For example, land use conditions in 2000 and 2005 are given, so that 2000 can be used as the base year, in order to predict year 2005 conditions—and then compare these to actual year-2005 conditions.

Calibration and validation rely on distinct data sets—essentially past/older data, in the case of original model calibration, versus the most recent data available to the planning agency, for model validation, which is ideally followed by further model calibration. Calibration emphasizes adjustment of model parameters, while validation emphasizes evaluation of a model's predictive capabilities. They can occur in concert (iteratively), as differences between observed and predicted values during the process of validation signal the need for further calibration (Wegmann and Everett, 2004).

The Federal Highway Administration's Travel Model Improvement Program (TMIP) has produced valuable documentation on validation (Barton-Aschman, 1997) that recommends validating model results after each step or stage of the model (e.g., a parcel subdivision model, followed by a land use change model, followed by a model of parcel intensity [jobs and housing, for example]). Comparing results after each model step facilitates the discovery of any discrepancies, along with sub-model calibration. TMIP documentation describes four common measures for quantifying such discrepancies: absolute difference (estimated minus observed values), relative difference (normalized absolute differences [such as average percentage mis-prediction]), correlation (i.e., r-squared between actual and predicted values [such as population across all zones]), and variance or standard deviation in residual errors (i.e., percent root mean squared error [RMSE] between actual and predicted values). Figure 7.1 illustrates potential stages in calibrating an integrated transportation-land use model system. Because each LUM is different, the exact steps are not outlined. As noted previously, calibration should be performed at each step of the modeling process to ensure reasonable results. Care must be taken to ensure that calibration does not unduly limit the model's flexibility. It is possible, for example, that upstream calibrations may lead to worse downstream predictions.



(Adapted from Barton-Aschman, 1997 Fig. 11-12)

Figure 7.1: Calibration Process for a Traditional Model of Travel Demand

The following two sections of this report detail the calibration process for UrbanSim and TELUM LUMs. Readers may consult Barton-Aschman's report (1997) for an in-depth discussion of calibrating a four-step travel model.

7.1 UrbanSim Modeling

This section describes specific features of the estimation and calibration process for the UrbanSim land use model. UrbanSim's version 4.0 is designed to estimate the coefficients of nearly all necessary choice functions and regression models. Calibration (outside the model software) then consists of using longitudinal data to check prediction reasonableness. This section discusses UrbanSim's built-in estimation process, suggested techniques for calibrating several parameters that are not treated by the built-in methods, and a new calibration technique that recognizes model input uncertainty.

7.1.1 Estimation of UrbanSim Parameters

A subset of the base year household and employment data is used for estimation of UrbanSim's over-1,000 parameters. The number of parameters to be estimated depends on such factors as the number of development types and employment sectors, and it will differ by region. For example, the Eugene, Oregon application for the 1980 base year has 1,865 estimated parameters. Equation 7.1 gives land price, and Table 7.1 shows a subset of the estimated parameters for UrbanSim's land price equation (based on a regression model), as applied to Volusia County, Florida in 2005:

$$\mathrm{p}_{ilt} = \alpha + \delta\left(\frac{\mathrm{V}_i^s - \mathrm{V}_{it}^c}{\mathrm{V}_i^s}\right) + \beta \mathrm{X}_{ilt} \tag{7.1}$$

where p_{ilt} is the price of per acre of development type i in location l at time t, V_{it}^c is the current vacancy rate of development type i at time t, V_i^s is the structural vacancy rate of development type i, X_{ilt} is a vector of characteristics relating to development type i in location l at time t, and α , δ, and the vector β are parameters to be estimated.

Table 7.1: Estimated Parameters for UrbanSim's Land Price Equation

(Adapted from Zhao and Chung's (2005) application in Volusia County, Florida)

Coefficient Name	Estimate	Standard Error	t-statistics	p-value
constant	0.45423	0.02746	16.54	0.0000
is_near_arterial	0.72954	0.01904	38.32	0.0000
is_near_highway	-0.80253	0.04120	-19.48	0.0000
ln_total_nonresidential_sqft_within _walking distance	0.31855	0.00341	93.33	0.0000
percent_high_income_households_ within_walking_distance	0.02679	0.00042798	62.60	0.0000
percent_mid_income_households_ within_walking_distance	0.02407	0.00045309	53.12	0.0000

In addition to regression models like Eq. 7.1, UrbanSim also estimates coefficients in MNL models. The estimation routines output meaningful information on R^2 and log-likelihood values, standard errors, and p-values, along with the coefficient estimates (UrbanSim, 2007).

7.1.2 Model Calibration

Calibration is similar to the estimation process, but it uses a different dataset (e.g., actual freeway speeds instead of freeway speeds resulting from the outputs of some volume-delay curves mixed with travel demand modeling [TDM] results). Because LUMs are designed to be dynamic by predicting forward in time steps, calibration typically requires data from two time steps. Multiple time steps of data are not always available, but often rough data related to current conditions can be obtained. If this rough data is available, and the base year differs from the current year, then the base year data can be run until the current year is reached and basic checks could be done on how well the prediction matches current conditions.

The research team has determined that four of UrbanSim's sub-models contain parameters that are not covered in the built-in estimation procedures. These models are shown in Table 7.2, along with suggestions of simple procedures for calibrating the parameters listed in the "Alter" column. To calibrate each entry in the "Alter" column, one can calculate the corresponding variables in the "Calculate" column and alter the appropriate "Alter" variable until the calculated quantity is considered reasonable and/or matches available data. These ideas are discussed in more detail below.

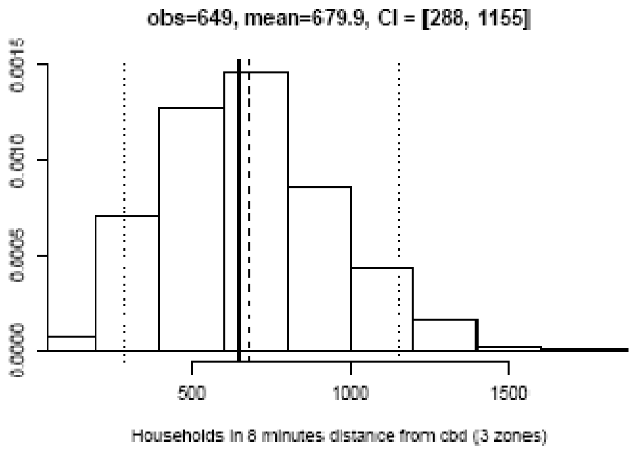
Table 7.2: Calibration of UrbanSim: 4 Key Models

Sub-model	**Calculate**	**Alter**
Household Location Choice	population densities by zone	household relocation probabilities, $P(h,t)$
Employment Location Choice	employment densities by zone	employment relocation probabilities, $P(j,t)$
Real Estate Development	densities of new development	construction costs, H, S demolition costs, D
Land Price	average price per zone	vacancy rates, V α δ β

The household and employment relocation probabilities, $P(h,t)$ and $P(j,t)$, represent the probability that a household of type h or a job of type j will relocate in time period t. These probabilities are initially used in the demographic and economic transition models, respectively, to ensure that control totals are met. Hard construction costs (H), soft construction costs (S), and demolition costs (D) are parameters used in the real estate development sub-model to determine the utility of each zone for each type of development. Current and long-term structural vacancy rates (V) as well as three parameters (α, β, δ) are used within the land price sub-model to determine the price of development type i in location l at time t. Refer to Chapter 5 of the report for more details.

The UrbanSim development team is presently developing an alternative model calibration method, called Bayesian Melding (BM). BM begins with prior probability distributions for the base year (y_1) input parameters based on historical data, and also a subsequent year (y_2) of data for comparison with model outputs (Sevcikova et al., 2007). Monte Carlo simulation produces model outputs in y_2 for numerous realizations of the y_1 input parameters and random number seeds. Weights are assigned to each model run based on the likelihood of the outputs given the actual y_2 data. For each parameter realization and random number seed, the model is then run until a third and future year, y_3, is reached; and the weights are used to form a probability distribution for each output measure. The model is considered to be "calibrated" if the actual data for y_3 consistently falls within a confidence interval (e.g., 90%) of the output probability distribution (Sevcikova et al., 2007).

Figure 7.2 shows a sample histogram of the number of households in eight minutes driving distance from Eugene, Oregon's central business district (CBD) based on this repeated, probabilistic simulation of UrbanSim runs. The solid line is the actual observation, the bold dashed line is the mean of the simulated outcomes, and the light dashed lines denote the lower and upper bounds of the 90% confidence interval.



(from Sevcikova et al., 2007 Fig. 9)

Figure 7.2: Histogram of the Simulated Posterior Distribution for the Number of Households Living within 8 Minutes of the CBD

UrbanSim 4.0's built-in estimation tools provide a great convenience to users who would otherwise have to rely on statistical software and would need to have expert knowledge of the estimation process. UrbanSim is also exemplary among other LUMs in its developer's interest in the issue of calibration. Although the BM process is not yet ready for public use, its development is a step in the right direction for the LUM community: Treatment of uncertainty in model inputs is critical to determining the range of feasible futures that a region may experience. A coherent expression of model uncertainty can better inform planning and policy making, and lead to decisions that work well under a host of potential future outcomes.

7.2 TELUM

TELUM consists of three sub-models, namely TELUM-EMP, TELUM-RES, and LANCON. TELUM-EMP forecasts the future locations of employment, while TELUM-RES generates the residential forecasts. LANCON estimates land consumption levels by zone, largely based on current employment and population levels. As noted earlier, TELUM (somewhat inappropriately) refers to parameter estimation as model calibration. The "CALIBTEL" module estimates the value of each parameter in the TELUM equations. The EMP submodel has five parameters per employment type; RES has six parameters per household type; and LANCON has 19 parameters. For more information on the TELUM model see Chapter 2 of this report.

Our research team has developed "open source" MATLAB code in order to better evaluate the calibration results of TELUM (which remains something of a black box, because TELUM documentation is inadequate and TELUM technicians are unable to answer most statistical questions). As a dynamic LUM, the TELUM parameter estimation process requires two time points of data. But this does not mean that validation is being performed (as it largely would in a cross-section/single-time-point model, like a standard TDM).

7.2.1 TELUM's Estimation

Some parameters in the EMP, RES, and LANCON models depend on the regional data and are calculated using calibration. The calibration process involves fitting the TELUM equation to the data of the region. The better the fit of the model the more reliable are the forecasts. In EMP and RES five and six parameters have to be determined for each employment type and household type respectively, while in the LANCON model 19 parameters have to be determined empirically. Calibration is a critical step in TELUM because the forecast can vary significantly across different combinations of parameter values.

In calibration, the base year forecasts are generated from the lag year data using the TELUM equations. The combination of parameters that generates the best base year forecast is used to forecast the employment and household location in the future. The parametric values of ITLUP equations are typically determined by solving an optimization formulation. The standard multiple regression techniques cannot be used to estimate the parameters of the model as the equations of TELUM are non-linear and the regional data may not be normally distributed.

The goodness of fit measure also has a significant influence in selecting the parametric values. Two commonly used measures of fit are R^2 and likelihood. Putman (1983) found that the R^2 criteria produced a flat surface close to the local optimum, complicating the optimization process. In comparison, he found that the likelihood criterion generated a steeper surface, facilitating model calibration. Thus, TELUM reportedly uses the entropy-based log-likelihood function as its goodness of fit criteria. It uses CALIBTEL, a gradient search technique, to solve the optimization formulation and thereby determine the model's parameter estimates.

7.2.2 MATLAB Code for TELUM's Estimation

As noted earlier, our team independently developed MATLAB code to try and replicate the TELUM and ITLUP software. The code was developed to overcome several of TELUM's restrictions (mostly on zone size, because TELUM greatly limits zonal resolution) and allow user access to the base equations (thus guaranteeing the model's specification). To be consistent with Putman's text descriptions of ITLUP (which TELUM documentation references), we used the following entropy-maximization formulation to determine parameter values. The formulation given in Putman (1983) is very nearly the same.[6]

$$\max \sum_{i \in I} N_i \ln\left(\hat{N}_i\right) \tag{7.2}$$

$$\textit{subject to} \sum_{i \in I} N_i = \sum_{i \in I} \hat{N}_i \tag{7.3}$$

where I is the set of all zones, N_i is the count (of jobs or population) in zone i, and $\hat{N}_i$ is the estimated value (of jobs or households) in zone i (essentially a complex function of other variables in the system, and a suite of unknown parameters). Our code solves the above formulation using MATLAB's built-in "Nelder-Mead" (1965) method, a numerical method for finding the optimal solution to a non-linear multi-dimensional problem.

---

[6] The User Manual neglects to ensure that total predicted counts equals total observed counts (Eq.6.3). This implicit constraint is quite fundamental and must have been simply overlooked by those providing the documentation. One other point to keep in mind is that entropy maximization of this sort relies on Sterling's Approximation (where the ratio of factorials is very nearly the sum of these multiplicative terms). This approximation works best with large values of N_i. Eq. 7.3's expression is equivalent to type of likelihood maximization for shares of jobs or housing across zones.

The effects of parameters that influence the household and residential spatial distribution but are not incorporated into TELUM are captured in "residuals." Residuals are computed as the difference between the actual data of the base year and the predictions generated by TELUM for the base year using the lag year data. TELUM allows users to meet base year targets by retaining all residuals and adding these back in to the estimates (essentially as a suite of fudge factors, one per zone per job and household category). The effect of these is said to diminish near-term mispredictions, but users do not know to what extent these values actually diminish. The TELUM-like application code the team has developed assumes a 20% reduction in these residuals every 5 years. TELUM does not allow users to fine-tune its parameters (evidently, to avoid misuse of the software), so further calibration or code modification simply is not feasible. Finally, validation is left up the user, and simply involves comparing predictions with actual, data, using recent out-of-sample counts (of jobs and housing). But, again, there is no way to fine-tune the TELUM parameters. This makes our MATLAB application code particularly useful. However, it was not able to come very close to TELUM's predictions in the long term Austin simulations we developed. For all these reasons, and several others (including challenges in getting the software to run properly and limitations on zone resolution), the team cannot highly recommend TxDOT's use of TELUM. We hope TxDOT personnel and others will be comfortable applying our MATLAB code, and/or can obtain very nearly the same code from TransCAD, which runs in a TransCAD environment (but is not yet available for public distribution).

Following estimation of all model parameters, integrated land use-transportation planning models generally require some form of calibration if analysts want to come close to meeting regional targets. Such targets are not always easily met, primarily because model systems are abstractions of reality, and model parameters are developed based on small samples and various assumptions. Policymakers and the public tend to scrutinize many key model results and they will lose faith in models that do not meet basic checks of reasonableness. Sequential calibration of each model step as it transpires is the norm, allowing analysts to fine-tune pieces of the model, as discrepancies emerge, rather than having to discern the origins of discrepancies that emerge in the final model outputs.

Both UrbanSim and TELUM allow for model estimation within the software systems. While no formal calibration procedures are yet available for UrbanSim, several suggestions are provided in this chapter of the report: household and job transition probabilities can be altered to achieve better predictions of population and employment densities by zone, construction and demolition costs can be altered to better match densities of new development, and vacancy rates and other parameters in the land price model can be altered to reflect average land prices by zone. The UrbanSim development team is working on a Bayesian Melding procedure that will input estimates of uncertainty surrounding each model parameter and output confidence intervals around each model output. This treatment of uncertainty should greatly aid planners in making decisions that work well for a wide range of feasible futures.

CALIBTEL is TELUM's internal procedure for estimation model parameters. To account for the effects of parameters that influence the household and residential spatial distribution but are not incorporated into the model, TELUM calculates residuals. These residuals are preserved for re-use in prediction as a way of hitting base year prediction targets, and then they slowly diminish over time. However, TELUM remains something of a black box for analysts, and the project team was unable to duplicate its parameter predictions, long-run jobs and housing predictions, or discern the fractions of residuals it chose to lose in each time step. Better documentation and more analytical flexibility are needed for widespread application of TELUM. The team's own MATLAB code and the code forthcoming in Caliper's TransCAD may be much more meaningful for TxDOT and Texas MPOs.

Chapter 8. Demonstration of System Integration, Modeling Approach, and Calibration Procedures

The modeling approach used to estimate all UrbanSim sub-model parameters as well as details of the integration of UrbanSim outputs with a typical travel demand model (TDM) are described in this section. Details on gravity-based land use model specification, calibration and integration was provided above[7].

Future year land use patterns depend on various factors including residential location (and relocation) choices, employment location (and relocation) choices, local and regional accessibilities, land prices, new development and land use policies. These factors are addressed in UrbanSim using calibrated sub-models. Each model has a distinctive specification and needs to be calibrated to best match Austin data. Calibration of land use models (LUMs) typically requires two distinct time points of data (e.g., land use patterns in 2000 and 2005); however, such detailed data is rarely available, so UrbanSim relies on land use data for only a single year. The control totals given as input values for households and employment also impact the model's predicted values.

8.1 UrbanSim's Model Specification

The models that have been estimated and calibrated are as follows:

8.1.1 Residential Location Choice Model

Households that are newly formed (e.g., newly married individuals splitting from their parents, a student moving to a new place for educational purposes) and households that are moving from one location to another within the region need to be placed based on personal attributes (CUSPA, 2006). The probability that these households choose a specific residential location (150m × 150m gridcells) is estimated at a micro-level resolution. This model has a multinomial logit (MNL) structure with explanatory variables like price (divided by household income), access to local employment, age of housing stock in each cell (note: there are 622,050 gridcells in our Austin case study).

From the household location choice specification, it can be observed that housing cost has a negative impact on the utility of household location choice model (HLCM) and the sensitivity of cost with respect to income also has a significant impact. Moreover, the number of residential units in the gridcell has a negative impact on the cell's utility, but the number of residential units in the gridcell within walking distance has a positive impact. Also, it can be observed that people belonging to the same income levels tend to co-locate.

8.1.2 Employment Location Choice Model

Similar to the residential location choice model, the employment location choice model locates newly created jobs and transferred jobs across gridcells. This model is also based on an MNL specification with parameter values varying by employment sector (i.e., industrial, commercial and home-based types of employment).

---

[7] UrbanSim is much more taxing of the two land use models (in terms of data and computing requirements) and thus required much more time and effort on the part of team members.

UrbanSim allocates jobs based on square footage required for the job and the total non-residential square footage available in the gridcell. Supply of space is defined as follows:

$$\boldsymbol{S_{blt} = S_{bl} / R_{bl},\ for\ l \in\ L_t} \tag{8.1}$$

$\boldsymbol{S_{blt}}$ is the number of job spaces available in a building type *b*, at location *l* and time *t*
$\boldsymbol{S_{bl}}$ is the total non-residential square footage for building type *b* at location *l*
$\boldsymbol{R_{bl}}$ is the space utilization in square footage for building type *b* and location *l*.

For each job, Monte Carlo sampling of location choices is undertaken, based on a subset of relative utilities across the region's 150m gridcells. (CUSPA 2006)

Industrial, Commercial, and Retail Employment Location Choice Models

Industrial employment location choice model specification indicates, industrial job locations are influenced by access to freeways and other arterials, the local presence of other service sector jobs (e.g., hospitals), households, and other industries, land values and age of the buildings on site. As one would expect, industrial square footage is viewed positively along with the non-residential improvements on site. Total cell employment has a negative effect (in large part because industrial uses tend to segregate), but total non-residential development (measured in square feet) within walking distance has a positive (though statistically not significant) effect. Overall industrial employment location choice model uses less predictive power than other land uses. The estimates of this mode are shown in Appendix E. Here, commercial employment refers to jobs in retail sector. And commercial employment location choice model specification implies that such jobs are attracted to gridcells enjoying more industrial square footage, higher overall value (land plus improvements) and lower non-residential improvement property values.

From the estimated results in UrbanSim's home-based employment location choice model, home-based jobs (essentially persons working from home) appear to be attracted to newer buildings, locations enjoying higher land values, lower levels of local population and higher average incomes.

8.1.3 Land Price Model

Land price is a key to match demand and supply of land across modeled locations. UrbanSim relies on ordinary least squares regression for estimates of how various attributes impact land values, as follows (CUSPA 2006):

$$P_{ilt} = \alpha + \delta\left(\frac{\mathrm{V}_i^s - \mathrm{V}_{it}^c}{\mathrm{V}_i^s}\right) + \beta \mathrm{X}_{ilt} \tag{8.2}$$

Where P_{ilt} is the price of land per acre of development type *i*, at location *l* at time *t*
δ is set by the user (often based on sensitivity testing, because data do not usually support direct estimation of this parameter)
V_{cit} is the current vacancy rate at time *t*, locally and regionally
X_{ilt} is a vector of location and site attributes
α and β are estimated parameters.

The land price model specification indicates that the variables that increase land price are improvement value, residential value, the year in which the building is built, commercial and industrial square footage in the surroundings, retail and service sector employment, and surrounding developed area. Vicinity to highways, basic sector employment, nearby low-income households, nearby non-residential square footage, residential area, and open space are estimated to decrease land price.

8.1.4 Residential Land Share Model

This model predicts the fraction of residential land in each gridcell (CUSPA, 2006). The fraction of residential land *y* is obtained as follows:

$$y = \frac{e^{\gamma}}{1+e^{\gamma}} \tag{8.3}$$

where *y* is a linear function of zone attributes and coefficients to be estimated. Residential land shares tend to fall with increases in basic and retail sector employment, local access to jobs and increases in the number of local residential units.

The residential land share model specification indicates that the fraction of residential land decreases with the increase in the basic and retail sector employment, and the accessibility to employment location from residential location and decreases with the increase in the residential units in the locality.

8.2 Integration of UrbanSim and Travel Demand Models

Using the sub-models described above (and as shown in Appendix E), and base-year values for all explanatory variables, UrbanSim is able to forecast a region's land use patterns, typically in one-year time steps. These outputs serve as inputs to a travel demand model (TDM), and future-year travel times and costs are obtained. These TDM outputs then serve as inputs to UrbanSim and future land use patterns (typically 5 years forward) are forecasted. This procedure was repeated here, until the land use patterns for Austin in the year 2030 were obtained. The entire process is illustrated by Figure 8.1.

UrbanSim has not been seamlessly integrated with TransCAD. Detroit's SEMCOG and Paul Waddell's team have been working on integrating the two, along with VISSUM, but the TDM is quite specific to SEMCOG's needs, rather than generic in nature.[8] Because we have access to TransCAD, we manually integrated UrbanSim and a four-step TDM, which was calibrated using 1996 Austin Travel Survey data (Lemp 2007).

Four iterations of full feedback (from network assignment to trip distribution) were used in running the TDM. The outputs obtained from the four-step model are matrices corresponding to travel times and travel costs across different times of day and modes.

Year 2030 results of UrbanSim with and without updating year 2000's travel times and costs (i.e., with and without integration of the TDM) are shown in Figure 8.2. Integration of UrbanSim with the TDM implies that the accessibilities terms obtained from the TDM are

[8] Supposedly, OPUS already can interface with Emme/2, MinUTP, and TP+. We are not sure why TransCAD does not.

introduced into the LUM every 5 years. These accessibilities are essentially nested logsums over modes, times of day, and destinations, based on the underlying utility expressions, as follows:

$$L_{ij} = \beta_{1nmt} + \beta_{GC,n}\left(\frac{Cost_{ijmt}}{VOTT} + TT_{ijmt}\right) \tag{8.4}$$

$$L_{ij} = \ln\left(\sum_{m,\in C} \exp\left[U_{nijmt}\right]\right) \tag{8.5}$$

$$A.I._i = \sum_{j=1}^{J} D_j e^{L_{ij}} \tag{8.6}$$

where *'n'* denotes trip type (only home-base-work trips are considered in this integration), *m* denotes modes (drive alone, transit & walk/bike modes here), *i* & *j* denote trip origin & destination, *t* denotes time of day (only AM peak period is considered here), D_j denotes quantity of activity in destination *'j'*, L_{ij} denotes the logsum for a trip which originates in zone *'i'* and is destined for zone *'j'*, and $A.I._i$ denotes the accessibility index for zone *'i'*.

As evident in Figure 8.2(c), households in Austin's more rural and sub-urban zones tend to move towards the CBD and urban zones in the scenario where the TDM has been integrated with the LUM. This is because the network congests over time, and central locations become relatively more attractive (to help households and workers save on travel times). If one does not update the year 2000 travel times, the population and jobs exhibit more dispersed patterns over the long run. This is as expected, which is reassuring. However, the results are not so striking. More comparisons will be helpful.

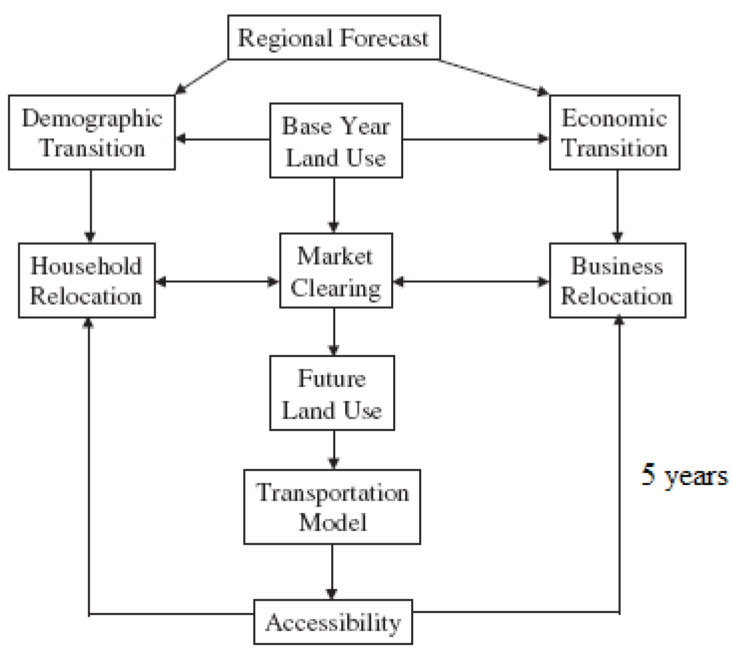
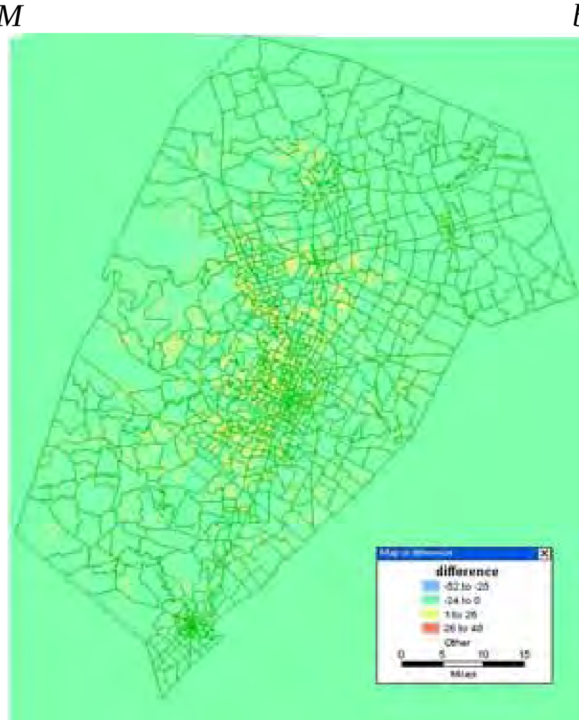


Figure 8.1: Flowchart Showing the Integration of UrbanSim with a TDM (Source: Dowling et al., 2000, Fig 7)

a) Before TDM *b) After TDM*



c) Difference in Estimates: After minus Before Densities

Figure 8.2: Land Use Patterns in 2030, Before and After Integration with a TDM

8.3 Gravity-based Land Use Model (G_LUM) Applications

TELUM and our MATLAB code were both used to generate employment and household forecasts for the Travis, Hays, and Williamson counties of the Austin metropolitan region. These applications are described here. For a description of Waco, Texas applications, refer to the report at http://www.ce.utexas.edu/prof/kockelman/G-LUM_Website/InvestigationOfTELUM.pdf.

TELUM has three submodels: DRAM®, EMPAL® and LANCON. The first two are trademarks of S.H. Putman Associates, Inc. While specification of the G-LUM's three components in MATLAB was designed to mimic TELUM and to follow published materials as closely as possible (e.g. TELUM 2007), actual model equations are no doubt slightly different from the trademarked, proprietary software. Therefore, these coded components are referred to as RESLOC, EMPLOC, and LUDENSITY in the model applications using MATLAB. Because the MATLAB code evolved and incorporated rules in order to ensure reasonable forecasts, the applications demonstrated here are based on the earliest version.

Williamson County is located in the northern part of the Austin, while Hays County is in the south. Travis County lies in between, and contains the region's center: the City of Austin. The three counties in total have 109 districts or 1,074 TAZs. The employees in the three-county region were classified into the following categories:

i) Basic employment (Bas)
ii) Services employment (Serv)
iii) Retail employment (Ret)
iv) Airport employees (Air)
v) Employees in colleges (ED1)
vi) Employees in school (ED2)

Their households were classified as:

i) Low income (Low): Annual income less than $25,000
ii) Below average (BAvg): Annual income between $25,000 to $40,000
iii) Above average (AAvg) Annual income between $40,000 to $75,000
iv) High income (High): Annual income more than $75,000

Due to its recency and data availability, year 2005 was chosen as the base year for implementation. Because TELUM makes predictions in five-year increments, data from year 2000 was used for calibration. The employment data and the total number of households in each zone were obtained from the CAMPO. The households were then divided into the above mentioned income categories based on year 2000 Census data. The land cover data for 2005 was obtained from the Capital Area Council of Governments (CAPCOG). The forecasts were generated by the TELUM and MATLAB codes using travel times from the base year 2005. The travel times are not updated based on the forecast of the previous iteration.

8.3.1 Model Calibration

Refer to Appendix I for an explanation of each of the gravity-based model's parameters. Model calibration results using TELUM are shown in Tables 8.1–8.3, and results from MATLAB code are in Tables 8.4 through 8.6.

Table 8.1: DRAM® Parameter Estimates—calibrated using TELUM across Austin's 109 districts

Parameters	Low-HH	BAvg-HH	AAvg-HH	High-HH
η	0.493	0.27	0.195	0.381
α	1.297	1.348	1.013	0.989
β	-0.186	-0.206	-0.056	-0.03
q	-0.064	-0.044	0.085	-0.013
r	0.43	0.5	0.42	0.435
s	0.359	0.342	0.105	0.244
B (HHtype,Low)	5.773	-0.549	-4.088	-3.36
B (HHtype, BAvg)	4.39	8.633	1.66	0.195
B (HHtype,Aavg)	-1.097	0.459	8.262	3.258
B (HHtype,High)	-3.003	-3.787	-0.365	5.239

Table 8.2: EMPAL® Parameter Estimates—calibrated using TELUM across Austin's 109 districts

Parameters	Basic	Retail	Serv	Air	College	ED1	ED2
λ	0.3132	0.3477	0.3306	0.146	0.3764	0.2308	0.3132
α	2.8096	2.8105	2.8103	2.8106	2.8102	2.809	2.8096
β	-0.009	-0.0086	-0.0108	-0.0062	-0.0178	-0.0093	-0.009
a	0.6172	0.6052	0.6067	0.7651	-0.0044	0.5669	0.6172
b	0.1176	0.1424	0.1157	0.0876	0.2787	0.0436	0.1176

Table 8.3: LANCON Parameter Estimates—calibrated using TELUM across Austin's 109 districts

Parameters	Residential	Industry	Commercial
Constant	0.089726	0.014728	0.026702
PerDev	0.7311	0.095	-0.551
PerBas	-0.297	0.633	0.951
PerComm	-0.565	-0.379	-0.898
PerLI	0.026	0.983	0.516
PerHI	-0.057	0.399	0.05
Developable	0.743	n/a	n/a

Table 8.4: RESLOC Parameter Estimates—calibrated using MATLAB across Austin's 109 districts

Parameters	Low-HH	BAvg-HH	AAvg-HH	High-HH
η	0.0004	0.000215	0.000236	0.000106
α	1.2382	0.89818	0.88763	0.89606
β	-0.0501	-0.04884	-0.04864	-0.04873
q	0.0896	0.011473	0.011458	0.010031
r	0.488	0.66018	0.65734	0.65772
s	0.4214	0.20567	0.20496	0.20391
B (HHtype,Low)	0.6935	-0.50495	-2.003	-3.1126
B (HHtype, BAvg)	0.6598	8.5297	5.019	0.21482
B (HHtype,Aavg)	0.5699	5.6832	8.8308	3.2166
B (HHtype,High)	0.4857	1.0617	-0.48058	9.3381

Table 8.5: EMPLOC Parameter Estimates—calibrated using MATLAB across Austin's 109 districts

Parameters	Basic	Retail	Serv	Air	College	ED1	ED2
λ	0.5177	0.0905	0.1895	0.2734	0.0283	0.0041	0.5177
α	0.2123	2.5664	4.9848	3.6195	3.8403	1.1559	0.2123
β	-0.0005	-0.0012	-0.0016	-0.0011	-0.0007	-0.0005	-0.0005
a	0.8209	0.4449	0.4455	1.0396	0.0046	0.0697	0.8209
b	-0.0701	0.2248	-0.1079	0.1503	-0.033	-0.4992	-0.0701

Table 8.6: LUDENSITY Parameter Estimates—calibrated using MATLAB across Austin's 109 districts

Parameters	Residential	Industry	Commercial
Constant	1.13E-05	1.0914	0.0495
PerDev	0.9725	0.4303	-1.19
PerBas	0.3983	0.77	0.3021
PerComm	0.0648	1.3301	-0.9004
PerLI	-0.3253	-0.2558	0.037
PerHI	-0.3544	0.7768	0.3942
Developable	1.2438	n/a	n/a

The entropy value (which was maximized in model calibration in order to find the estimated parameters) obtained using the MATLAB code is greater than the entropy value of the parameters computed by TELUM in all three sub-model cases. Therefore, parameters computed by the MATLAB code are expected to yield more accurate predictions than TELUM equations for the three-county Austin region.

8.3.2 Model Forecasts

A comparison of the outputs produced by TELUM and the MATLAB code is presented in this section. The spatial distributions of three main employment classes and total employment along with the spatial distributions of all four household classes in the forecast years, as produced by TELUM and MATLAB codes, are discussed here. The forecasts are produced in increments of 5 years. The zones have been classified into four classes based on the number of employment and household types present in that zone (Figures 8.3–8.16). Each class is assigned a different color.

Basic employment

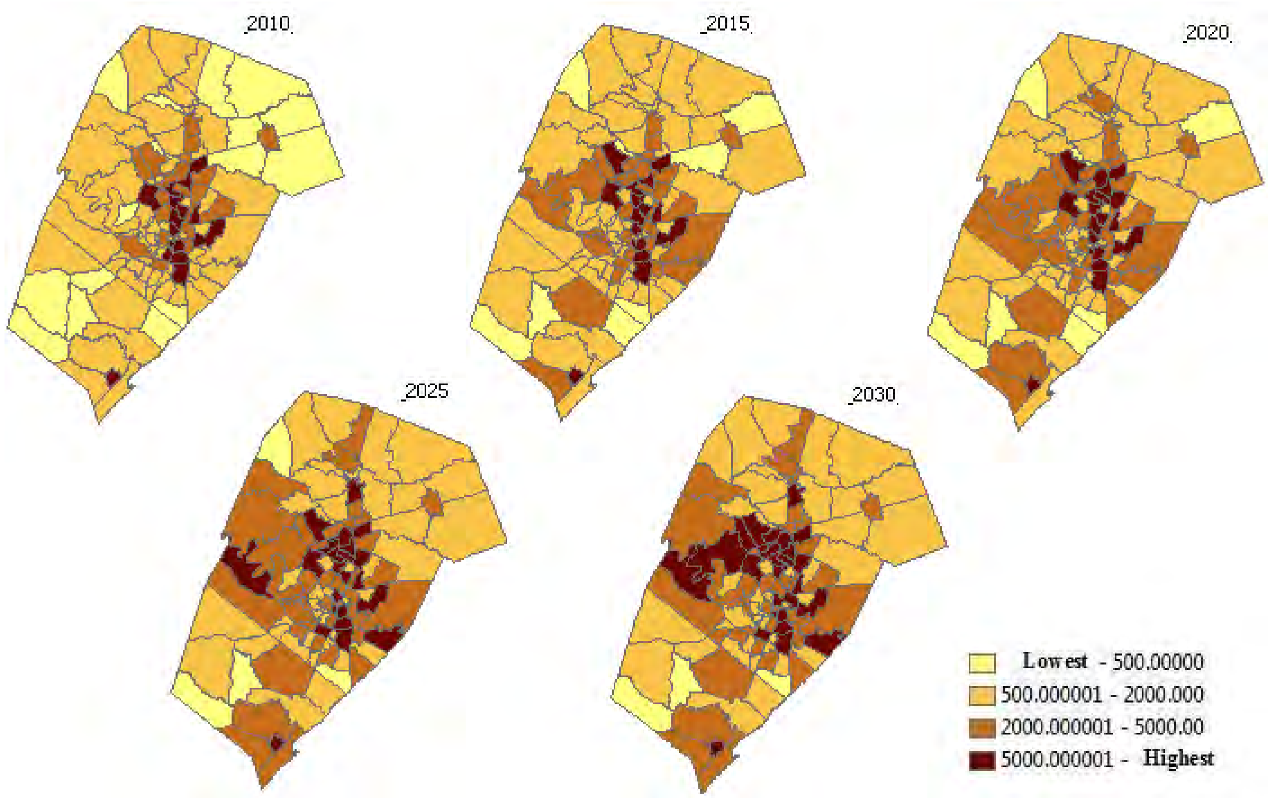


Figure 8.3: TELUM Forecast by District for Basic Employment for the Three-county Austin Region

TELUM predicts that the basic employment will increase in almost all zones in the three-county region. The basic employment is predicted to be high in the outskirts of the city of Austin, particularly in zones west of the city by 2030.

Figure 8.4: MATLAB Code Forecast by District for Basic Employment for the Three-county Austin Region

In contrast, MATLAB code predicts the density of basic employment to be relatively high in central Austin. The majority of growth areas are located close to the city, and only a few zones in Hays and Williamson counties are predicted to experience significant increase in basic employment.

Retail employment

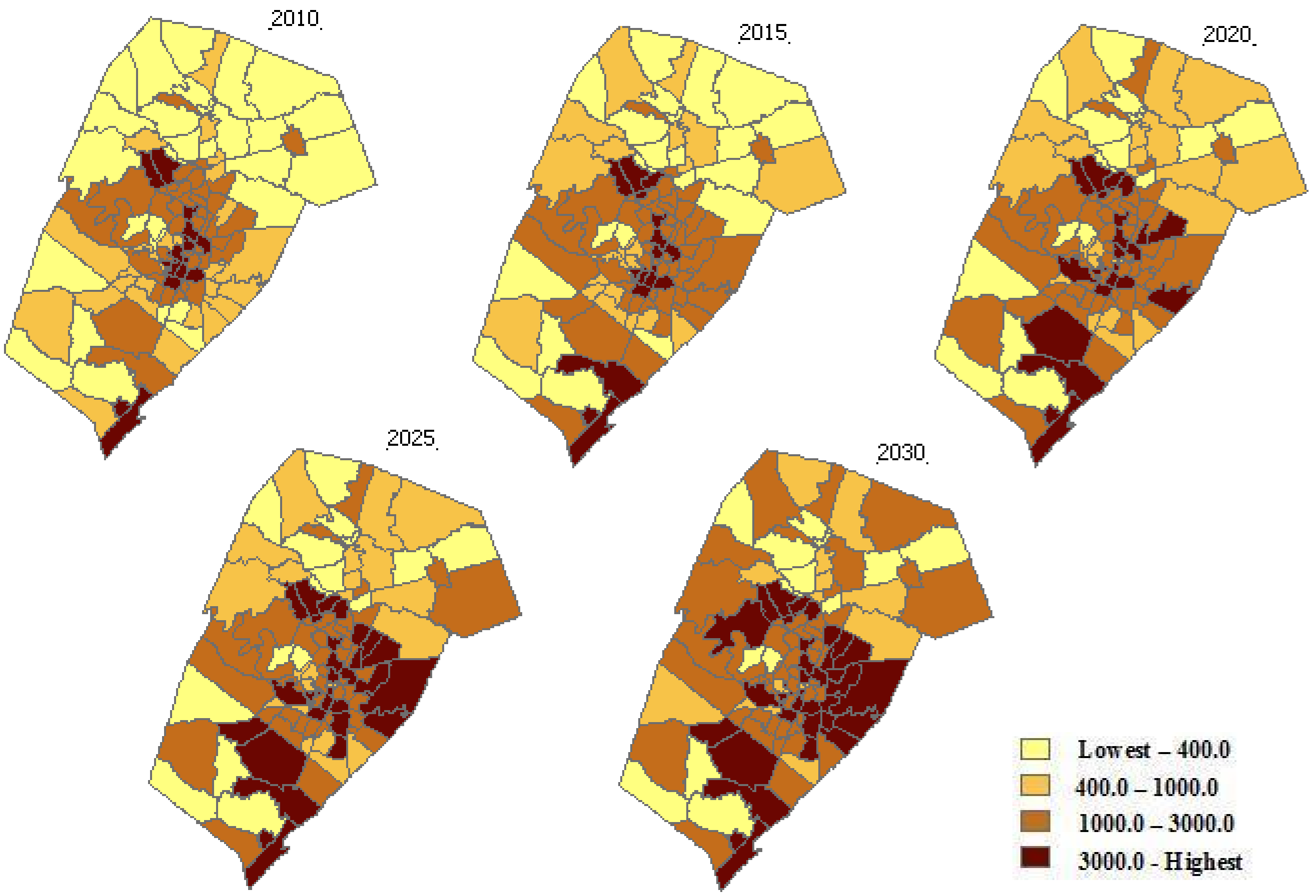


Figure 8.5: TELUM Forecast by District for Retail Employment for the Three-county Austin Region

Almost all zones in Travis County are predicted by TELUM to experience a dramatic increase in retail employment. By 2030, the outskirts of the city are forecasted to have high numbers of retail employment. Some of the zones in Hays and Williamson counties are also predicted to experience a large increase in retail employment.

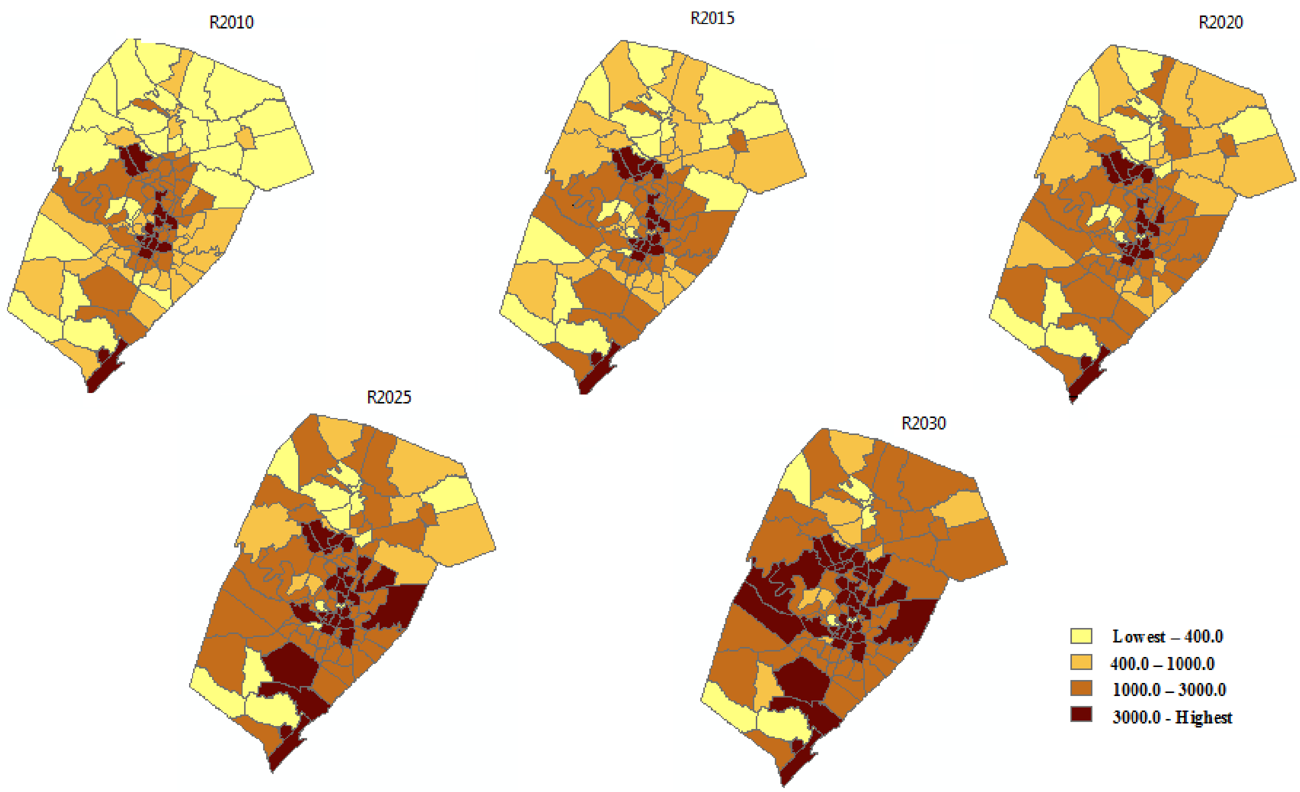


Figure 8.6: MATLAB Code Forecast by District for Retail Employment for the Three-county Austin Region

MATLAB code also predicts a dramatic increase in retail employment levels in the outskirts of the city. However, the spatial distribution of the retail employment predicted by both the MATLAB and TELUM codes differ substantially. The MATLAB code forecasts that areas west of the city will experience high retail employment, while TELUM predicts increases mainly east of the city. TELUM's forecast of retail employment in Hays and Williamson counties largely matches the MATLAB code predictions.

Services employment

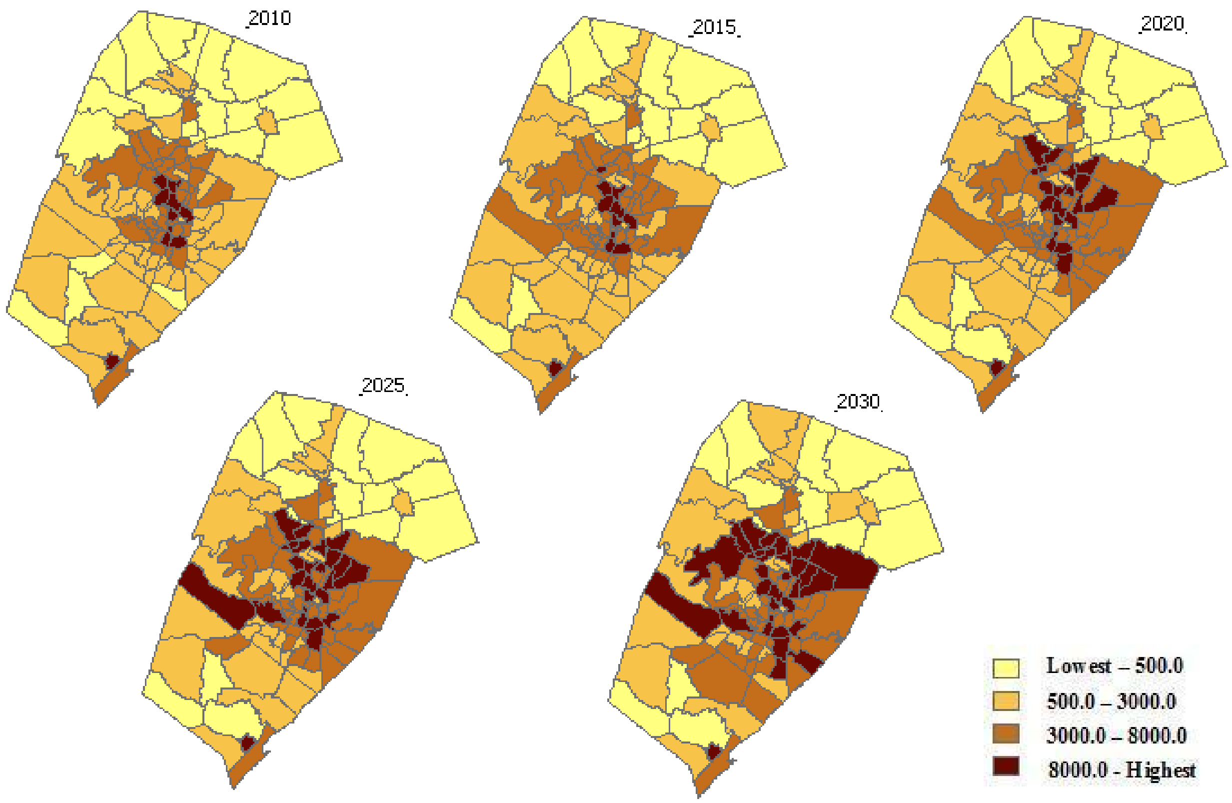


Figure 8.7: TELUM Forecast by District for Service Employment for the Three-county Austin Region

Service sector employment is predicted by TELUM to be high in districts located north of the city by 2030. Two zones south of the city and in Hays County are also predicted to experience dramatic increases in service employment. The zones in Williamson County are predicted to experience only small changes in service employment.

Figure 8.8: MATLAB Code Forecast by District for Services Employment for the Three-county Austin Region

The MATLAB code predicts central Austin to have high levels of service employment by 2030, but TELUM forecasts this growth to happen north of the city center. The MATLAB code predictions and TELUM predictions are nearly the same for the zones in the Hays County. However, their predictions differ significantly for Williamson County. The MATLAB code forecasts a significant increase in service employment levels in Williamson County zones, but TELUM predicts only small increases in such zones.

Low-income households

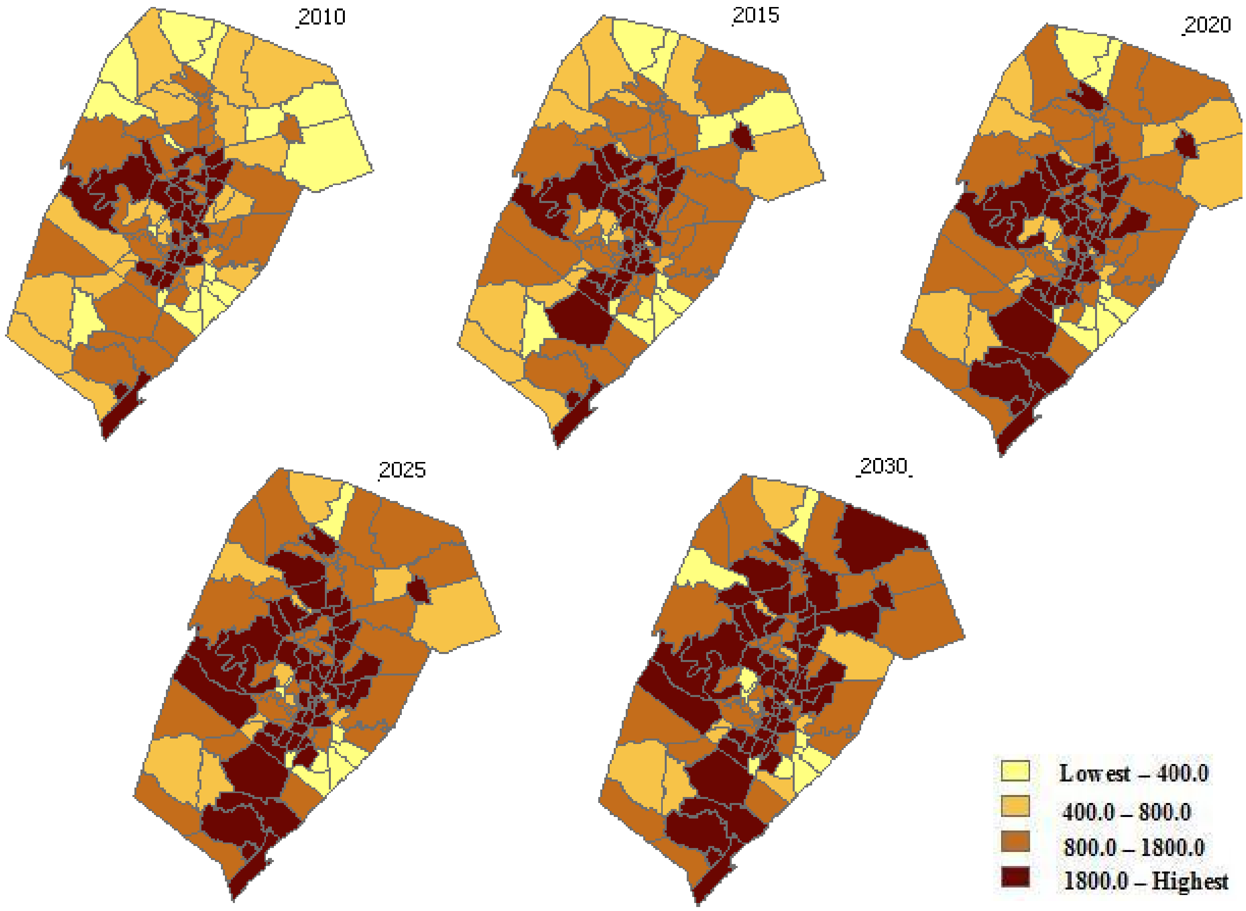


Figure 8.9: TELUM Forecast by District for Low-income Households for the Three-county Austin Region

TELUM predicts the number of low-income households to increase in most of the zones. A large number of these households are predicted to be located in zones close to the center and to the west of Austin. Some zones in Hays and Williamson counties are also predicted to have a high number of low-income households by 2030.

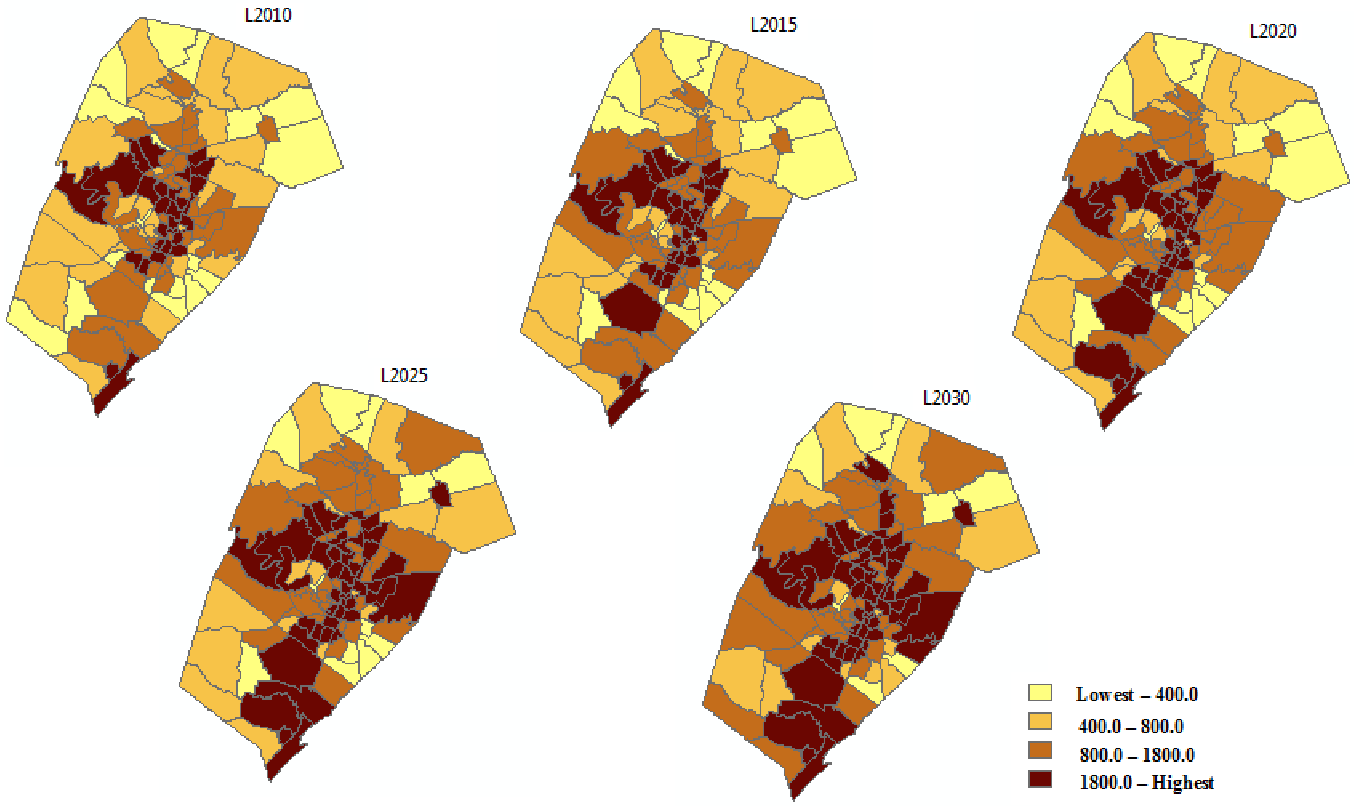


Figure 8.10: MATLAB Code Forecast by District for Low-income for the Three-county Austin Region

The MATLAB code forecast matches the TELUM forecast for most of the zones in Travis and Hays counties. In Williamson County, the MATLAB code predicts small changes in the number of low-income households. TELUM, on the contrary, predicts a large increase in the number of low-income households here.

Below-average-income households

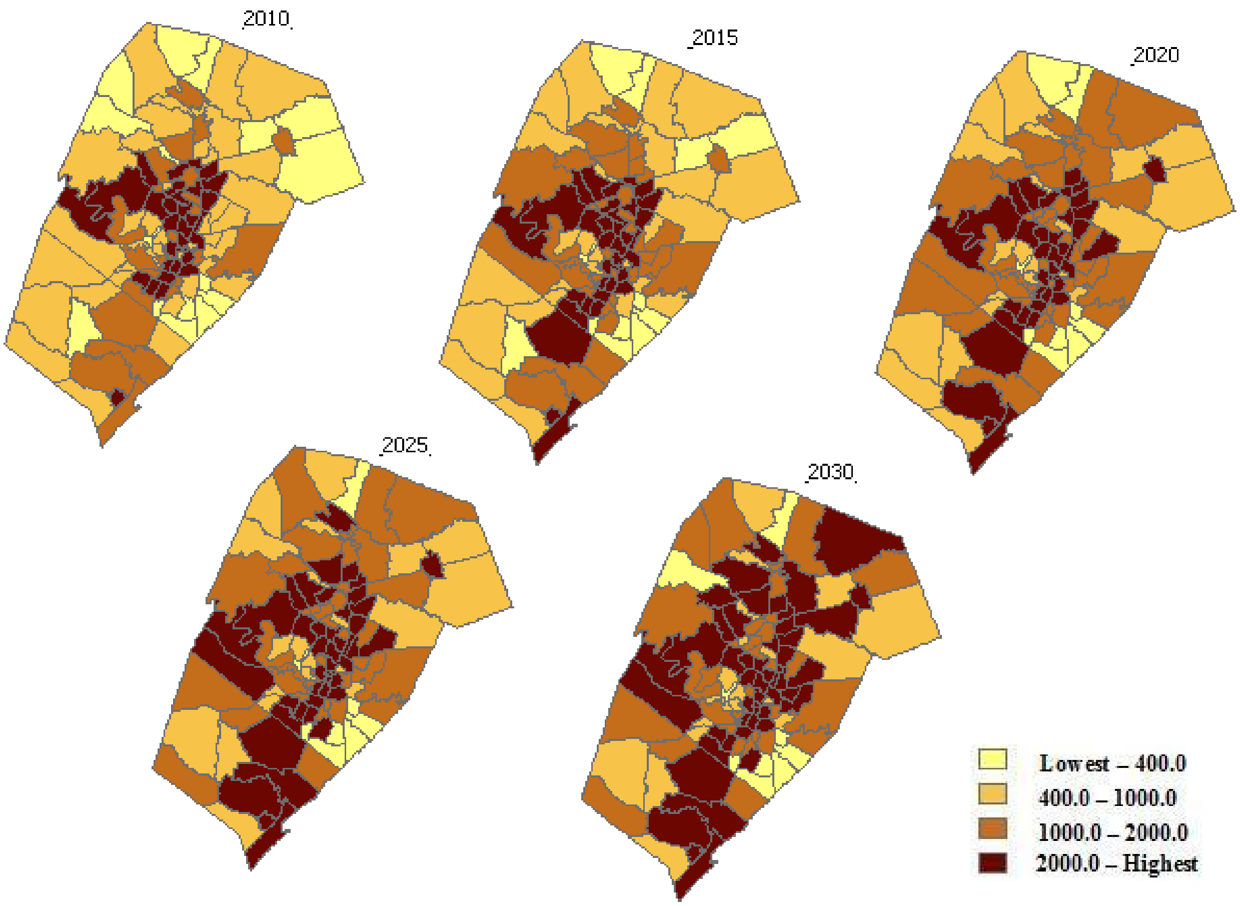


Figure 8.11: TELUM Forecast by District for Below-average-income Households for the Three-county Austin Region

Below-average-income households are predicted to increase in most of the zones in the three-county region by TELUM. Zones close to city center and to the west, along with two zones in Williamson and Hays counties, are predicted to have a high number of these households by 2030.

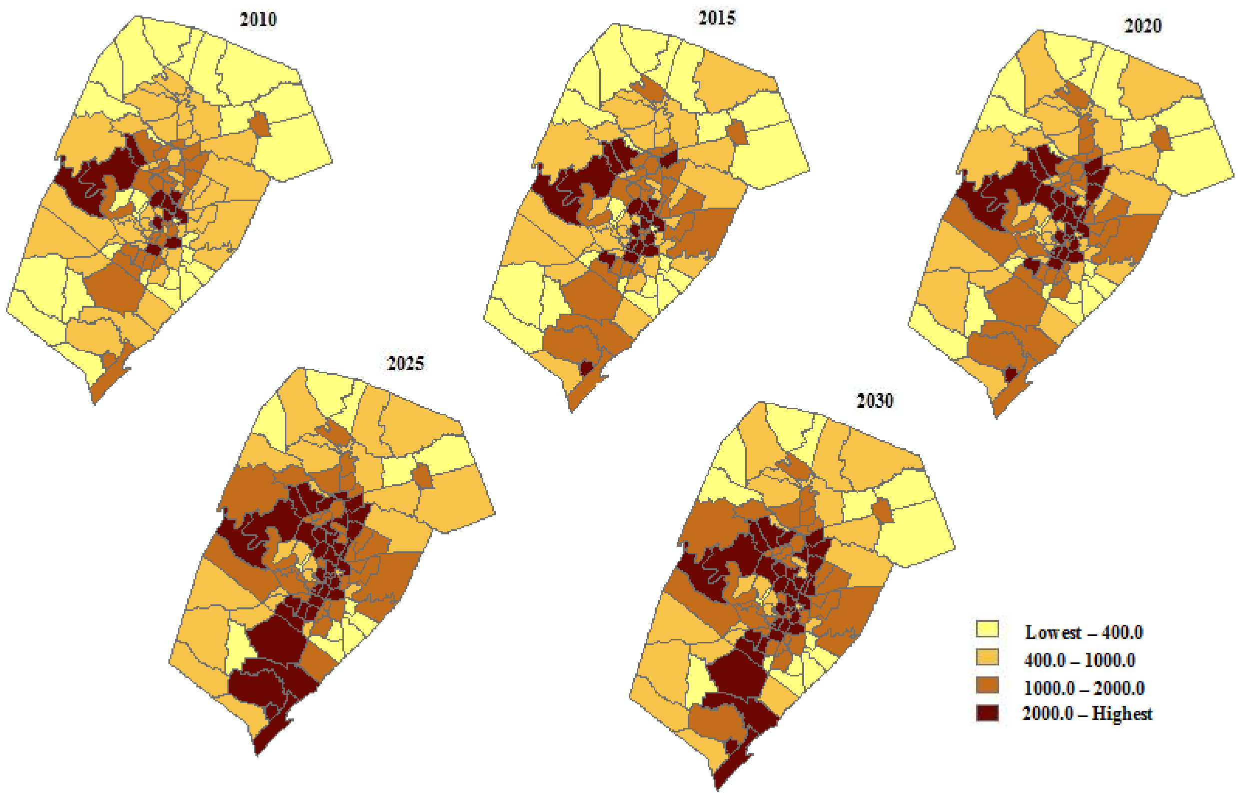


Figure 8.12: MATLAB Code Forecast by District for Below-average-income Households for the Three-county Austin Region

MATLAB code also forecasts the below-average-income households to be concentrated in zones to the west of city and in some zones in Hays County. However, the predictions of the MATLAB code are significantly different from the TELUM predictions for zones in Williamson County and zones located to the east of the city. TELUM predicts the number of below-average-income households in these zones to be much higher than the MATLAB code predictions.

Above-average-income households

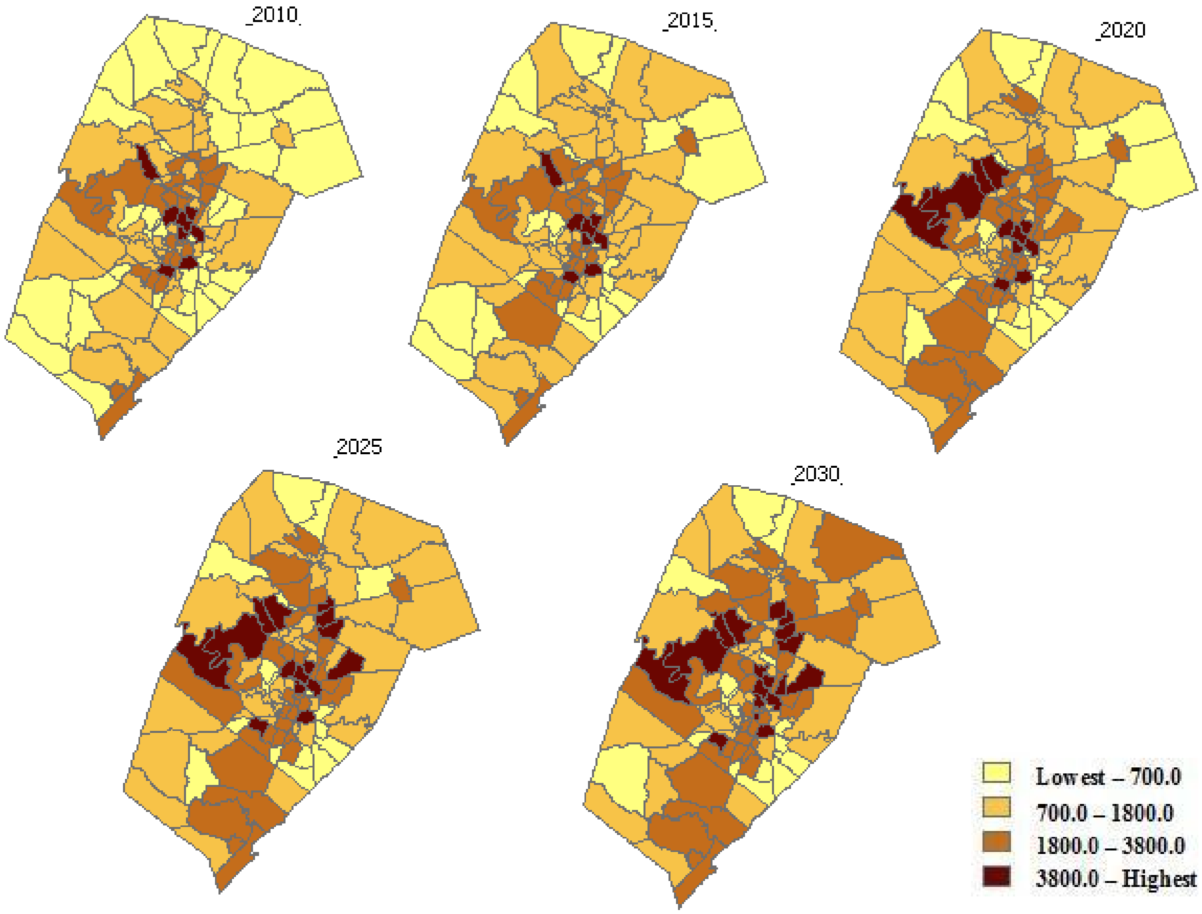


Figure 8.13: TELUM Forecast by District for Above-average-income households for the Three-county Austin Region

The number of above-average-income households is also predicted by TELUM to increase in most of the zones. A large number of above-average-income households are predicted by TELUM to be located to the west and north of Austin in 2030.

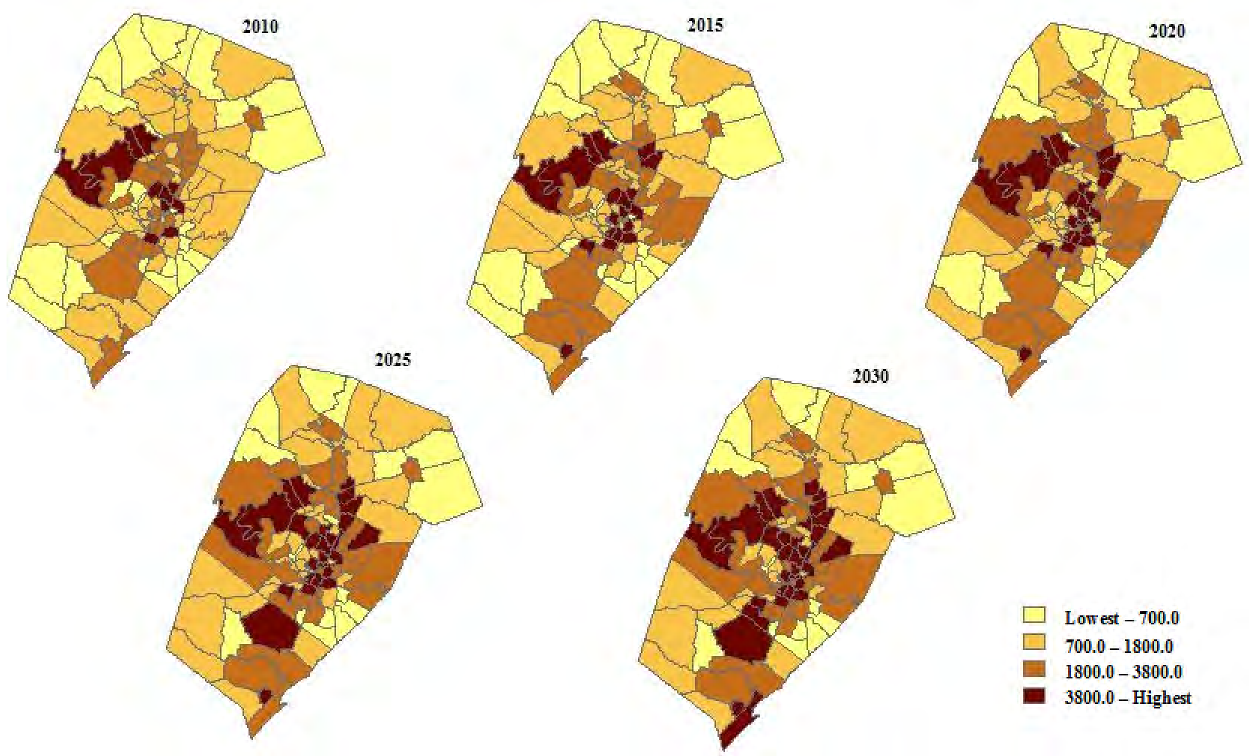


Figure 8.14: MATLAB Code Forecast by District for Above-average-income Households for the Three-county Austin Region

The forecast of spatial distribution of the above-average-income households by MATLAB code matches the TELUM forecast for most of the zones in Travis County. MATLAB code predicts two zones in Hays County to have large number of these households but TELUM does not. MATLAB code also predicts small changes in the number of above-average-income households in Williamson County but TELUM, on the other hand, predicts significant change in the number of above-average-income households in these zones.

High-income households

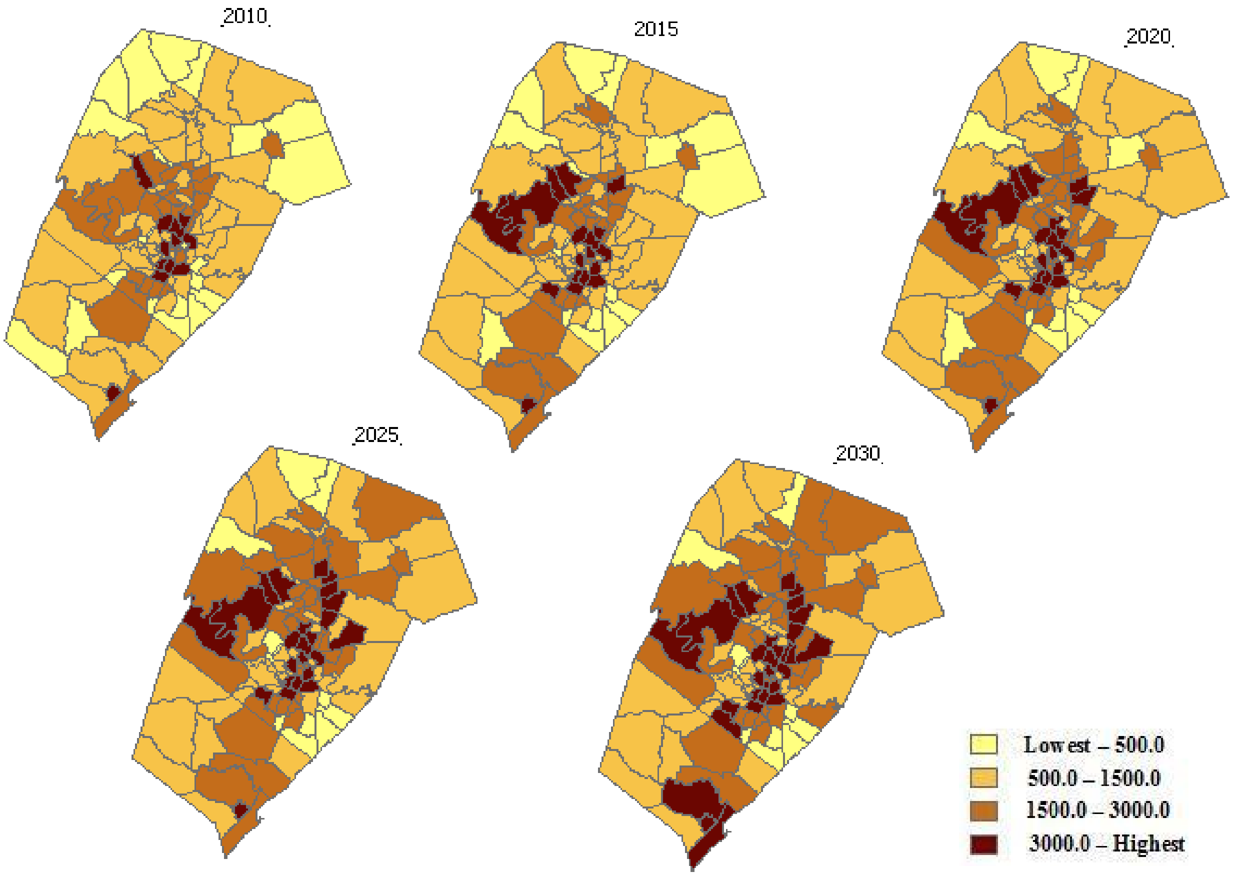


Figure 8.15: TELUM Forecast by District for High-income Households for the Three-county Austin Region

A large number of high-income households are predicted by TELUM to be located in central Austin, some zones to the west of Austin, and in two zones in Hays County by 2030. TELUM also predicts the number of high-income households to increase in most of the zones.

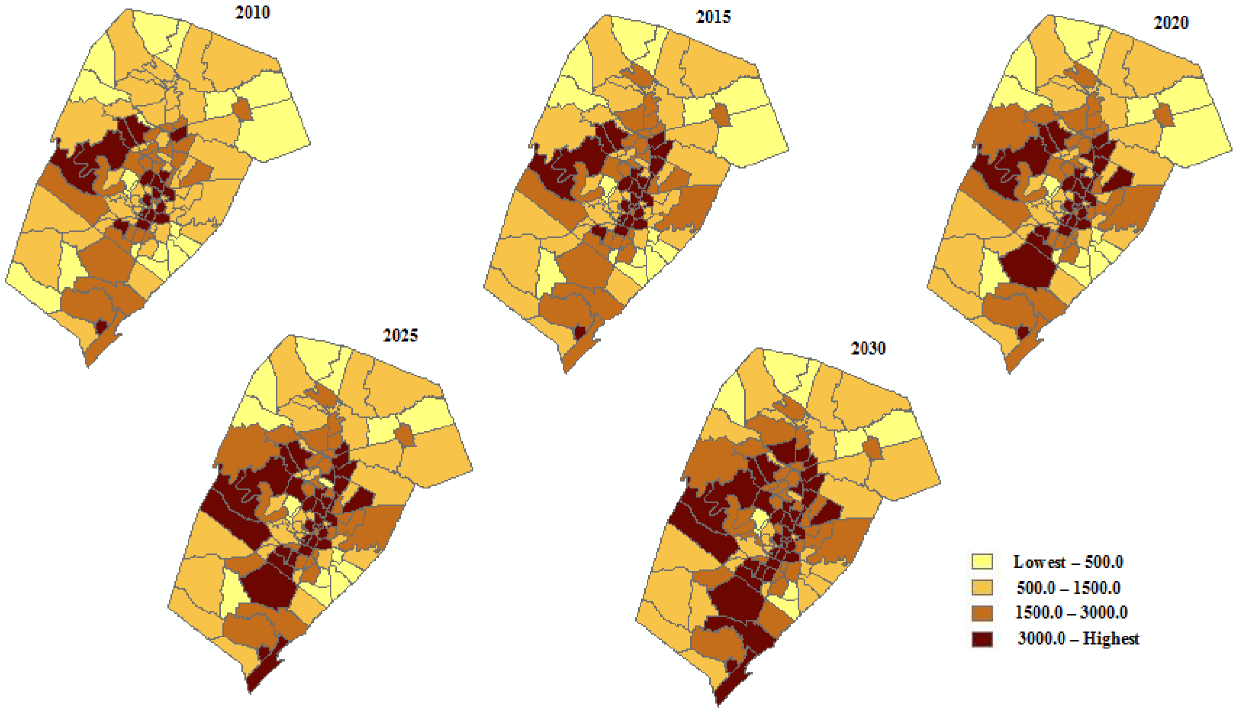


Figure 8.16: MATLAB Code Forecast by District for High-income Households for the Three-county Austin Region

The predictions of the spatial distribution of the high-income households of MATLAB code are nearly the same as the TELUM predictions. The difference is again in some zones of Williamson County where TELUM predicts the number of high-income households to be greater than the MATLAB code predictions.

Summary

TELUM predicts an increase in total employment in most Travis County zones, while MATLAB code predicts job increases really only in City of Austin zones. MATLAB code predicts a significant increase in total employment in many Williamson County zones, while TELUM does not. The spatial distribution of low-income households and below-average-income households predicted by TELUM differs dramatically from that of MATLAB code. However, there are close similarities between forecasts for above average income and high-income households. Hence, in general, there is a significant difference in the spatial distribution of employment and household forecasted by TELUM and MATLAB code. The source for such distinctions is very likely from the unknown formulation of LANCON in TELUM. TELUM's user manual neglects LANCON's specification, and our MATLAB code follows the formulas in Putman's book (1991).

8.3.3 Model Limitations

As one of the most widely applied land use techniques in the world, gravity-based allocation methods enjoy a relatively simple model structure, moderate data demands, and relatively straightforward estimation. Such benefits can be critical for agencies and modelers that do not have the resources on hand for more sophisticated modeling approaches. Clear limitations in opportunities for policy analysis (see, e.g., Lemp, 2007) and poor performance with relatively disaggregate zonal systems and/or lightly developed zones (PBQD 1999) are noted issues.

Given their limitations, such models may not survive close scrutiny. As done in places like Dallas-Ft. Worth, manual adjustments of results in each period may be needed (Dankesreiter 2008). Another option is to make the household and job allocation models more sensitive to land allocation model forecasts. Under the current specification, LANCON does not provide direct feedback to employment allocation model, and enters the household allocation model only via three land use variables. Land availability forecasts should impact the maximum number of new households and jobs that can be allocated to a zone.

In general, the land allocation formulation does not make great sense, and tends to generate unreasonable average land consumption values for households and jobs (as compared to base and prior year land conditions in each zone). This sub-model may be dramatically improved by adding a lagged density term, allowing it to "pivot off" of recent land use densities in each zone, rather than relying on model averages (which presently causes it to ignore current zone densities).

Of course, existing gravity-based LUMs are aggregate in nature and therefore limited by level of detail. For example, land use information at the level of neighborhood zones is generally too coarse for detailed analysis of land cover change and associated biogenic emissions estimation. Such models also are totally unresponsive to certain policies, such as land taxation and subsidies, adequate public facilities ordinances (APFOs), and many styles of zoning. Modelers need to be fully aware of these limitations before seeking to develop and then apply such models. If not, they may be sorely disappointed late in the modeling process.

Chapter 9. Examination of UrbanSim's Sensitivity to Congestion Effects

This section describes the process used to evaluate the sensitivity of UrbanSim's land use predictions to congestion. This was done in order to appreciate how sensitive land use patterns may be to transport costs, under the UrbanSim framework, and to simply ensure that such a relationship emerges when transport conditions vary. It should be noted that recent work[9] by a student of the RS examined the effects of road pricing using the gravity land use model (G-LUM), in tandem with a travel demand model. Those results are not discussed at any length here. However, it is interesting to summarize the results: Essentially, road pricing (at about 4 ct/mile, and higher during peak periods on congested freeways) had a negligible impact on near and long-term land use patterns, though it did have a significant effect on VMT (roughly a 15% reduction) (Zhou et al., 2008). Interestingly, the urban growth boundary scenario tested had the same overall VMT effect, along with a significant land use effect (when using the G-LUM, calibrated for the Austin region).

9.1 Model Development

Congestion is a major negative externality imposing costs on busy regions and their travelers. Addressing congestion is a major objective of most transportation agencies, and land use and transport policies in tandem may offer some benefits. This chapter describes the procedures used to ascertain changes in UrbanSim's land use pattern predictions, relative to the business-as-usual (BAU) scenario, following an adjustment of travelers' sensitivity to travel costs.

As described in Chapter 8 of this report, outputs from UrbanSim were fed manually into TransCAD for 5-year periods for the TDM, much as TxDOT and Texas MPOs are likely to do (and as illustrated by Figure 9.1). In this study, in order to effectively congest the network, the parameter on generalized cost (β_{GC}) was doubled. This essentially makes all travelers twice as sensitive to travel costs, behaving as though travel is very costly (or very time-consuming, because time is money), making trip distances shorter and the relative accessibility of well-located zones even higher.[10] The relevant equations are as follows:

$$U_{nijmt} = \beta_{1nmt} + (2 * \beta_{GC,n}) * (\frac{Cost_{nijmt}}{VOTT_n} + TT_{nijmt}) \tag{9.1}$$

[9] A US EPA STAR grant project funded much of the G-LUM modeling work. Lengthy reports on this topic are available for anyone who is interested.

[10] Trying to "double congestion" is not a straightforward operation. For example, if analysts tried to keep doubling travel times after every network assignment, the outputted results would be incompatible with travel choices. A much better approach to forcing congestion on the network is to reduce capacities of all links, but halving capacities will likely more than double travel times on congested corridors. A final approach considered briefly here was to double the number of travelers (by doubling population and jobs), but that approach completely confounds the LUM's operation, which is the basis for this experiment. Thus, we chose to go with the doubling-cost-sensitivity approach (which means less travel, and a less congested network, but the land use market response is what we are seeking, as translated through the lower AI values).

where the generalized cost coefficient, $\beta_{GC,\,n}$, is now doubled, and the associated logsums, which feed into the accessibility indices (as described in chapter 8), are calculated as follows:

$$L_{ij} = \ln\left(\sum_{m,\in C} \exp\left[U_{nijmt}\right]\right) \tag{9.2}$$

$$A.I._{i} = \sum_{j=1}^{J} D_{j} e^{L_{ij}} \tag{9.3}$$

9.2 Model Results

As noted in Chapter 8, the travel times and accessibility indices for home-based work trips by households with and without cars in the AM peak were calculated and used in UrbanSim. As a result of a doubled “sensitivity” to travel costs in the TDM specification, changes in AI values lead to changes in long-term land use patterns, as illustrated by Figure 9.2 (between the BAU and “congested” scenarios).

As shown in Figure 9.2(d), which is the difference between Figure 9.2(b) and 9.2(a) results, households tend to move closer to the freeways and towards cities such as Leander, Jonestown, Lago Vista, and Lakeway. The reason for the attractiveness of these relatively rural zones is probably their lower land and improvement values, coupled with relatively high local accessibilities (simply due to being town centers). Regional accessibilities are the variables that rely on travel costs, but they do not play a role in the current model specification (as shown in the Appendix E of this report), due to a lack of statistical significance and sometimes unexpected/improper sign in early model development. These earlier specifications are being revised, to avoid multi-collinearity in accessibility terms, allowing for regional accessibility variables to play more appropriate roles. Those results will be described in this project’s final report.

Finally, Figure 9.3 illustrates the predicted patterns of land use, and 9.3(d) shows that there is no significant difference in employment location choices over time when sensitivity to transport cost is doubled. This insensitivity in employment patterns is due to the statistical insignificance of regional accessibility terms in the model specification (Appendix E of the report) as well as the lack of predictive power for at least two of the employment-based location choice models. More interesting results can be expected when specifications are enhanced. In the meantime, we are pleased that this computationally demanding, data intensive and complex model system is running “properly.” And we look forward to further case studies and specification modifications.

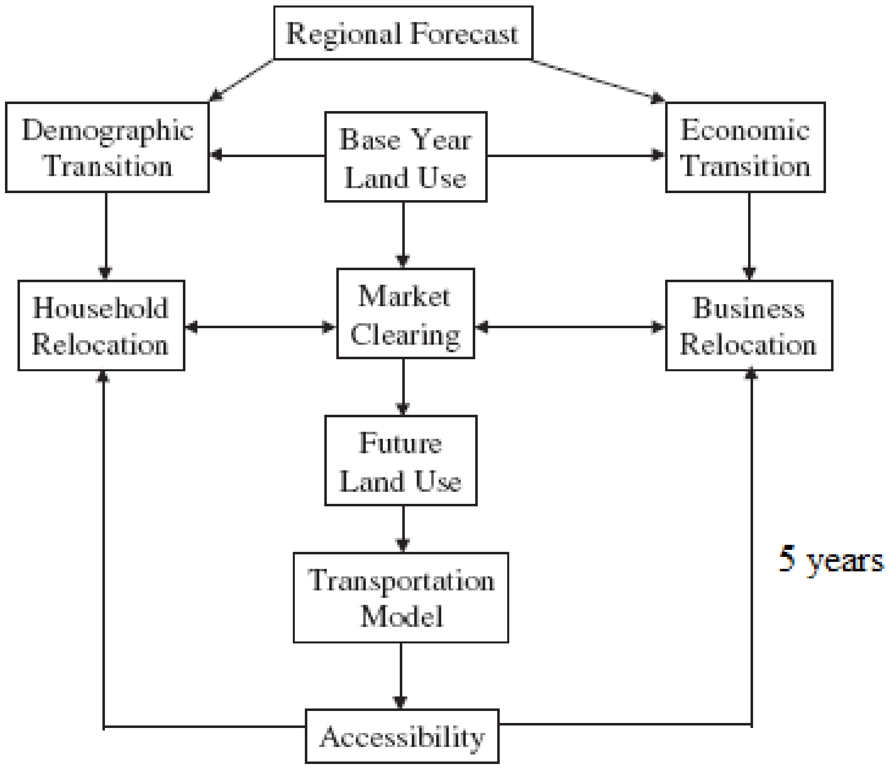
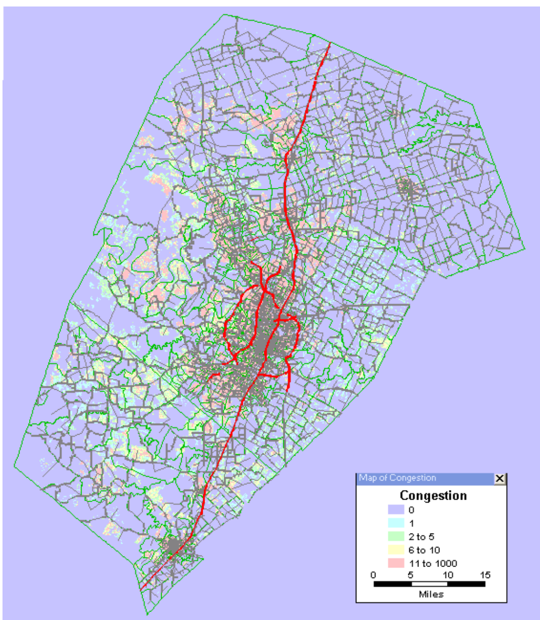


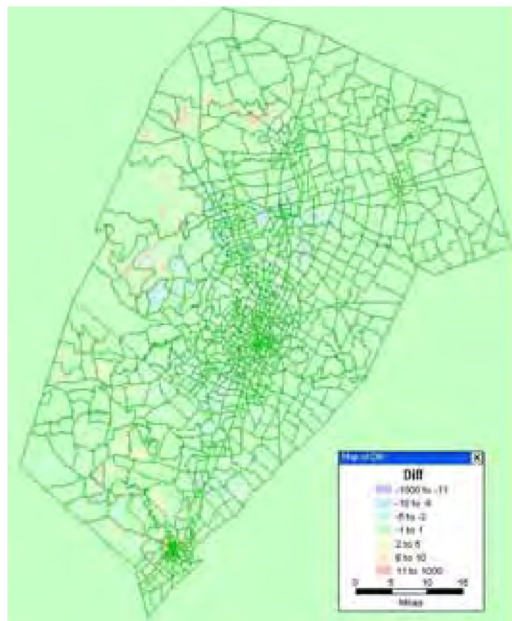
Figure 9.1: Flowchart for Integration of UrbanSim and a TDM

a) BAU

b) Doubled travel-cost sensitivity



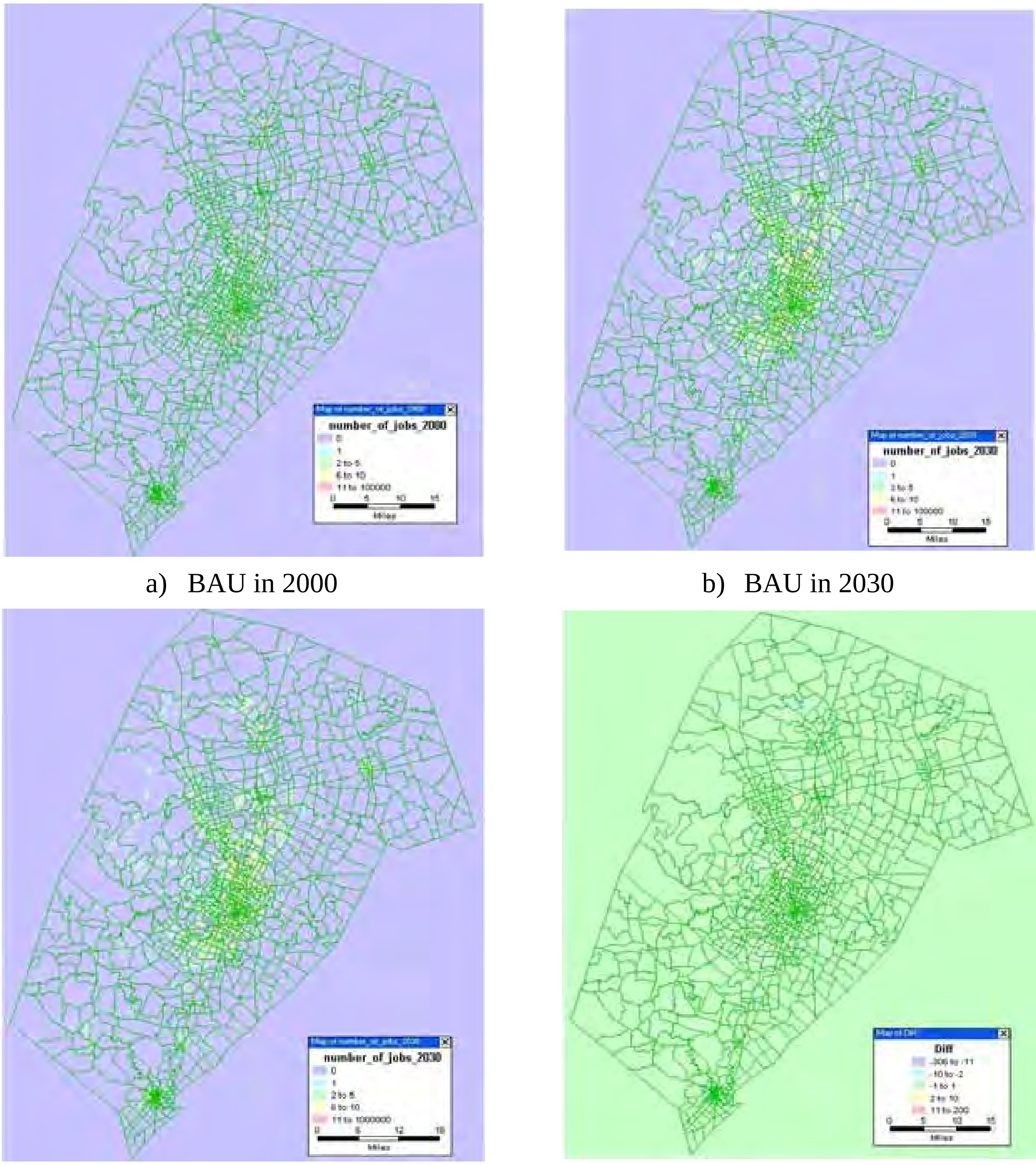
c) Depiction of (a) inserted on a network file



d) Difference in values between (b) and (a)

Figure 9.2: Year 2030 Household Locations, with and without Added Congestion

Note: Units are persons per grid cell, and 150 m x 150 m is approx. 5.2 acres.



a) BAU in 2000

b) BAU in 2030

c) Doubled travel-cost sensitivity

d) Difference in values between (c) and (b)

Figure 9.3: Employment Locations with and without Added Congestion

Note: Units are jobs per grid cell, and 150 m x 150 m is approx. 5.2 acres.

Chapter 10. Examination of Uncertainty in an Integrated Model: Simulation Errors versus Parameter and Input Data Errors

Major transportation improvements are costly investments, often made with the help of mathematical models to anticipate future transportation conditions. Given budget constraints and the wide-reaching and long-term nature of major project impacts, selection of system improvements should be done with care. This section of Research Project 0-5667 is focused on how variations in model parameters and inputs affect land use and transportation predictions, and how the formal recognition of uncertainty in integrated land use-transport modeling exercises may change investment decisions, thereby resulting in decision-making that is more robust to future conditions. The decisions considered in this paper are limited to roadway improvements (based solely on predicted v/c ratios) and land use policy (based on four simplified scenarios).

Integrated modeling of land use and transport is complex, and relatively few studies have evaluated the effects of input and parameter uncertainty in this context. Those that have done so focus primarily on the variation in model outputs. For example, Smith and Shahidullah (1995) evaluated the performance of 10-year population forecasts (a key land use model input) by age group for three very different census tracts, and found mean absolute percent errors to range from 15% to 20%. Zhao and Kockelman (2002) measured the propagation of uncertainty through a four-step travel demand model and concluded that trip assignment (to the network) serves to reduce uncertainty levels in outputs to initial levels of input uncertainty (rather than widening, for example), and that employment estimates, trip generation rates, and mode choice parameters are critical determinants of output variations (e.g., link flows and total VMT)—and that such final outputs can be closely predicted via relatively few input values (using a regression of their 100 simulated outputs). Pradhan and Kockelman (2002) extended Zhao and Kockelman's study by examining propagation through an integrated transportation-land use model (ITLUM) using UrbanSim and a four-step travel model for Eugene, Oregon; they concluded that in the long run, only those inputs that have a cumulative effect (such as population and employment growth rates) are likely to have a significant effect on model outputs.

Krishnamurthy and Kockelman (2003) conducted another study of uncertainty in ITLUMs, this time using a Disaggregated Residential Allocation Model (DRAM®) and Employment Allocation Model (EMPAL®) based on the DRAM®-EMPAL® Lowry-type models developed by Putman (1983), along with a four-step travel model in TransCAD. This research was conducted on the three-county Austin metropolitan region and found output variations to be most sensitive to the exponent (of the volume-to-capacity ratio) in standard link performance functions, the split of trips between peak and off-peak periods, and trip generation and attraction rates.

More recently, Sevcikova et al. (2007) illustrated the benefits of Bayesian Melding over simple random sampling in their quest to calibrate traffic analysis zone-level household counts using UrbanSim for Eugene, Oregon. Random numbers were used within the economic and demographic transition models, employment and household mobility models, employment and household location choice models, and the land development model for a total of 843 uncertain parameters. Two sets of 100 simulations were run, using a different random number seed for each set. The Bayesian Melding procedure led to wider confidence intervals for various outputs that are more likely to contain the true result than if a simple random sampling procedure had been used. Related to this, Gregor's (2007) new land use model (LUSDR) makes uncertainty explicit through multiple model runs, in order to find the best extension of Jackson County, Oregon's urban growth boundary.

While this study may be the first to consider how investment decisions change given uncertainty in an ITLUM context, it is not the first to consider this topic in the travel model context, particularly with respect to traffic assignment. Lam and Tam (1998) used Monte Carlo simulation methods to study the impact of uncertainty in traffic and revenue forecasts for road investment projects. They assumed normal distributions for each of several uncertain parameters, including population and demand elasticity. Moreover, Waller et al. (2001) assigned independent distributions for each origin-destination pair's future year demand in three test networks (ranging from two to 100 origin-destination [OD] pairs) and demonstrated how assignment models relying on expected values of all inputs will tend to underestimate future congestion and may (in 14% of cases studied) lead to selecting projects with higher average (future) travel costs (i.e., lower net benefits) than ideal, and higher variance in such costs (which implies more risk). Duthie et al. (2006) extended this earlier research to allow for correlations in trip-making between OD pairs and showed how neglecting correlations when they exist will lead to errors in future travel cost predictions and suboptimal project selections (2% to 50% of the time, depending on the correlation structure and the objective function used). Also, Rodier and Johnston (2002) found that errors in county population forecasts can affect whether or not the Sacramento region meets air quality conformity. The authors' focus, however, was sensitivity analysis (as opposed to uncertainty analysis), so uncertain parameters (i.e., population growth, income, and fuel prices) were only varied by a few percentage points, and counties with outlying population growth rates were eliminated. Harvey and Deakin (1995) considered uncertainty in population growth, fuel prices, and household income levels in their Short-Range Transportation Evaluation Program (STEP) model and found that plausible ranges of the input variables resulted in VMT values that differed by -25% to 15% (relative to the original prediction).

This work builds off the existing literature and seeks to fill the gap that exists in regards to improved land use and transport decision making. The great need for incorporating uncertainty into the decision making process has been noted by many researchers and practitioners (e.g., Ascher, 1979; Khan, 1989; Gifford, 1994) and the types of uncertainties presented in ITLUMs have been categorized many different ways (e.g., Mahmassani, 1984; Niles and Nelson, 2001; Armoogum et al., 2002). The research presented in this chapter focuses on the uncertainty inherent in regional population and employment forecasts, as well as parametric uncertainty, and uses a Lowry-type land use model and a four-step travel model to appreciate how decision making may differ when uncertainty is recognized.

The following sections describe the modeling process and the Latin Hypercube (LH) methodology used for uncertainty analysis. These are followed by experimental analysis on a sample region, results, conclusions, and insights gained from the research.

10.1 The Model

The Integrated Transportation and Gravity-Based Land Use Model (ITGLUM) described in this chapter was created by the research team for this project and a related US EPA-sponsored project. The transportation module of ITGLUM (contained within Figure 10.1's bottom gray box) is a very basic four-step model of trip generation, trip distribution, mode choice, and route choice. Readers may wish to review Martin (1998) for a description of these steps. The land use module of ITGLUM (contained within Figure 10.1's top gray box) is based on Dr. Putman's documentation of ITLUP equations (Putman, 1983), which represent a Disaggregated Residential

Allocation Model (DRAM®) and Employment Allocation Model (EMPAL®). The land allocation module, LUDENSITY, was developed by the authors (however, it's placement in the model series is based on Putman's LANCON model [Putman, 1983]). Calibration determines the value of parameters that generate the best forecast for the base year based on the lagged year data. Each sub-module is described in Appendix F.

10.2 Methodology

This research focuses on the decision-making impacts of uncertainty in future regional control totals of households (by income group) and employment (by sector), and in the trip generation and distribution parameters. Each uncertain parameter is assumed to follow a lognormal distribution with coefficients of variation (COVs) equal to 0.1, 0.3, or 0.5 (as tested in Zhao and Kockelman [2002]). Sampling is done via the LH technique (described in Appendix G), which Mathé (2000) proved to have asymptotic superiority over simple random sampling for all distributions. LH sampling is a multi-dimension extension of the Latin Square, where a two-dimensional sample space is split into a grid (where each cell along each dimension has an equal probability of occurrence) and exactly one sample point is chosen from each row and each column. McKay et al. (1979) first introduced LH sampling and proved its superiority over random sampling for estimating a function mean and distribution function given certain monotonicity conditions. LH has also been shown to work well for the traffic assignment problem (Duthie et al., 2006).

The decision evaluated here is as follows: how many miles of roadway need improvement? Essentially: how many miles of roadway are in danger of a volume-to-capacity (v/c) ratio above 0.85 by year 25 (i.e., 25 years after the base year)? (Note: Each roadway is one lane in each direction, and lane-miles and roadway miles are defined equivalently here.)

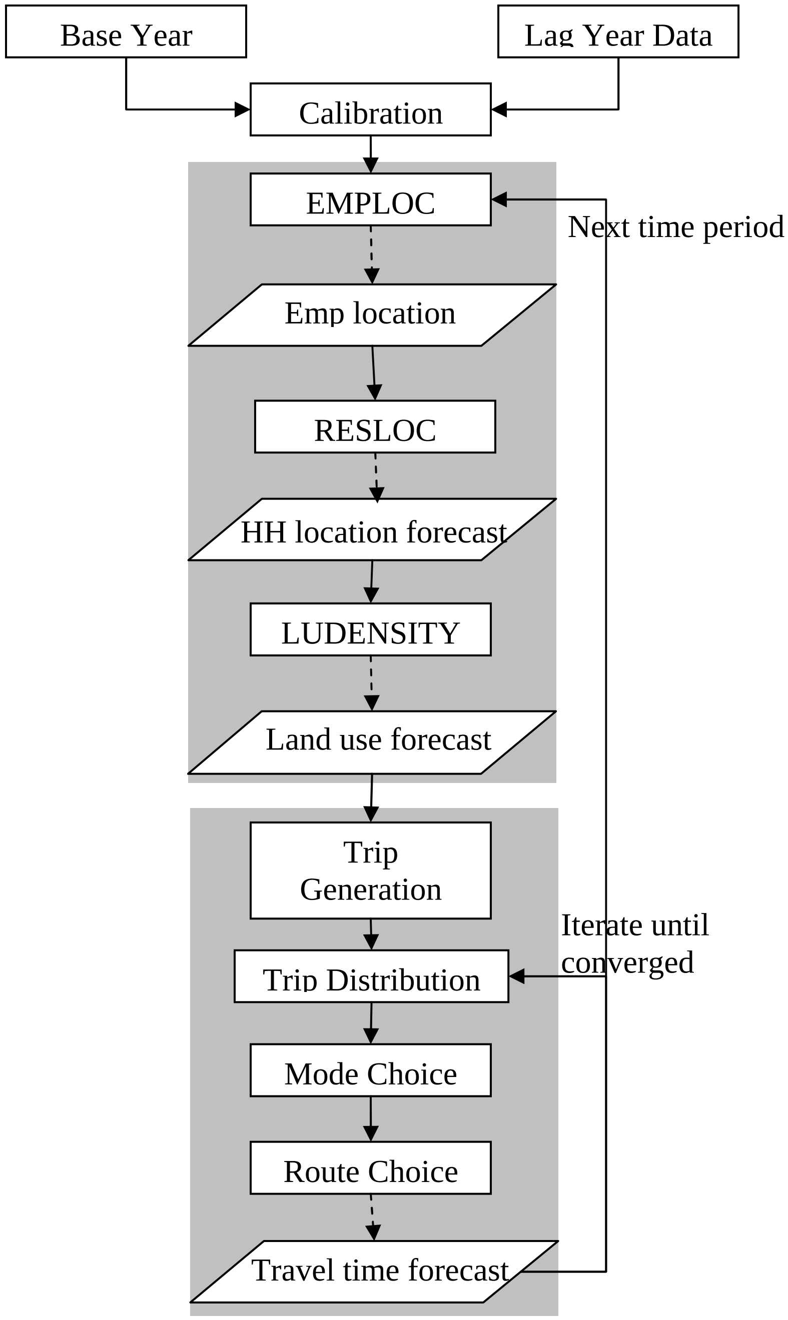


Figure 10.1: Block Diagram of ITGLUM

Five land use policies are considered: The first places no restrictions on development. The second requires all basic employment to be allocated to the 16 outer zones. The third constrains 50% of employment growth to occur in the region's core (zones 8, 12, 13, 14, and 18). The fourth constrains 50% of residential growth to occur in the region's core. And the fifth constrains 50% of all growth to occur in the region's core. It is assumed that roadways with volumes greater than 85% of link capacity in the future year (in this case, 25 years into the future) will be improved, so a comparison is made between the number of miles improved when uncertainty is neglected and considered (i.e., when traffic volumes are evaluated at their expected values). These land use policies were enforced by applying the relevant portion of the control total to the appropriate section of the region. For example, if 50% of the employment has to be in the core, then development in the core was scaled to meet 50% of the employment control totals. Similarly the rest of the region was scaled to meet the other 50% of the employment control totals.

For each covariance level and each land use policy, six batches of 80 samples of the uncertain inputs and parameters were generated (based on the rules-of-thumb provided in Stein [1987]). Average v/c ratios were calculated according to the formulas in Appendix G. These average v/c ratios are then compared with the ones obtained from a deterministic analysis that uses mean values for all input data and parameters.

10.3 Experimental Analysis

As shown in Figure 10.2, a highly idealized sample region (with zones denoted by gray lines) and roadway network (with nodes and links denoted by black lines) allow for ready analysis while helping clarify the impact and sources of modeling results. (Modeling an actual region, such as Austin, would likely obscure the sources of differences in deterministic and stochastic investment results.)

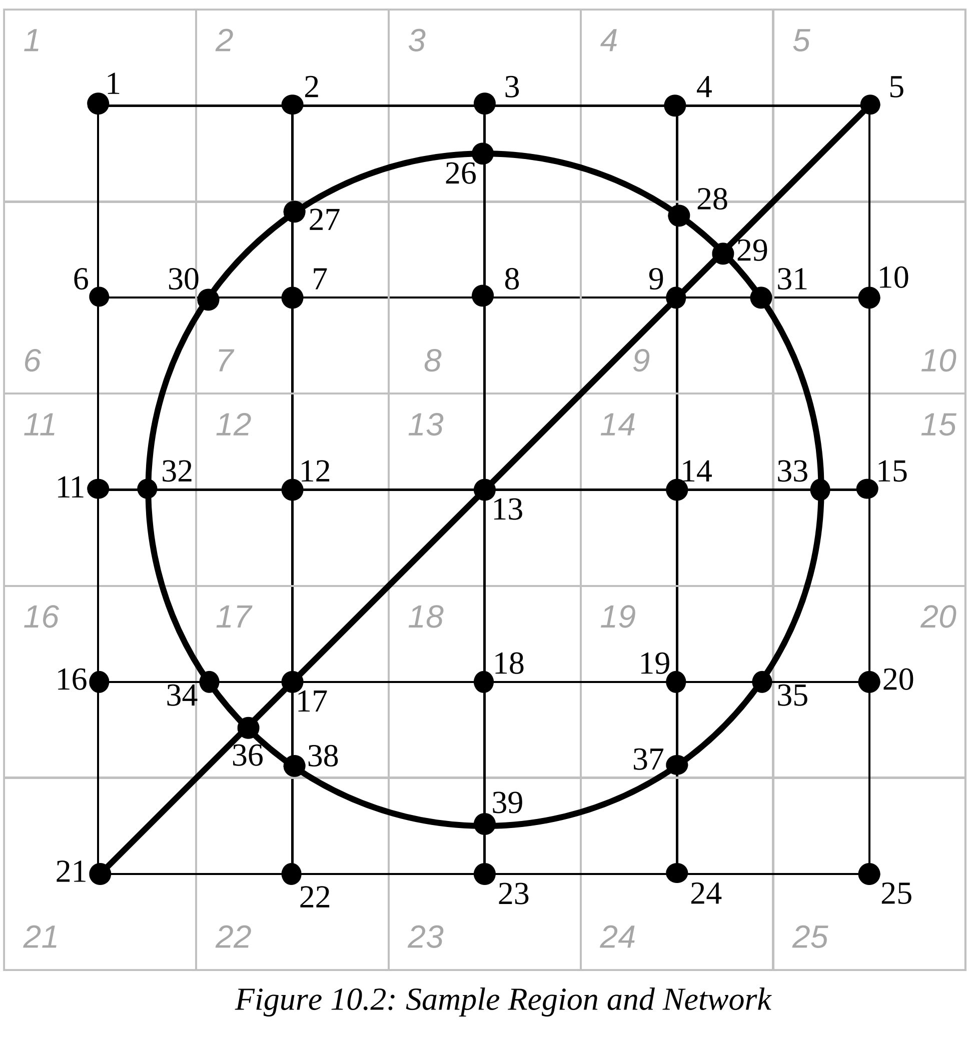


Figure 10.2: Sample Region and Network

Each zone is 1 square-mile (or 640 acres). Highways (circumferential and diagonal) are denoted by heavy black lines, while local roads (in a grid pattern) are shown by lighter black lines. Capacity on local roads is assumed to be 400 vehicles per hour per lane (vphpl), while highway capacity is 2000 vphpl. All roads are bi-directional, having one lane in each direction. Free flow travel speeds are assumed to be 63.6[11] miles per hour (mph) on highways and 30 mph on local roads. Mean values of regional control totals for employment by sector and household by income, developed using ad hoc rules of thumb for "reasonable" development, are given in Table 10.1.

Table 10.1: Regional Control Totals (Mean Values Used)

	Employment (by sector)			Households (by income)			
Time Period	**Basic**	**Retail**	**Service**	**Low**	**Med-Low**	**Med-High**	**High**
1	4313	7188	17250	345	1150	1495	460
2	4959	8266	19838	397	1323	1719	529
3	5703	9505	22813	456	1521	1977	608
4	6559	10931	26235	525	1749	2274	700
5	7543	12571	30170	603	2011	2615	805

Table 10.2 gives the parameters for the EMPLOC and RESLOC sub-models obtained from calibrating Equations E1 and E2 in Appendix F, based on assumed base and lag year conditions presented in Appendix H's Tables H.1 and H.2 (also created using ad hoc rules of thumb for "reasonable" development). Table 10.3 gives the parameters for the LUDENSITY sub-model, obtained from calibrating Equation E3 in Appendix F. The parameters listed in the following tables and their associated equations are explained in Appendix F. Equation F3's L_0 value was assumed to be 0.5 acres per basic job, 2.5 acres per commercial job, and 6 acres per household (based on approximations of the maximum amount of land typically allocated to these uses).

Table 10.2: Calibrated Parameter Values for EMPLOC and RESLOC Sub-Models

	EMPLOC (by sector)			RESLOC (by income)			
Parameter	**Basic**	**Retail**	**Service**	**Low**	**Med-Low**	**Med-High**	**High**
η	0.17	0.089	0.33	0.056	0.064	0.05	0.085
ρ	0.13	-2.7	-1.76	-0.012	-0.02	-0.001	-0.0017
r	-	-	-	-1.6	-1.22	0.62	2.72
q	-	-	-	1.37	0.36	0.86	0.89
b	-	-	-	-45.9	-7.31	-11.1	-16.4
ω	0.019	0.008	0.0025	-	-	-	-

[11] The odd number for free-flow speed derives from averaging reasonable free-flow travel times for approximate freeway link lengths.

Table 10.3: Calibrated Parameters for the LUDENSITY Sub-Model

	LUDENSITY (by use)		
Parameter	**Basic**	**Commercial**	**Residential**
β_1	0.0021	0.1347	0.0001
β_1	1	0.5498	1.1182

All modeling was done in MATLAB on a Dell Optiplex GX620 with 4 CPUs, 3.4 GHz, and 2GB of RAM. Each run of ITGLUM took approximately four minutes.

Several interesting results arise from the LH simulations (with details given in Table 10.4). First, as the coefficient of variation increases, the miles of roadway experiencing a year-25 v/c ratio above .85 decreases for all land use policy scenarios. An interesting extension of this research would be to examine the effect of positive correlations between household and employment control totals, to check whether the Waller et al.'s (2001) result—that positive correlations (versus an assumption of independence) between interzonal travel demands lead to significantly higher total system travel times—hold here as well.

Second, the sections of roadways with v/c exceeding 0.85 in the stochastic cases are always a subset of the roadway improvements recommended in the deterministic analyses (though rankings generally differ). Third, in every case except one (out of 15 cases), the roadways expected to have v/c ratios over 0.85 based on the deterministic analysis were the ones with volumes closest to capacity in the actual (simulated/uncertain-input) cases. The one exception was the '50% Total in Core with COV = 0.5' scenario, where there is one link that was not selected in deterministic analysis (DA) that has a higher volume to capacity ratio in uncertain analysis (UA) than another link that was selected in DA. This suggests that deterministic analyses, which are the norm, are likely to provide a reasonable idea of the roads most likely to experience future year congestion, though the rankings of these roads and predicted v/c ratios will differ. If the level of uncertainty is very high, then it may be only the most congested link as determined by DA that is also considered highly congested in UA.

It is also useful to note that the fewest roadway improvements (according to the UA) are needed when 50% of the region's *total* growth is assigned to the region's core. This may be expected because shorter trip distances lead to increased transit use and fewer trips loading roadways outside the core. As expected, the majority of the roadways improved in this case are within the core. In all other cases, all roadways needing improvements were those connected to the ring road. Interestingly, all congested roadways, in this example, were local roads. (Evidently, there were too many freeways of too high capacity to become congested.) In slight contrast to the *total* growth in the core scenario, the DA results suggest that constraining 50% of *residential* growth to the core leads to the fewest roadways needing improvement. The scenario with the most roadways needing improvement was consistently the 'no control' scenario, where no constraints were placed on the land use model, allowing growth to pattern itself more 'naturally.'

Table 10.4: Miles of Roadway Improvements

COV	LU Policy	Selected in Deterministic Analysis (DA)	Selected in Uncertain Analysis (UA)	Selected in UA, but not in DA	Selected in DA, but not in UA
0.1	None	6	5.33	0	0.67
	Basic Outskirts	5	3.83	0	1.17
	50% Emp in Core	4.67	3	0	1.67
	50% Res in Core	2.5	2.5	0	0
	50% Total in Core	5.5	0.5	0	5
0.3	None	6	5	0	1
	Basic Outskirts	5	3.5	0	1.5
	50% Emp in Core	4.67	2.67	0	2
	50% Res in Core	2.5	2.5	0	0
	50% Total in Core	5.5	0.5	0	5
0.5	None	6	0.5	0	5.5
	Basic Outskirts	5	0.5	0	4.5
	50% Emp in Core	4.67	0.5	0	4.17
	50% Res in Core	2.5	0	0	2.5
	50% Total in Core	5.5	0	0	5.5

A gravity-based ITLUM was used to assess the impact of uncertainty on estimates of congested lane-miles in year 25. Household and employment control totals, as well as parameters in the trip generation and distribution equations, were assumed to follow independent lognormal distributions with three levels of uncertainty (embodied in COV assumptions). Sampling was done using the Latin Hypercube (LH) technique, and five growth policies were examined. Two hundred and forty simulations were generated in total (six batches of 80 samples).

Results suggest that greater uncertainty in parameters and other model inputs (i.e., higher COVs) lead to better distribution in traffic flows and thus less congestion (links with v/c >0.85), on average. Moreover, the ordering of congested roadway links is quite similar regardless of the treatment of uncertainty. This is important because it suggests that the roadways ranking the highest in terms of v/c ratios in the deterministic analysis (DA) will likely also be selected for improvement in an uncertainty analysis (UA) based on average v/c ratios. However, the results of the analyses do differ slightly in terms of the growth policy that leads to the fewest congested roadways: the DA results suggested that constraining 50% of residential growth to occur in the core would limit the number of links with v/c ratios above 0.85, while the UA suggested this would occur when constraining 50% of all growth to occur to the region's core.

Future work is needed to determine the impact of correlations amongst uncertain model inputs, specifically amongst the regional control totals so that scenarios representing economic booms and busts, or decisions where growth is expected to happen in either one part of the region or another, can be analyzed. Correlations were neglected in the current work because these are often not known and the independent probability distributions used are a first step towards assessing the role of uncertainty in roadway improvement recommendations. Another venue for future work is to apply the method developed here to an actual region such as Austin, and to use a more detailed model set that captures behavior in a more detailed way.

Chapter 11. Summary and Conclusions

Federal legislation mandates that metropolitan planning organizations (MPOs) consider the likely effects of transportation policy decisions on land use and the environment. Various tools exist to forecast elements of transportation and land use separately; however, little has been done to integrate these co-dependent processes and ascertain which integrated approaches are likely to be most effective and practical for various MPOs and other planning staff.

The U.T. Austin research team interviewed staff at 30 MPOs from around the U.S. regarding land use modeling practices and available data sets. The team examined literature and experiences across a wide variety of predictive land use modeling paradigms, including gravity-based, spatial input-output, microscopic and cellular automata methods. The U.T. Arlington team members examined a more straightforward approach, called suitability analysis (including *What if?*, UPlan, and ArcGIS's Spatial Analyst tool) and then applied several such approaches to the Waco, Texas region.

The U.T. Austin team developed an open-source gravity-based land use model (G-LUM) using MATLAB code, for calibration and application of a system of equations based on Stephen Putman's popular (but proprietary) ITLUP model. The team compared the FWHA's freely available (but closed-source and somewhat inflexible) version, called TELUM, to G-LUM results using Waco data. The team also ran G-LUM on the basis of Austin data under three scenarios, using 5-year time steps (for both the land use and travel demand models) and traffic analysis zone geography.

The U.T. Austin team also coded this land use model and a streamlined travel demand model in C++ in order to appreciate how variations in predictions may impact decision-making under parameter and input uncertainty. Finally, they developed 150-meter grid cell data sets for application of the UrbanSim model, and ran four scenarios at one-year time steps. Both teams then provided workshops to TxDOT and Texas MPO planning staff and consultants in Houston and Austin during the summer of 2008.The team concluded that the presence and complexity of any land use modeling applications are largely a function of MPO size, with the biggest and best-staffed regions typically pursuing some form of land use modeling, often via outside consultants. A great need exists for in-house capabilities, and straightforward tools.

Unfortunately, existing land use models appear to be inadequately documented, and solutions to specification questions are difficult to come by. For example, TELUM's objective function (for parameter calibration) remains something of a mystery, as do elements of its land-consumption functions. The same is true of UrbanSim's land development functions (for construction of new built space). PECAS documentation is still evolving, and complete code is not presently available to the public. For these and other reasons, well documented and highly transparent models (such as G-LUM) are valuable. Nevertheless, it appears difficult to achieve perfectly reasonable performance from zone-based models over time, as zones can depopulate or grow too fast. Many modifications to ITLUP's base equations were made in order to "rein in" certain tendencies evident via the G-LUM applications.

While UrbanSim and some others enjoy the property of strict controls on growth of total population and total jobs (helping ensure these stay in balance), limited move tendencies (helping ensure zones do not depopulate too quickly, for example), and demand for built space never able to exceed supply, the Austin UrbanSim applications demonstrated strong centralizing tendencies (in both populations and jobs) that are inconsistent with past trends and unlikely to emerge in practice, without significant changes in policy and/or behavioral preferences.

Moreover, the data demands and programming skills required by current and past versions of UrbanSim appear excessive for the vast majority of regions, at least in the near term.

For these reasons, most MPOs may be very interested in making use of existing tools like TransCAD's new gravity-based land use model (as presented in the second project workshop, but currently without parameter-calibration capabilities) and land suitability analysis through ArcGIS, for subjectively scoring parcels or grid cells (on the basis of proximity to nuisances, attractions, and sensitive sites, as a function of land use type) and then allocating expected growth using simple rules. Whatever the approach, the opportunities to apply (and enhance) existing data sets are many. The value of these and other modeling tools should be evident in the multitude of information provided staff, decision makers, and public stakeholders across the state and around the world, as they sort through the complex process of refining land use and transport policies for the local, regional, and global challenges.

References

Abraham, J.E., and J.D. Hunt. “Policy Analysis Using the Sacramento MEPLAN Land Use Transportation Interaction Model.” *Transportation Research Record*, 1685 (1998):199-208. Washington, D.C.

Abraham, J.E. and J.D. Hunt, “Random Utility Location/Production/Exchange Choice, The Additive Logit Model, and Spatial Choice Microsimulations.” *Presented at the 86th Annual Meeting of the Transportation Research Board* (2007), Washington, D.C.

Adina M. Merenlender, Colin Brooks, David Shahbazian, Shengyi Gao, and Robert Johnston, “Forecasting exurban development to evaluate the influence of land use policies on wildland and farmland conservation.” *Journal of Conservation Planning*, 2004 1(1) pp. 64-88.

Alan Borning, Paul Waddell, and Ruth Förster, UrbanSim: Using Simulation to Inform Public Deliberation and Decision-Making. Preprint of paper to appear, Digital Government: Advanced Research and Case Studies, Hsinchun Chen et al. (eds.), Springer-Verlag, in press.

Anjomani, A. 1984. "The Overlaying Map Technique: Problems and Suggestions," *Journal of Planning Education and Research*, 4:111-119.

Anjomani, A. 1992. "A Planning and Design Support System for Land-Use and Environmental Planning: GIS, Optimization Models and Spatial Decision Support Systems," a paper presented at the 34th Annual Conference of the Association of collegiate Schools of Planning, Columbus, Ohio, October-November 1992.

Anjomani, A., Shad, N., Saberi, A. 2005. “SDSS for Land Use and Environmental Planning: An Application Example,” in GIS Planet, Proceedings, Spatial Decision Support Systems, S23, 2005.

ArcGIS Spatial Analyst, Retrieved August 31, 2007, from http://www.esri.com/software/arcgis/extensions/spatialanalyst/index.html

ArcView Spatial Analyst, Retrieved July 31, 2007, from http://www.esri.com/software/arcview/extensions/spatialanalyst/about/model.html

Argonés, V., (2006). *Personal Communication*, El Paso Metropolitan Planning Organization.

Armoogum, J., Madre, J.-L. and Y. Bussière. Uncertainty in Long Term Forecasting of Travel Demand from Demographic Modeling. *Proceedings of the 13th Mini EURO Conference*, Bari, Italy, 2002.

Ascher, W. Problems of Forecasting and Technology Assessment. *Technological Forecasting and Social Change*, Vol.13, 1979, pp.149-156.

Ball, R., (2007). *Personal Communication,* Kern Council of Governments. (rball@kerncog.org).

Caliper Corporation. "STEP3: Microsimulation-Based Small Area Demographic Forecasts for Clark County," 2006, Newton, MA.

Canon, A., (2007). *Personal Communication*, Hidalgo County Metropolitan Planning Organization. (956-618-0803)

Center for Urban Simulation and Policy Analysis (CUSPA), 2006.*Opus: The Open Platform for UrbanSim Version 4.0 Reference Manual and Users Guide*. Available at: http://www.urbansim.org/opus/releases/opus-4-1-2/docs/opus-userguide.pdf

Chapin, F., Stuart, J.R., and Kaiser, E. 1979. *Urban Land Use Planning*, Third Edition, University of Illinois Press, Urbana.

Clark, K.C. and L. Gaydos, "A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area." *Environment and Planning B: Planning and Design*, Vol. 24 (1995): 247-261.

Clarke, K.C. and L. Gaydos. "Loose coupling a cellular automaton model and GIS: long-term growth prediction for San Francisco and Washington/Baltimore." *International Journal of Geographical Information Science.* Vol. 12 (1998): 699-714.

Dickey, John W. and T. M. Watts, 1978. *Analytic Techniques in Urban and Regional Planning.* McGraw Hill .

Dijkstra, E.W. A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik*, Vol. 1, 1959, pp. 269-271.

Ducrot, R., C. Le Page, P. Bommel, and M. Kuper. 2004. Articulating land and water dynamics with urbanization: an attempt to model natural recourses management at the urban edge, *Computers, Environment, and Urban Systems*, 28 (1-2): 85-106.

Duthie, J., Unnikrishnan, A., and S.T. Waller. Network Evaluation with Uncertain and Correlated Long-Term Demand. *Proceedings of 85th Transportation Research Board Meeting* Washington, D.C., 2006.

Ellis, P.L., (2007). *Personal Phone Communication,* Texas Transportation Institute. (979-845-9934).

Evilia, C., (2007). *Personal Communication,* Waco Metropolitan Planning Organization. (254-750-5650).

Farhan, M., (2007). *Personal Communication,* Wasatch Front Regional Council. (801- 363-4230).

Gifford, J.L. Adaptability and Flexibility in Urban Transportation Policy and Planning. *Technological Forecasting and Social Change*, Vol.45, 1994, pp.111-117.

Gliebe, J., (2006). *Personal Communication,* Parsons Brinckerhoff.

Goodchild, M.F, 2000. "The current status of GIS and spatial analysis," *Journal of Geographic Systems,* Vol. 2: 5-10.

Gregor, B. "The Land Use Scenario Developer (LUSDR): A Practical Land Use Model Using a Stochastic Microsimulation Framework." *Transportation Research Record*, 07-0438. (2006) Washington, D.C., pp 93-102

Gribbin, D.J. and J.S. Kaleta. Integration of Planning and NEPA Processes. 22 February 2005. Accessed 15 March 2007. <http://www.fhwa.dot.gov/HEP/plannepalegal050222.htm>.

Griffin, G., (2007). *Personal Communication,* Capital Area Metropolitan Planning Organization. (512-974-2084).

Hartshorn, T. A. 1992. Interpreting the City: An Urban Geography (2nd edition) John Wiley and Sons, New York.

Harvey, G. and E. Deakin. Description of the STEP Analysis Package, Berkeley, California, 1995.

Hodges, M., (2007). Personal Phone Communication, Texas Department of Transportation - Austin. (512-486-5122).

Hopkins, L.D. 1977. "Methods for Generating Land Suitability Maps: A Comparative Evaluation," *Journal of the American Institute of Planners*, 43, 4:386-400.

Hunt, J.D. and J.E. Abraham, "Design and Application of the PECAS Land Use Modeling System," presented at the 8th Computers in Urban Planning and Urban Management Conference, Sendai, Japan, May 2003.

Hunt, J.D. and J.E. Abraham. Theoretical Formulation: System Documentation Technical Memorandum 1. October 2007.

Johnston, R. A., and D. R. Shabazian. "UPlan: A Versatile Urban Growth Model for Transportation Planning." *Transportation Research Record*, 03-2542. (2002) Washington, D.C.

Johnston, R. A., Clay, M., and M. C. McCoy. "Assessment of Regional Integrated Transportation/Land Use Models Study." Caltrans Contract 65A0186 (2005). Davis, CA.

Johnston, B., Lehmer, E., Gao, S., and Roth, N. 2006. *UPlan Land Use Allocation Model 2.21 User's Manual*. Information Center for the Environment, University of California, Davis, pp. 1-41.

Kaiser, E. J., David R. Godschalk, and F., Stuart Chapin, Jr. 1995. *Urban Land Use Planning,* Fourth Edition, University of Illinois Press, Urbana.

Kakaraparthi, S. K., and K.M. Kockelman. 2008. "An Application of Urbansim to the Austin, Texas Region: Integrated-Model Forecasts for the year 2030". Under review for presentation at the Transportation Research Board's 88th Annual Meeting. Washington, DC, January 2009.

Khan, A.M. Realistic Planning for Transportation—A Flexible Approach, *Long Range Planning*, Vol.22, 1989, pp.128-136.

Kitten, C., (2007). *Personal Communication*, Capital Metropolitan Planning Organization. (512-974-2748).

Klosterman, R. E., Siebert, L., Hoque, M. A., Kim, J. and A. Parveen. "*What if?* Evaluation of Growth Management Strategies for a Declining Region." *International Journal of Environmental Technology and Management*, Vol. 6 (2006): 79-95.

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

Kruse, D., (2006). *Personal Communication,* Alamo Area Council of Governments.

Lam, W.H.K and M.L. Tam (1998) Risk Analysis of Traffic and Revenue Forecasts for Road Investment Projects. *Journal of Infrastructure Systems,* Vol. 4(1), pp. 19-27.

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

Lemp, J.D.2007 Travel Demand Forecasting Models: Development, Application, and Comparison of Aggregate and Activity Based Approaches for the Austin, Texas Region. Master's Thesis, The University of Texas at Austin.

Lindstone, H.W., and M. Turoff (eds.) 1997. The Delphi: Technique and Application, 2nd ed., Reading, MA: Addison-Wesley Publishing Company.

Lobb, J., (2006). *Personal Communication*, Resource Systems Group, Inc. (Previously Salt Lake City's forecasting manager for the Wasatch Front Regional Council.)

Lund, M., (2007). *Personal Communication*, Brownsville Metropolitan Planning Organization. (956-548-6154).

Mahmassani, H. S. Uncertainty in Transportation Systems Evaluation: Issues and

Approaches. *Transportation Planning and Technology,* Vol. 9, 1984, pp. 1-12.

Martin, W. National Cooperative Highway Research Program Report 365. Travel Estimation Techniques for Urban Planning**.** Washington, D.C.: Transportation Research Board, 1998.

Mathe, P. Hilbert, 2000. Space Analysis of Latin Hypercube Sampling. *Proceedings of the American Mathematical Society,* Vol. 129(5), pp. 1477-1492.

McDaniel, H., (2007). *Personal Communication,* Amarillo Metropolitan Planning Organization. (806-378-4219).

McHarg, I. L. 1969. *Design with Nature*. New York: Natural History Press.

North Carolina Center for Geographic Information and Analysis, *Land Suitability Analysis-User Guide*, December, 2005. Retrieved from http://dcm2.enr.state.nc.us/Planning/user_guide_lsa2005.pdf.

McKay, M.D., Conover, W.J., and R.J. Beckman. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, Vol. 221, 1979, pp. 239-245.

McLarty, J., (2007). *Personal Communication,* Northwest Arkansas Regional Planning Commission. (479-751-7125).

Messen, D., (2006). *Personal Communication,* Houston-Galveston Area Council.

MySQL (2007). MySQL AB. http://www.mysql.com/

NCTCOG. North Central Texas Council of Governments. Main webpage. Accessed 15 February 2007. http://www.nctcog.org/.

Nick, H., (2007). *Personal Communication,* Tyler Metropolitan Planning Organization. (903-531-1174).

Niles, J. S. and D. Nelson. Identifying Uncertainties in Forecasts of Travel Demand.

Preprint for the Transportation Research Board's 80th Annual Meeting, Washington D.C., 2001.

Opus: The Open Platform for Urban Simulation and UrbanSim Version 4, Reference Manual and User's Guide (April 24, 2007). Center for Urban Simulation and Policy Analysis, University of Washington. http://www.urbansim.org/opus/candidate-releases/opus-2007-04-25/html/userguide/

O'Sullivan, A. 2003. *Urban Economics,* Fifth Edition, McGraw-Hill/Irwin, New York, N.Y.

O'Sullivan, A. 2007. *Urban Economics*, Sixth Edition, McGraw-Hill/Irwin, New York, N.Y.

Pearson, D., Bell, C.E., and G.B. Dresser, "TRIPCAL5 User's Manual," sponsored by TxDOT, Research Study Number 2-10-90-1235, November 1990.

Pradhan, A. and K. Kockelman. Uncertainty Propagation in an Integrated Land Use-Transportation Modeling Framework: Output Variation via UrbanSim. *Transportation Research Record*, Vol. 1805, 2002, pp. 128-235.

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

Python Programming Language (2007). Python Software Foundation. http://www.python.org/

Richard E. Klosterman, “The *What if?* Collaborative Support System,” *Environment and Planning, B: Planning and Design*, 26 (1999): 393-408.

Richard E. Klosterman, *What if?* 2.0 User’s guide. *What if?*, Inc. 2007: 1-54.

Rodier, C.J. and R.A. Johnston. Uncertain Socioeconomic Projections used in Travel Demand and Emissions Models: could plausible errors result in air quality nonconformity? *Transportation Research Part A*, 2002, Vol. 36(7), pp. 613-631.

Ruiz-Juri, N. and K. Kockelman, “Evaluation of the Trans-Texas Corridor Proposal: Application and Enhancements of the RUBMRIO Model,” *Journal of Transportation Engineering*, Vol. 132(7), 2006.

Schell, R., (2007). *Personal Communication,* North Central Texas Council of Governments. (817-608-2387).

Sevcikova, H., Raftery, A.E., and P.A. Waddell. Assessing Uncertainty in Urban Simulations using Bayesian Melding. *Transportation Research Part B,* Vol. 41, 2007, pp. 652-669.

Shaw, K., (2007). *Personal Communication*, City of Amarillo. (806-378-3020).

Sheffi, Y. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1985.

Smetana, E., (2007). *Personal Communication,* San Angelo Metropolitan Planning Organization. (325-657-4210).

Smith, S., (2007). *Personal Communication,* Central Texas Council of Governments. (254-933-7075).

Smith, S.K. and M. Shahidullah. An Evaluation of Population Projection Errors for Census Tracts. *Journal of the American Statistical Association,* Vol. 90(429), 1995.

Stein, M.L. Large Sample Properties of Simulations using Latin Hypercube Sampling. *Technometrics,* Vol. 29, 1987, pp. 143-151.

Texas Transportation Institute (TTI), “ATOM2 User Manual,” Prepared for TxDOT, 2001.

Tinajero, R., (2007). *Personal Communication,* El Paso Metropolitan Planning Organization. (915-591-9735).

TRANUS (English version). Modelistica. Accessed 2 February 2007, at http://www.modelistica.com/tranus_english.htm.

User Manual: TELUM (Transportation Economic and Land Use Model) Version 5.0. March 2005. http://www.telus-national.org/telum/TELUMUserManual.pdf

Waddell, P., "UrbanSim: Modeling Urban Development for Land Use, Transportation and Environmental Planning." *Journal of the American Planning Association*, Vol. 68 No. 3 (2002): 297-314.

Waddell, P. and G. F. Ulfarsson. *Handbook in Transport, Volume 5: Transport Geography and Spatial Systems.* (2004) Elsevier Science.

Waddell, P. *Email Communication*, University of Washington (pwaddell@u.washington.edu), March 18, 2007.

Waller, S.T., Schofer, J.L., and A.K. Ziliaskopoulos. Evaluation with Traffic Assignment under Demand Uncertainty. *Transportation Research Record,* Vol. 1771, 2001, pp. 69-75.

Wies, K., (2007). *Personal Communication,* Chicago Metropolitan Agency for Planning. (KWies@cmap.illinois.gov).

Wood, R., (2007). *Personal Communication,* Sherman-Denison Metropolitan Planning Organization. (903-813-3534).

Zhao Fang and Chung Soon. A study of Alternative Land Use Forecasting Models, Lehman Center for Transportation Research, Miami, Florida, April 2006, pp. 1-33.

Zhao, Y. and K. Kockelman. The Propagation of Uncertainty through Travel Demand Models: An Exploratory Analysis. *Annals of Regional Science,* Vol. 36(1), 2002, pp. 145-163.

Zhou, B. K. Kockelman, and J. Lemp. Transportation and Land Use Policy Analysis Using Integrated Transport And Gravity-Based Land Use Models. Under review for presentation at the Transportation Research Board's 88th Annual Meeting. Washington, DC, January 2009.

Appendix A: Land Use Model Descriptions

Eight land-use models reviews are included in this appendix: UrbanSim, TELUM, PECAS, TRANUS, *What If?*, SLEUTH, UPlan, LUSDR, STEP 3, and MEPLAN. This is by no means an exhaustive list of the available land use models, and some models will be discussed in more detail than others depending on the research team's experience with each model to date. Nevertheless, these are felt to be the key competitors for TxDOT's future land use modeling efforts across Texas's various regions.

UrbanSim

UrbanSim, open-source software developed at the University of Washington by a team led by Dr. Paul Waddell, is perhaps the most comprehensive land use modeling package available. Eight models are used within the UrbanSim package to predict the household, employment, and land characteristics for each 150 square meter grid cell covering a region. While its comprehensiveness is attractive to many land use modelers, others are deterred by the numerous data requirements.

UrbanSim uses Census data, specifically household data from the Public Use Microdata Sample and the Summary Tape File 3A, to synthesize a base year household database with characteristics including household size, income, number of vehicles, and number of workers. Base year employment must be geocoded and assigned to grid cells with characteristics such as employment sector and number of employees. A base year parcel-level map of land cover with data on lot price, building price, land use code, and zoning must be available in GIS.

The models used in UrbanSim can be categorized into accounting models, probabilistic choice models, and regression models. Accounting models include the household and employment transition model. The household transition model simulates births and deaths using factors such as household income, age, household size, and the presence or absence of children. Iterative proportional fitting is used to ensure that the population equals a pre-specified control total. Created households are not assigned to a specific location until the implementation of the household location model. The employment transition model uses a similar methodology to model job creation and loss (Waddell, 2002).

Probabilistic models can be further broken down into rate-based models for use in relocation and logit models for use in location. The household relocation model uses historical data to simulate whether or not a household decides to move. Each household that has made the decision to move is then assigned using the household location model and the status of its previous location is updated to vacant. The multinomial logit household location model is a function of the characteristics of the location (i.e., housing price, density, and age), neighborhood characteristics (i.e., land use mix, density, average property values, and accessibility to stores), and accessibility to jobs. The employment relocation and location models use a similar methodology to simulate the movement of jobs (Waddell, 2002).

The land price model is a linear regression model for determining the price of each grid cell over time. It is based on urban economic theory, which states that the more valuable a piece of land is, the more expensive it will be to purchase. The value of the land is determined by neighborhood characteristics, accessibility measures, and policies.

One advantage of UrbanSim is that it allows users to create and test different scenarios based on alternative policies by changing modeling constraints. The outputs of UrbanSim can be summarized at any level of aggregation including grid cells, TAZ, or the entire region. While several metropolitan areas are experimenting with UrbanSim, it has yet to be fully implemented anywhere for forecasting purposes.

TELUM ("Transportation Economic and Land Use Model")

Dr. Stephen H. Putman, developer of TELUM, was among the first to identify the interrelationship between transportation policy and land use. He did this using the Integrated Transportation and Land Use Package (ITLUP), which he developed in the 1970s. Over 20 U.S. metropolitan areas have applied all or a portion of ITLUP. Since the initial work, ITLUP has been extensively revised and updated, aided in part by advancements in the computer technology sector (i.e., development of GIS) in the past two decades. In 1997, METROPILUS, a new land use model was developed. The entire modeling system along with a Graphical User Interface was embedded in the ESRI's ArcView GIS operating environment. It contained several models for location analysis, and could perform data analysis, statistical analysis and display mapping of the output. Each version of TELUM is available in a free format, and may also be available in a format with more functionalities for a fee.

In 1999, Professor Putman began retooling METROPILUS as a land use component of Transportation Economic Land Use System (TELUS). The TELUS Land Use Model or TELUM can predict the location and growth of residential and nonresidential development for up to 30 years based on the current and lag year residential and nonresidential development, the locations of transportation improvements, and data on congestion in the transportation system

TELUM consists of sub-models for residential development (TELUM-RES), employment (TELUM-EMP), and land consumption (LANCON). Land use forecasting is done in time increments of 5 years. Each time increment begins with the execution of TELUM_EMP to forecast the locations of employment in four to eight sectors based on previous year data including the number of each type of employment in each zone, zonal characteristics (i.e., population, area), and zone-to-zone travel time.

Once the regional employment forecast is complete, TELUM-RES forecasts all residential locations. TELUM-RES requires the classification of households into four or five groups based on income. Residences are located based on employment locations and previous year data including number of each type of residence in each zone, percent of each zone's area devoted to residential use, percent of each zone's area that is vacant and developable, percent of each zone's developable area that is developed, and zone-to-zone travel cost.

LANCON computes the land consumption in each zone based on the demand for land for employment and residential purposes, and the supply of land in each zone (User Manual: TELUM, 2005).

PECAS ("Production, Exchange and Consumption Allocation System")

PECAS, a freely available JAVA-based model developed at the University of Calgary, consists of two submodels, the first being the spatial input-output model. The spatial input-output model uses nested logit models for three different decision types. The first decision type, which is at the top of the nest, is the location of activities. The second decision is how much a certain commodity should be made or used at a given location of activity. The third decision, at the bottom of the nest, is where to buy or sell the commodities, including the decision to travel or have the goods shipped (Johnston, 2005).

The second submodel PECAS utilizes is the land development model, which is available in aggregate (zone based) or disaggregate (grid cell based) models. The land development model utilizes calculations from the spatial input-output model to calculate profitability for developers by development type for each unit of land. The choice to develop is modeled as a logit model (Johnston, 2005).

PECAS interacts the models through discrete time-steps, recommended to be one year, and through land use zones, recommended to be set at a maximum of 750 zones. These recommendations allow for quick response between land development and land prices, while minimizing the burden on the operating system running the model (Johnston, 2005).

The developers of PECAS expect to come out with PECASim, a microsimulation based model, in the near future (Abraham and Hunt, 2007).

TRANUS

TRANUS is a freely available Windows-based integrated land use/transportation modeling system that draws upon modeling theories ranging from spatial economics, gravity models, input-output economics, random utility, and network optimization. It consists of four basic phases, iterated over a number of time periods.

The first phase is an input-output econometric model to distribute productions and demands. It starts with a user-supplied assumption about the growth of basic sectors. In turn, the model creates additional demands and productions related to the needs of those basic sectors. The model also contains provisions for supply scarcity (especially of land) which would increase the price of that sector, and thus limit its usage by low-performing sectors, which in turn would limit the growth of that sector. However, the model also contains the notion of substitution, whereby a sector can substitute less desirable, but less expensive resources. After determining the amount of total basic sector growth, the model probabilistically places that growth in zones where that growth would be most attractive to that sector.

In the second phase of the model, the total production for each zone/sector combination is established. This includes the generation of induced demand and the calculation of production costs (based on costs of inputs plus the transportation costs of those inputs and based on the transportation network costs of the previous time period). Production locations are then established through an iterative process where costs and demand are adjusted until equilibrium is achieved.

Next, the activity-transportation interface is called. Based on locations of productions and demands, items (people or freight) are moved from one location to another. The spatial distribution from a consumption zone to the production zones is made with a discrete choice model. The demand function in the model includes the consumption cost as well as transport disutility between zones.

Finally, the transportation model is run. Multiple shortest paths are found to match a particular supply to a particular demand, and the supply is divided probabilistically into each path. However, because use of a link may cause congestion on the link, which would reduce travel time, the transportation model is repeatedly run with adjusted travel times until equilibrium is established (Tranus, 2007).

What if?

What if? is a rule-based model that is compatible with any GIS software. It determines land suitability for development, and creates projections of demand for land consumption.

Applications are limited to fifteen types of urban land use, fifteen types of non-urban land use, three residential uses, and four time periods. The spatial resolution is user defined.

Land suitability is determined by first allowing the user to rank the importance of each suitability factor, and then specify the relative suitability of each factor type for each land use and the land uses that are candidates for conversion from their current use to another use.

Four types of land demand are projected: residential, regional, preservation, and local. The assumptions used to project demand differ between types. To project demand for residential land, users must specify future year projections for total number of households, group quarters, and average household size. Also, the future percentage breakdown by housing type, housing density, vacancy rate, infill percentage, and the loss rate for each type of residential housing must be specified. To project regional demand, users must specify projected employment, density, and infill rate. To project the future locations of preserved land, the user must specify the total amount of land that should be preserved for various purposes in the future year. To project local demand, the user must specify the quantity of land that will be required per thousand new residents in each political jurisdiction.

The user has control over the following aspects of allocation: the order in which projected land use demands are allocated, influence of infrastructure plans (i.e., utility extension) on allocation, land use plans and regulations, and parameters related to the growth pattern. Important outputs include the land use impacts on open space, nutrient loading, water quality, and air quality.

What if? does not rely on behavioral theory, and so its success is largely dependent on the user's ability to specify reasonable assumptions. The interactions between markets for land, housing, non-residential uses, labor, and infrastructure are not modeled (Klosterman et al, 2006).

SLEUTH ("Slope, Land-use, Exclusion, Urban extent, Transportation, and Hillshade")

SLEUTH is a freely available cellular automation model developed at the University of California, Santa Barbara. Required input files include raster image maps of initial conditions and at least three historical conditions, a map of areas unlikely to be urbanized, at least two maps of some historical transportation system configurations, and a digital elevation model (to create a grid of slopes).

In addition to these, a set of numerical inputs is needed for each of five factors or coefficients, which control the behavior of the system. A diffusion factor determines the dispersive nature of the land use distribution. A breed coefficient specifies how likely a newly generated detached settlement is to begin its own growth cycle. A spread coefficient controls how much diffusive expansion occurs from existing settlements. A slope resistance factor influences the likelihood of settlement extending up steeper slopes. A road-gravity factor attracts new settlements toward and along roadways. Related to these factors and coefficients are four different types of modeled urban growth: spontaneous growth, diffusive growth, organic growth, and road-influenced growth (Clarke and Gaydos, 1998).

The modeling procedure could be simply described as follows. First, potential cells are randomly selected for urbanization, and then the uniform growth rules are applied to evaluate the cell and its neighbors on a cell-by-cell basis, and then the weighted probabilities decide whether the potential cells are urbanized. Second, changes are updated at the end of this time span, and the modified urbanization configuration serves as the basis for the next year; whenever a historical configuration is available, the predicted result is compared with the historical conditions to calibrate the model. Finally, the modified parameters are used for prediction, which is generally the ultimate goal of our urbanization modeling (Clarke and Gaydos, 1995).

One unique aspect of SLEUTH is its ability to modify its parameters during a model run if, for example, a growth rate exceeds a pre-specified “critical high.” Another unique aspect of the model is its capability to visualize uncertainty, using multiple Monte Carlo runs.

UPlan

UPlan is a freely available rule-based model which is used in conjunction with ESRI’s ArcGIS, requiring the use of its Spatial Analyst Extension. Rule-based models use a set of rules along with GIS to distribute population, land distribution and employment (Waddell and Ulfarsson, 2004).

The model distributes land distribution to a set of predefined, discrete land use categories. UPlan is also flexible in the fact that the user can program different land types (Johnston and Shabazian, 2002). Johnston and Shabazian (2002) explain that categories may be matched to those of the MPO and a residential category may be defined for very-large-lot residences.

In order to calculate the land distribution, a user must input demographic and land use factors. These data are converted to hectares of land consumption for each land use. The user must input specific data to calculate land use consumption for each category. For example, if one wants to investigate future housing consumption, he or she must input the household size, percent of households in each density class and average parcel size (Johnston, 2002).

UPlan assumes that the closer a vacant lot is to an attraction, the more likely it will be bought and developed. UPlan also assumes that the attractive lots are developed because of their location relative to urban areas and transportation facilities. Lands that cannot contain development may also be identified and excluded in the modeling process (Johnston and Shabazian, 2002).

UPlan may be well-suited for organizations starting up a land use planning section, because its rule-based nature facilitates the accessibility of the model, and no calibration is required. One major consequence of having a model based on rules is that one cannot interpret the results from the model as being behavioral (Waddell, 2004). Also, UPlan does not include an economic model to simulate travel costs, land costs, or the general state of the economy in a region (Johnston et al., 2005).

LUSDR (“Land Use Scenario Developer”)

LUSDR, a model developed using the R programming language at Oregon Department of Transportation, is designed to create a balance between model completeness and model comprehensiveness. Model completeness relates to the fact that the number of dynamic microsimulations be run many times with different policy inputs. Numerous model runs minimizes the uncertainty due to the dynamic nature of the model, and allows the user to test a range of policy inputs (Gregor, 2006). According to Gregor (2006), a comprehensive model incorporates a significant amount of behavior in the model, including modeling decisions as a result of market interactions and modeling activity-based travel demand. Model completeness requires run times to be reasonably short, while a comprehensive model may take days or weeks to run (Gregor, 2006).

LUSDR balances the conflicts of model completeness and comprehensiveness by performing stochastic microsimulations using samples. Virtually all modeling processes in LUSDR are Monte Carlo processes (Gregor, 2006). The outcomes are found by sampling from distributions which come from the Census Public Use Microsample data, "terminal node probabilities from decision trees, probabilities derived from inventory size distributions, logit model probabilities, and expert judgment of land use compatibilities used as probabilities" (Gregor 2006, p. 8).

Employment establishments and households are generated using joint probabilities derived from household and employment data entered by the user. Developments are allocated to zones based on land supply, constraints, and "preferences." Also, the model considers supply and demand when allocating these developments. These steps repeat until all developments are allocated to a zone (Gregor, 2006). LUSDR source files will be freely available in the future.

STEP 3

STEP 3 is a microsimulation-based tool created by Caliper Corporation (2006) for small area demographic forecasts. Unlike most land use models, STEP 3 does not rely on control totals. Instead, population is progressed each year via migration and aging. Uniform grid cells are used instead of political boundaries such as parcels because it is easier to detect trends over time. Redevelopment is not accounted for. The key components of STEP 3 are the labor force participation model and the location choice models.

The labor force participation model first determines if each person is a worker based on gender, household structure, age, and race. For persons over 65 years in age, a retiree model is run to determine if a person retires based on gender, household structure, and age.

The location choice models run in four steps. First, a work location is determined for each head of household employed in key sectors, which necessitate the choice of workplace before choice of residence. In the Clark County, Nevada application, the key sector was the hotel industry, and the work location was modeled as function of accessibility to city business district (CBD), accessibility to the "strip," and the natural logarithm of available hotel jobs. Second, a residential location is determined for each household considered in the first step. This choice depends on the pre-determined work zone and is a function of travel time, ratio of housing cost to income, residential density, manufacturing job density, and the natural logarithm of housing units. Third, the residential location is determined for all remaining households as a function of average travel time. Lastly, a work location is determined for all remaining workers, dependent on residential zone and a function of travel level of service, access to CBD and "strip," and the natural logarithm of available jobs.

Several assumptions are made about household structure. Young adults are assumed to leave home at age twenty-two and form a new household. The presence of an auto in the new household is based on the ratio of autos to people in the original household. A marriage rate by age is assigned to males, and prospective female partners are chosen based on age and race. Divorce rates are based on age and gender of the head of household. If a divorce occurs, the income and autos are split. Children are randomly assigned to each parent using gender defined custody probabilities.

Rules are in place to determine whether or not a cell is suitable for growth. A cell is considered for residential growth only if it has two neighboring cells with population above a given threshold. If growth does occur, it is distributed according to a pre-specified owner-renter ratio. A cell is considered for retail employment growth if it has a high population and small amount of retail. Retail deficits are found by analyzing the gradient of retail by population density surface. The number of retail jobs required is based on the height of the "deficit peak" in relation to overall population density (Caliper Corporation, 2006).

MEPLAN

MEPLAN is a market-based urban economic model (Johnston, 2005). The model focuses on two markets that exist, the market for land and the market concerning transportation. Within the land market, prices and generalized cost signals drive production, consumption and activity locations. Within the transportation market, money costs and congestion delay drive mode and route selections (Abraham and Hunt, 1998).

The land market uses logit models to allocate different amounts of activities to various zones in the project area. The zone attractiveness used in the logit models is based on the various costs of the activities, location specific costs, and the cost of transportation. These interactions lead to the development of origin-destination tables of different types of trips throughout the region. The origin-destination tables are then loaded onto a network which includes nested logit models for mode choice and stochastic user equilibrium for the assignment of traffic. The network derives impedances which, in turn, impacts the attractiveness of zones and locations of activities (Abraham and Hunt, 1998).

This model is reiterated through a time step with the land market interaction being performed before the transportation market interaction. The impedances from the transportation market are fed into the land market for the next time step. This introduces a lag in the land market in response to the transportation network (Abraham and Hunt, 1998).

Appendix B: Summary of MPO Modeling Staff Responses to Interview Questions

Texas MPOs

Amarillo's MPO is currently working closely with localities to develop a new travel model and land use modeling procedure and to enhance GIS capabilities (McDaniel, 2007). Kelley Shaw, planning director for the City of Amarillo, expressed disappointment with current trend-based methods for predicting population and employment. He hopes to move the region towards a fast running land use model with visualization capabilities such as Klosterman's *What if?* to model alternatives for growth in this fast developing region (Shaw, 2007).

The Brownsville MPO's Technical Committee uses its knowledge of the region and data on subdivision activity to modify the state data center's county forecasts of population and employment. Then, the 3-person staff allocates the county numbers to TAZs and sends the information to TxDOT to run the TDM. While MPO coordinator Mark Lund does not feel that an advanced land use model would be of use to his region due to its inability to accurately capture constraints imposed by the environment and by the will of the public, he is initiating a suitability process to engage politicians in the public in a discussion of possible growth scenarios. The city of Brownsville has no zoning restrictions, and the GIS data currently available does not contain information on land cover (Lund, 2007).

While the Capital Area Metropolitan Planning Organization (CAMPO) is looking seriously at implementing an advanced land use model, currently population and employment are allocated from state data center county control totals to traffic analysis zones (TAZs) with a Delphi approach. According to planner Greg Griffin, CAMPO is looking most seriously at UrbanSim, especially its integration into ArcGIS. Griffin likes using the grid cell unit of analysis, which will hold steady over time, as opposed to using political boundaries such as parcels or TAZs, which will change. Other issues he sees with the data are that it changes dynamically and is often outdated by the time it is used in a model, citizens may have privacy concerns over the use of parcel level data, and data in each county is updated on different time scales. To lessen the impact of these data issues on modeling results, UrbanSim will likely be implemented with 5-acre grid cells. Also, a tool such as CommunityViz may be considered for use in public meetings (Griffin, 2007).

El Paso MPO's Roberto Tinajero and New Mexico State University's Quinn Korbulic are working on implementing **UrbanSim** for the region. Quinn is gathering the necessary data, and then Roberto will be responsible for calibration and operation of the model. A 10 kilometer grid cell resolution will be used, and the results will be compared to the results of using resolutions of 150 meters and the TAZ level. At this time, El Paso MPO does not have a land use model (Tinajero, 2007).

Hidalgo County MPO director Andrew Cannon is pleased with his organization's procedure for forecasting land use with aerial photographs, ArcMap, and TransCAD. Year 2000 Census data and a **rooftop count** of 2004 aerial photos will be used to find a linear trend. Andrew estimates that the Hidalgo County region houses between 20,000 and 30,000 undocumented migrants, who substantially impact the transportation system performance, but are not represented in the Census. Preparation of the 2004 population and employment counts took approximately four months. The MPO has a close relationship with the county, which is providing parcel data in GIS for 2005 and 2006 (Cannon, 2007).

Houston-Galveston Area Council is the only MPO in Texas to have both calibrated and implemented UrbanSim to date, and it is in fact the first MPO in the U.S. to rely on UrbanSim for long-range planning. Though it has not been able to integrate its land use and travel demand forecasting work in past UrbanSim applications, it hopes to do that soon. Such integration and internal consistency is the goal following the agency's transition to CUBE Voyager software later this year, when the new software will be made available to both the land use and travel demand modeling groups. (Presently, the agency's EMME2 software is used only by its travel modeling group.) UrbanSim's model calibration has been contracted out to a consultant. Each of the region's eight counties was modeled separately, as results at the regional level were felt by staff to be unreasonable (Messen, 2006).

Steve Smith of the Central Texas COG covering the Killeen-Temple region suggested that advanced models are not needed for small urban areas. The Central Texas COG employs just four-part time staff, each of whom works part-time in another city or regional agency. This system of employment has the benefit of facilitating inter-agency communication. Parcel data is maintained by the cities and employment data is gathered from the Texas Workforce Commission. Permits and subdivision plats were used to take the 2000 Census data to the 2006 model base year. City land use plans and the Delphi method were used to forecast population and employment in each TAZ (Smith, 2007).

The San Angelo MPO is currently updating its travel model and gathering GIS parcel data as well as data from the Texas Workforce Commission with the help of a consultant. With a staff of only 2.5 employees, they will not consider their land use model options until after the travel model development is complete at the end of summer (Smetana, 2007).

San Antonio's Alamo Area Council of Governments has developed a 2000 base-case data set and is working on their 2005 employment data (Kruse 2006). Satellite data provide land slope estimates across the region's 150-meter grid cells, while permitting data are mined to provide construction and demolition information from the 1980s and 1990s. Appraisal District data offer information on parcel dimensions and locations, as well as a wide sample of sales prices for a variety of property types. These are the standard approaches to acquiring the data critical to UrbanSim's module calibrations.

Sherman-Denison is the smallest MPO in Texas with a population of approximately 56,000. Transportation Director Robert Wood is happy with the region's method of predicting population and employment—in-house Delphi to allocate SDC control totals to TAZs using Census data and septic tank permits, however he wishes he and the one other staff member had more time to devote to the evaluating the data. The MPO is including in their budget a demographic study through TxDOT's TP&P department. GIS parcel data is available as well as data from 911. No zoning controls are in place (Wood, 2007).

The Tyler MPO uses a consultant to lead a Delphi process informed by data from the Census, Texas Workforce Commission, Claritas, and development permits to get socioeconomic forecasts. According to Planner Heather Nick, Tyler MPO is happy with their process, but is not very knowledgeable about more advanced models that are available. The MPO has GIS data at the parcel level, and 2003 aerials that will be updated this spring (Nick, 2007).

The Waco MPO is currently working with a consultant on scenario planning. While this visioning process will give the region a relative idea as to how the transportation plan will affect land use patterns, the population and employment forecasts input into the travel model are obtained through a modified Delphi process. Seven years ago a consultant tried to use a land use model and the MPO was not happy with the results (Evilia, 2007).

The North Central Texas Council of Governments (NCTCOG) currently uses the Dram/Empal land use model to forecast population and employment to districts. The district level information is then allocated to TAZs using a Delphi process. NCTCOG is interested in using a more advanced land use tool in the future such as UrbanSim or PECAS. GIS parcel data is available for 2005. While the quality of data varies widely between counties, it is of high quality in the "core counties" of Dallas, Tarrant, Collin, and Denton (Schell, 2007).

MPOs Outside Texas

Chicago Metropolitan Agency for Planning (CMAP), the newly formed entity combining the previously separate land-use and transportation planning agencies, is working from a full application of Dram/Empal created in the 1990s. CMAP is currently evaluating UrbanSim, PECAS, and LEAM. The combination of LEAM—"Land Evolution Assessment Model"—and CMAP's travel model is referred to as CLUTE or "Chicago Land Use Transportation Evolution." According to Kermit Wies, CMAP is still working on data exchange protocols and is far from testing the model. Wies expressed dissatisfaction with Dram/Empal, saying that it is "good as <an> instructional tool for indicating various first-order relationships," however, "overall <it> is not sufficiently advanced to satisfactorily explain the policy interactions necessary to inform current planning debates." He suggested that regions new to land use modeling should develop their model incrementally "to educate and instruct policy makers on simple intuitive points" (Wies 2006).

Oahu MPO was to be one of the first to implement UrbanSim, but has yet to rely on it for the region's land use forecasts (Gliebe 2006). This is because the MPO staff is not yet satisfied with the results at the zonal level, relative to their land inventory approach. Their current (and past) approach relies on rather straightforward land allocation rules, developed by modeler Steve Young, who is very familiar with the region's past development and present patterns (Gliebe 2006). Oahu's experience is evidence that it may be possible for expertise and logic to perform more accurately than the best of the large-scale, microsimulation-based land use models. .

Kern Council of Governments, the MPO for the Bakersfield California region, recently considered several suitability and land-use models (*What if?*, PLACE3, and PECAS) before choosing UPlan. UPlan was chosen because two of the eight MPOs in the region have had experience with it and the software and support is free. Although there are concerns that UPlan's raster-based processing may be less efficient than the polygon-based processing of WhatIf?, one MPO in the area had a bad experience with a previous implementation of WhatIf? (Ball, 2007).

Northwest Arkansas Regional Planning Commission (NWARPC) expressed a desire to integrate travel demand and land-use modeling in the future. Currently, NWARPC uses data from the Census and the Arkansas Department of Workforce Services to get population and employment projections (McLarty, 2007).

Utah's Wasatch Front Regional Council (WFRC) was one of the first to test UrbanSim, working with it since the late 1990s. John Lobb, who led that modeling effort, indicated that he feels that Salt Lake City is now at the stage Houston was 2 years ago (Lobb 2006). Approximately 3 years was required to collect the necessary data. One of the key data challenges is that data sources used rarely provide data for a consistent point in time. Some of these sources used are the Census, Work Force services, and GOPP (Government of Public buy back from city). The geographic unit used in UrbanSim is a 25 square-acre grid cell. The grid cell data is converted to TAZs (1500) while modeling (Farhan, 2007). Until the UrbanSim model is ready, WFRC is continuing to project population and employment at the TAZ level using the Modified Stratified Iterative Disaggregation (MSID) process. MSID is based on the theory that density decreases and growth rates increase as distance from the CBD increases.

Appendix C: The Legal Foundation for Integrated Land Use-Transport Modeling

At one time, an improperly created long range plan could be the subject of a lawsuit. See *Sierra Club v. U.S. Dept. of Transportation*, 962 F.Supp. 1037 (N.D. Ill., 1997). In this case, which involved the building of a toll road in the Chicago area, the U.S. District Court ruled the U.S. Department of Transportation and the State of Illinois had failed to comply with NEPA and MIS procedures when it used a socioeconomic forecast that only one alternative could satisfy. However, as part of an effort to streamline the transportation planning process, Congress barred judicial review of much of the planning process by passing The Safe, Accountable, Flexible, and Efficient Transportation Equity Act of 2003 (SAFETEA). For instance, 23 U.S.C. § 134(o) states the following:

> Since plans and programs described in this section are subject to a reasonable opportunity for public comment, since individual projects included in the plans and programs are subject to review under the National Environmental Policy Act of 1969 (42 U.S.C. 4321 et seq.), and since decisions by the Secretary concerning plans and programs described in this section have not been reviewed under such Act as of January 1, 1997, any decision by the Secretary concerning a plan or program described in this section shall not be considered to be a Federal action subject to review under the National Environmental Policy Act of 1969.

And, US code states the following, immediately after the seven factors mentioned in 23 U.S.C. § 134 (f) (1):

> Failure to consider factors.--The failure to consider any factor specified in paragraph (1) shall not be reviewable by any court under this title, subchapter II of chapter 5 of title 5, or chapter 7 of title 5 in any matter affecting a transportation plan, a transportation improvement plan, a project or strategy, or the certification of a planning process. (23 U.S.C. § 134 (f) (2))

Thus, the courts have been removed from considering whether a long range plan or transportation improvement plan is in compliance with the seven factors listed in 23 U.S.C. 134 (f), or in their accompanying regulations. Courts also cannot judge whether a plan meets the requirements of NEPA. Notwithstanding the above, this does not mean that such factors or NEPA can be ignored. The Secretary of Transportation still has the authority and the requirement to certify any long range plan or transportation improvement plan, and failure to have certified plans can subject a transportation management area to a 20% withholding of federal funds until a certified plan is in place. 23 U.S.C. § 134 (i)(5)(C)(i).

Furthermore, in a memorandum entitled “Integration of Planning and NEPA Processes” (Gribbin and Kaleta, 2005), the chief counsel of FHWA indicates that “current law provides authority for and even encourages the integration of the information and products developed in highway and transit planning process into the NEPA process.” As projects listed in the transportation plans are still subject to NEPA, “because of the continuity between the planning and project development processes, the NEPA analysis for a transportation project needs to be reviewed in the context of the transportation planning process.” Furthermore, information gathered during the planning process can be used as a basis for NEPA documentation, including the purpose and need statement, the consideration and even elimination of reasonable alternatives, and the description of environmental consequences. As the nature of planning is to look broadly at future land use, development, population increases, and other growth factors, it can provide the basis for the assessment of cumulative and indirect impacts required under NEPA. As for the level of review,

> "FHWA and FTA should independently review regional analyses or studies of transportation needs conducted during the transportation planning process at a similar level. FHWA and FTA reviewers do not need to review whether assumptions or analytical methods used in the studies are the best available, but, instead, need to assure that such assumptions or analytical methods are reasonable and scientifically acceptable. This review would include determining whether assumptions have a rational basis and are up-to-date and data, analytical methods, and modeling techniques are reliable, defensible, and reasonably current."

Thus the incorporation of detailed planning metrics such as land use impacts, despite being seen as an aspirational goal by Congress, is still seen as practically obligatory by the Department of Transportation.

Appendix D1: Suitability Analysis

The following discussion and images offer more details on the process of suitability analysis and key software packages.

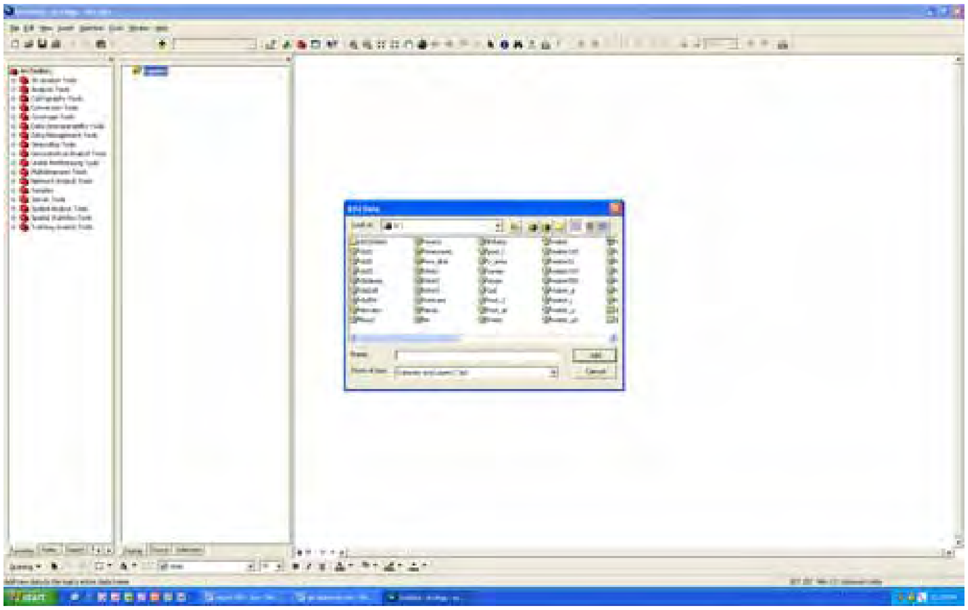
Steps of suitability analysis

Suitability analysis involves the following steps (Kaiser et al., 1995):

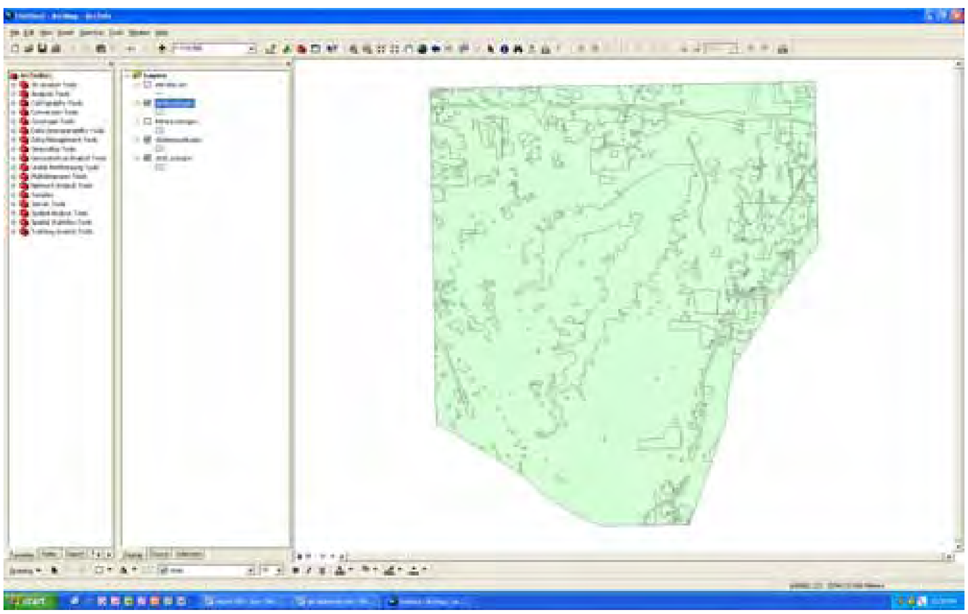
1) Identify land use to be analyzed.
2) Determine the suitability factors to be considered for the particular use (e.g., slope, soil, distance to expressway and airports, access to waterline, land cover).
3) Specify factor ratings based on the internal characteristics of each suitability factor depending upon their contribution (for example slopes of 1 to 6 percent are given a higher rank, say 2, than steeper slopes of more than 6 percent, which are ranked lower, with a 1).
4) Determine factor weights for each suitability factor in terms of its relative importance for suitability for use under study (e.g., because distance to expressway and waterline is deemed twice and three times as important for industrial location as slope, they are weighted as 2 and 3 respectively whereas slope is weighted 1).
5) Multiply each factor rate by the factor weight. Combine the weighted factors into a single suitability score (addition, multiplication…).
6) Add the weighted factors to generate suitability score for each site in the case.
7) Reclassify the resulting range of numerical scores into a simplified composite suitability score.
8) Transform the outcome into a suitability map by choosing a set of patterns to represent the different degrees of suitability.
9) Generate a statistical report showing for each suitability class, the site identification, number of acres, and other relevant data.

The process of land suitability analysis in ArcMap using ArcGIS

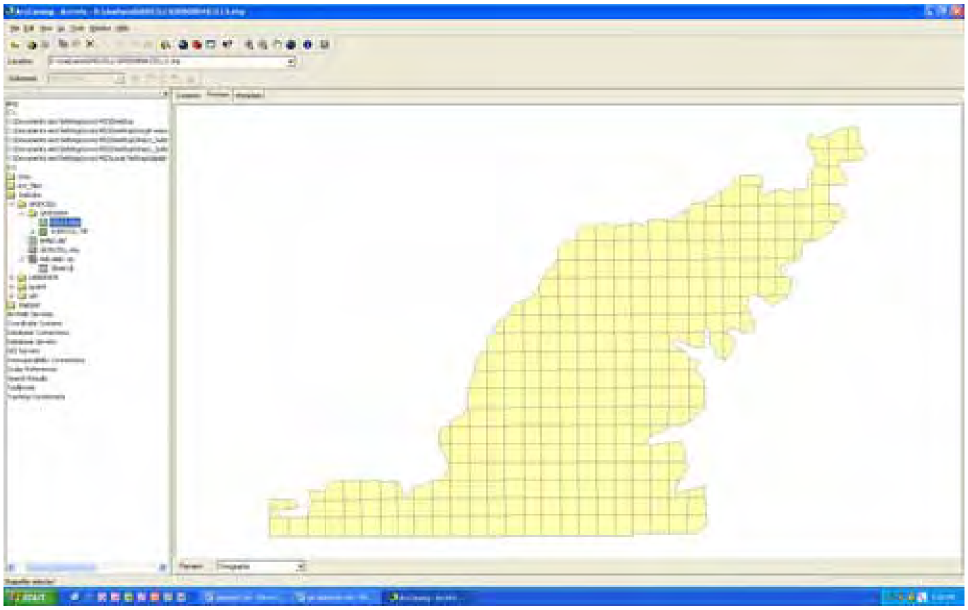
1) As an example, figures below show JoePool Lake land use suitability study.



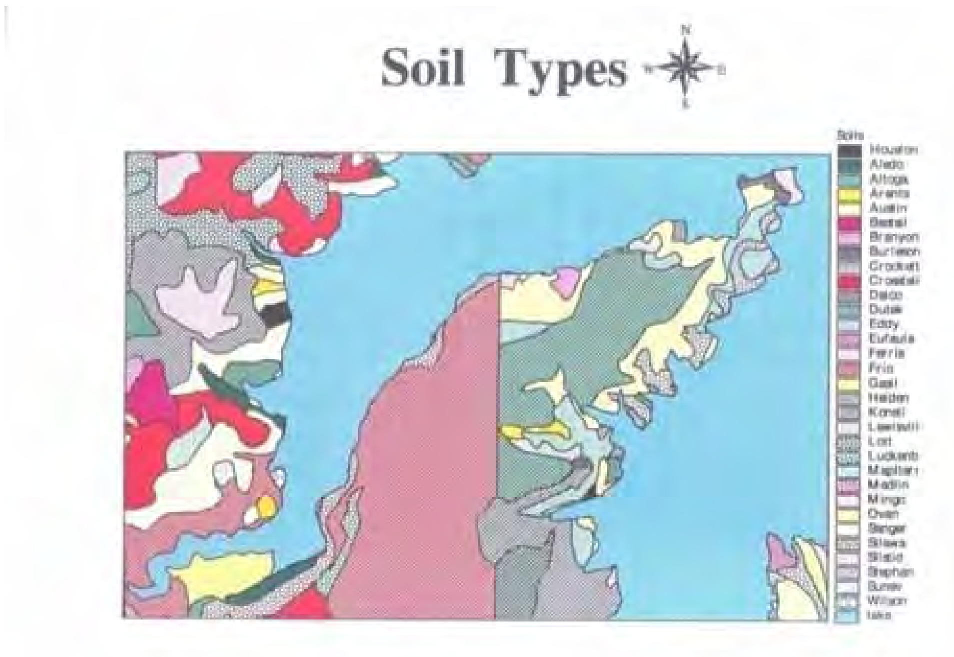
2) Identify the land uses to be studied and define the boundary for the study area.



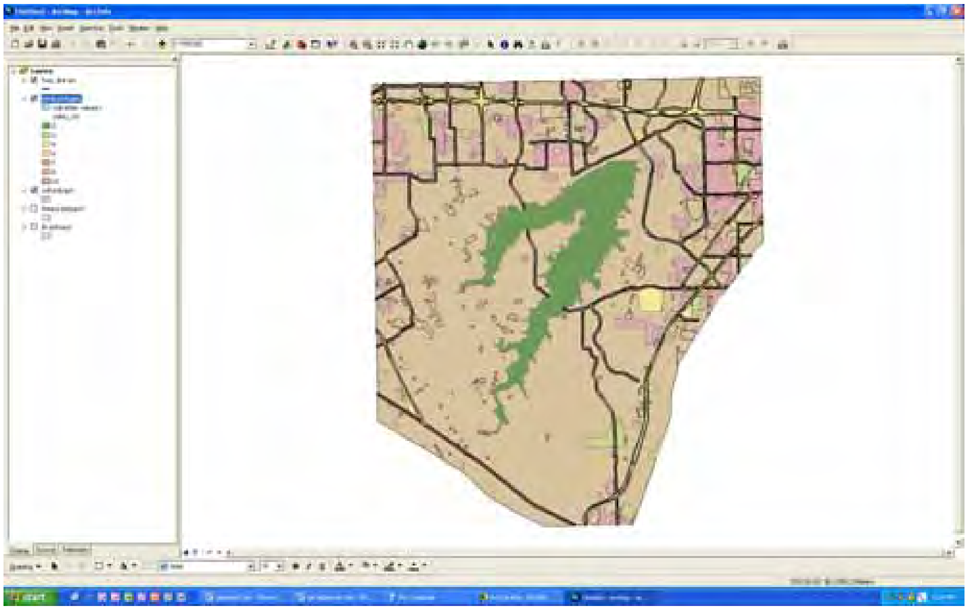
3) Select a grid size (e.g., 150X150 meter) and prepare a grid/fishnet for the study area identified.



4) Determine the suitability factors to be considered for the analysis (e.g., soil, slope, vegetation, distance to expressways and airports etc.).

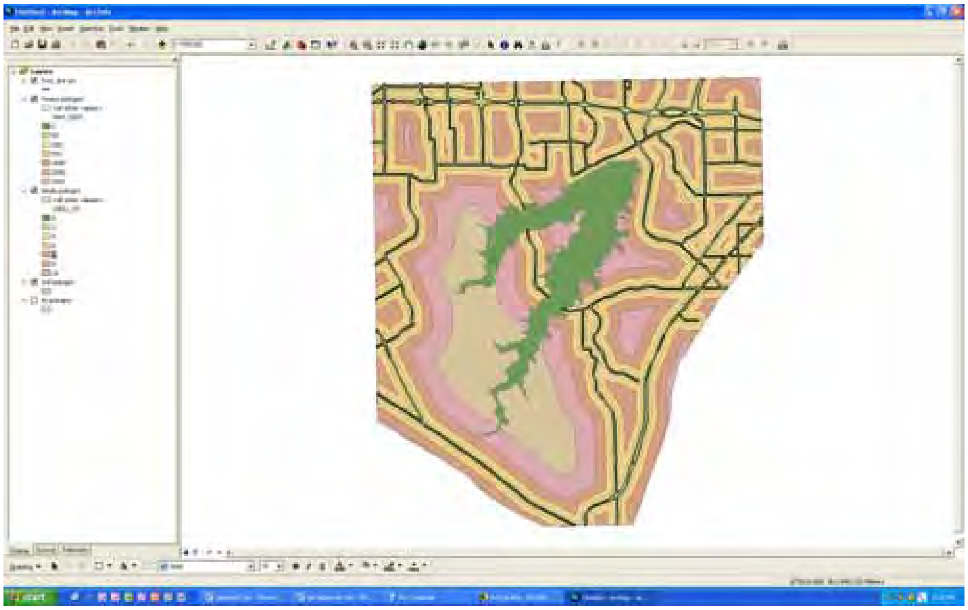


5) The following figure shows existing land uses and undeveloped areas in which the identified land uses need to be allocated. For example, open space, single-family, multi-family, low-density commercial, and high-density commercial will need to develop.



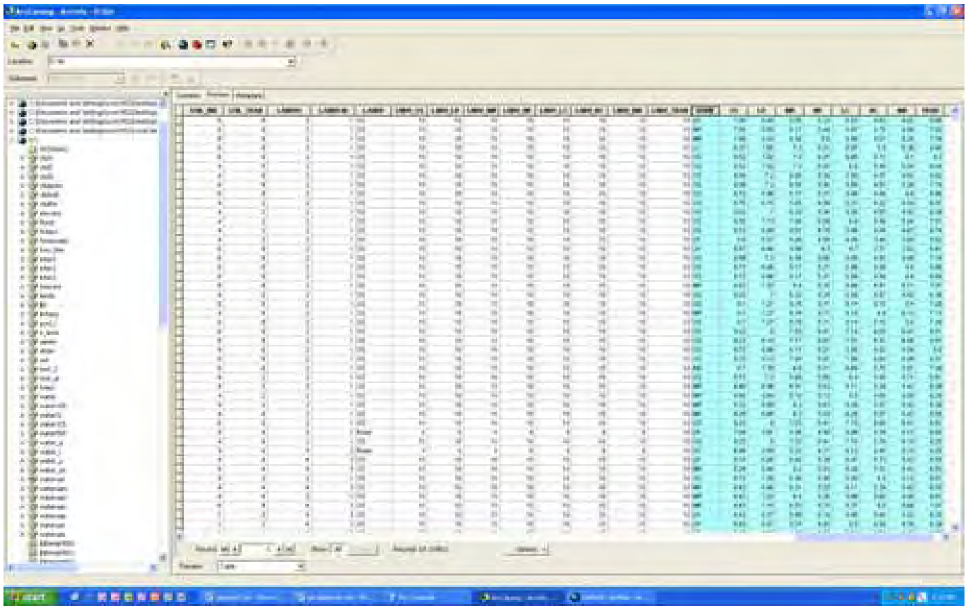
6) Specify factor ratings.

7) Buffer command is used for rating each cell under the study area.

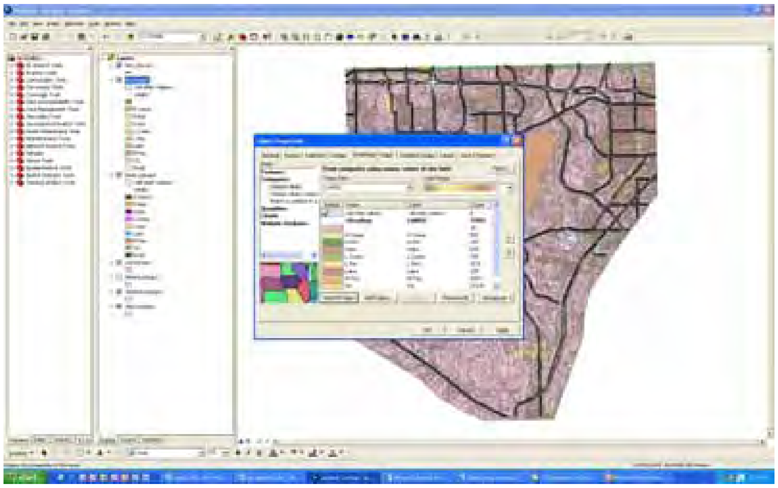


8) Specify factor weights.

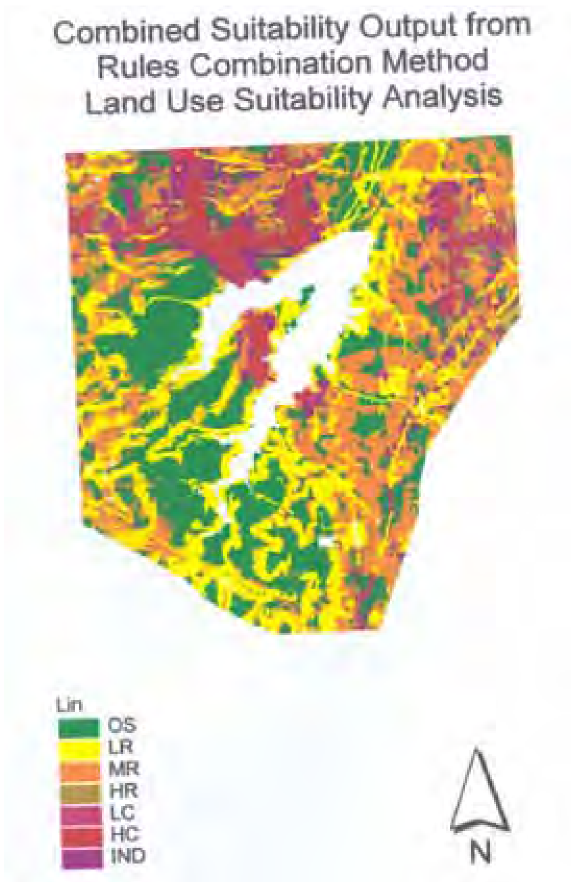
9) Suitability for each land use is the sum of the adjusted ratings for all the factors.



10) The final output map could be obtained via the standard Rule-based Combination method.

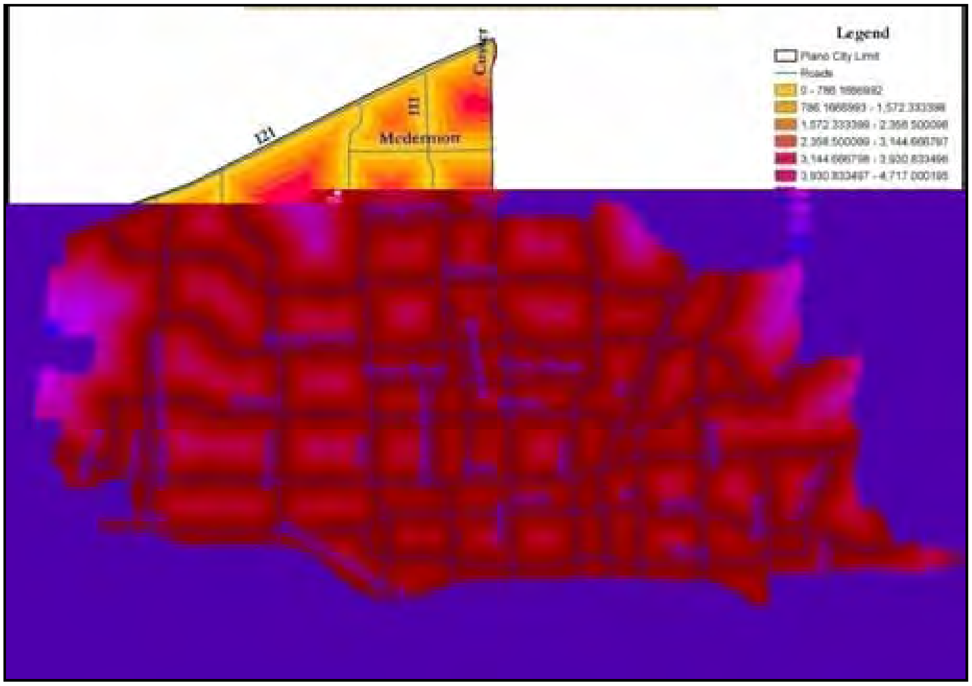


11) Final, conventional, rule-based composite land use suitability map.



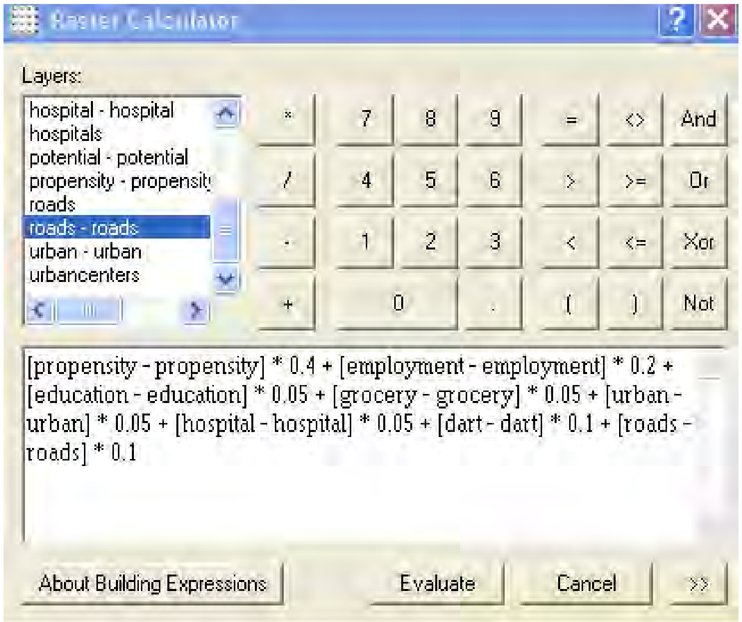
The process of land suitability analysis in ArcGIS Spatial Analyst

1. After the selection of the study area, convert the shapefiles into raster format for GIS spatial analyst operations. For further analysis, perform buffer operation on the selected area.



2. Reclassify the buffered map for assigning the rates and weights.

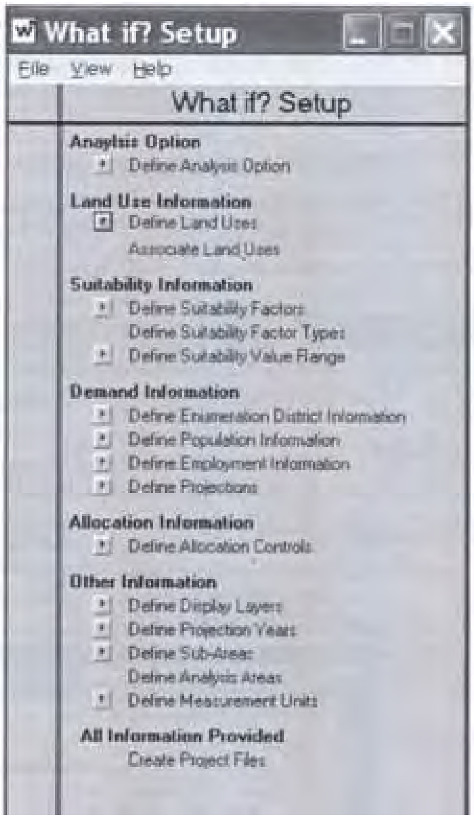
3. Depending on demand, assign percentages to each considered land use.



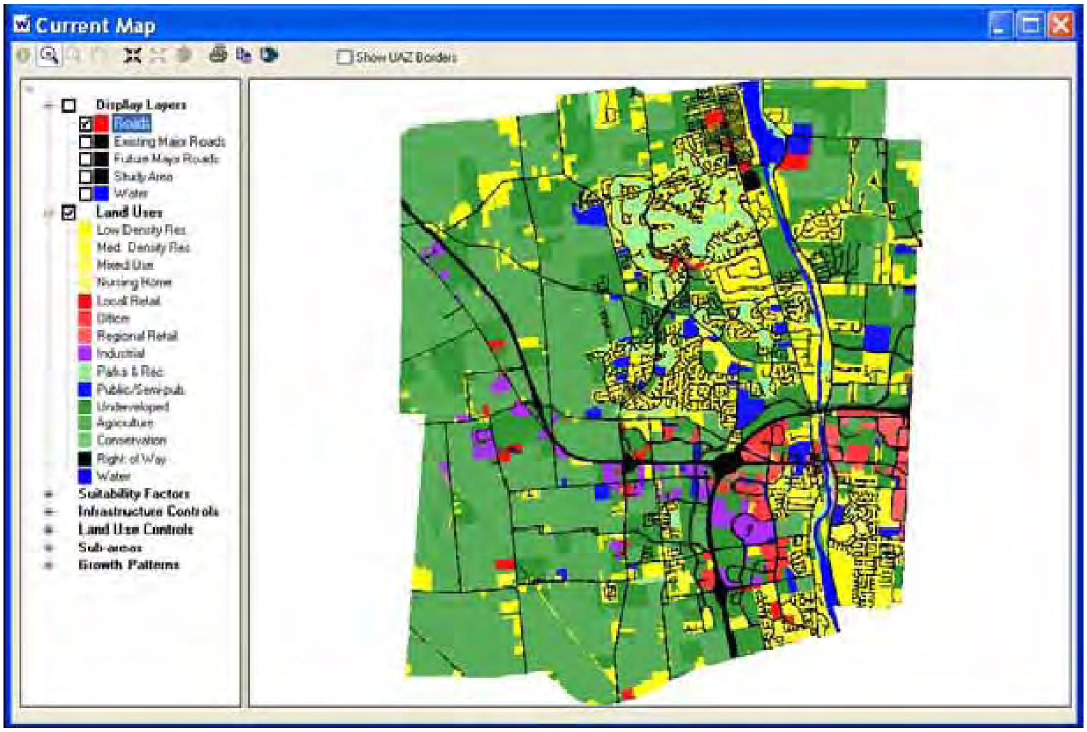
4. The result will produce separate suitability maps for each land use.

The *What if?* Process

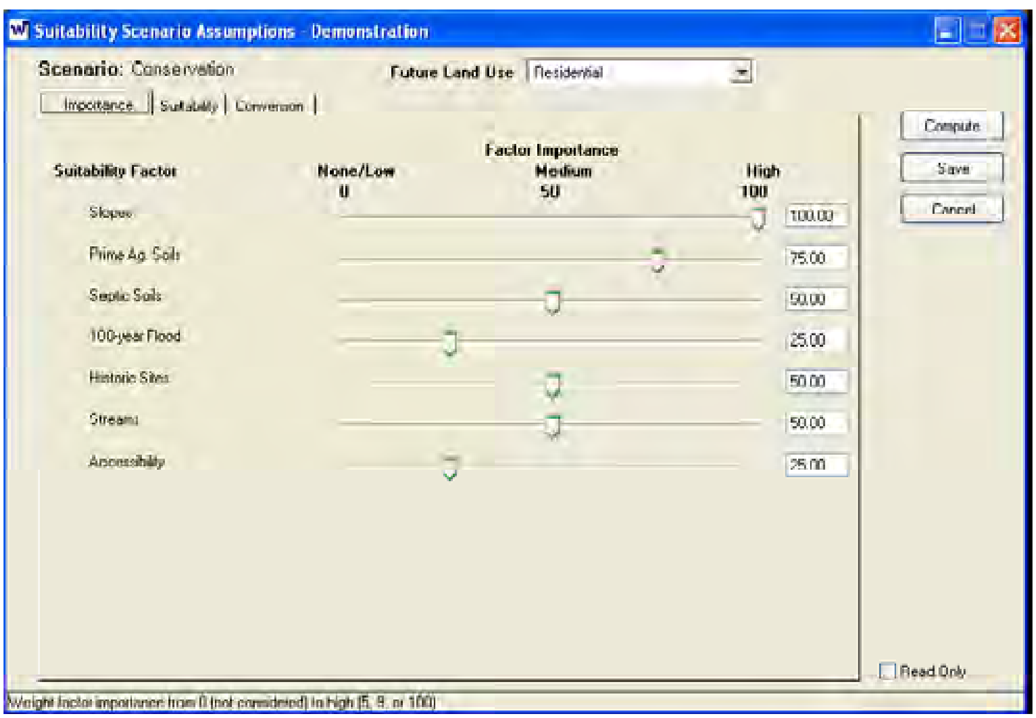
1. Define the analysis option.



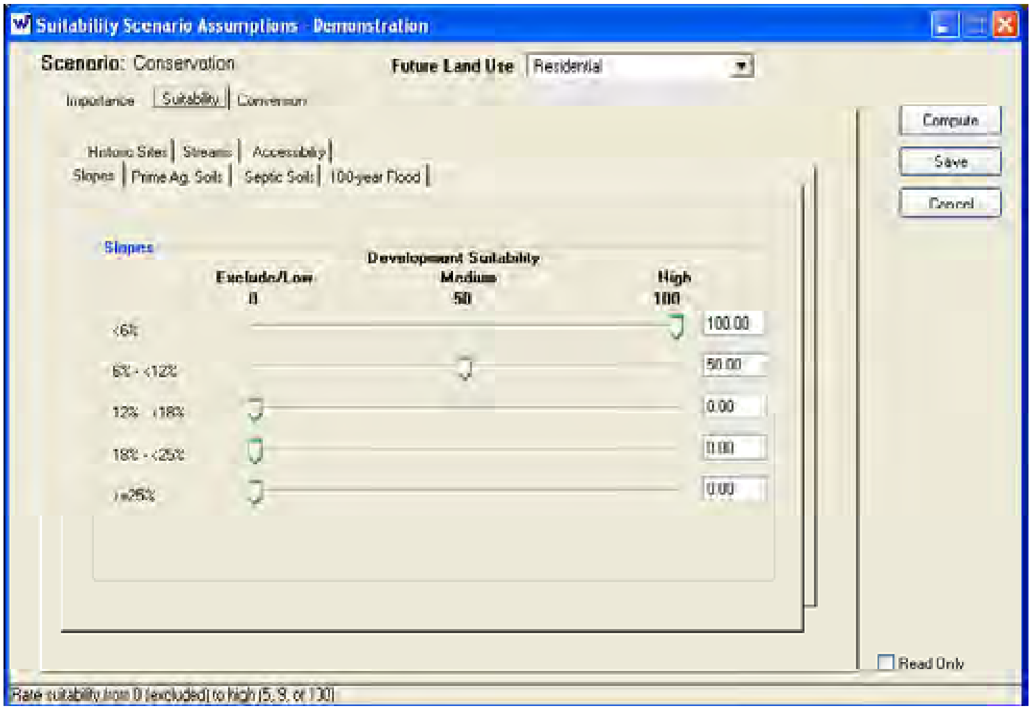
2. View the current map.



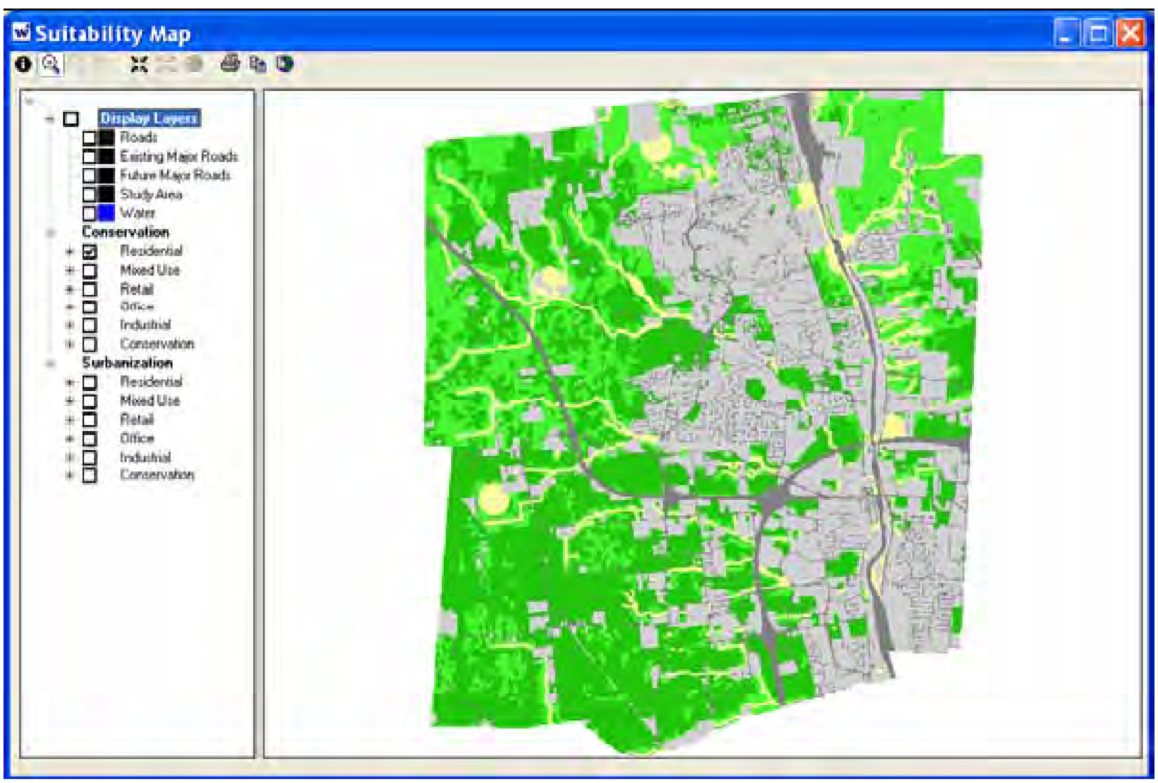
3. This form provides slider bars that make it easy to determine the relative importance of different suitability factors for locating each land use.



4. Allocate the rates and weights for the factors considered.

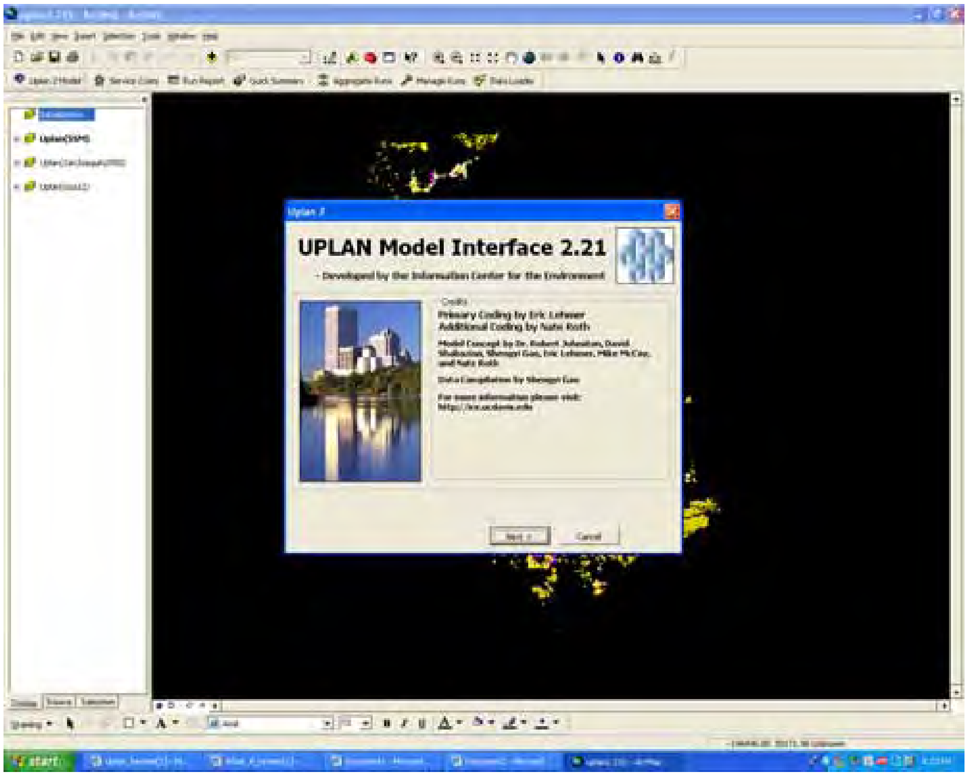


5. Generate the map after assigning the rates and weights. The Suitability map shows each location's suitability for a particular land use, scaled from "Not Suitable" to "High," as determined by the suitability scenario assumptions. The map also identifies areas that are not developable and cannot be converted from their current use.

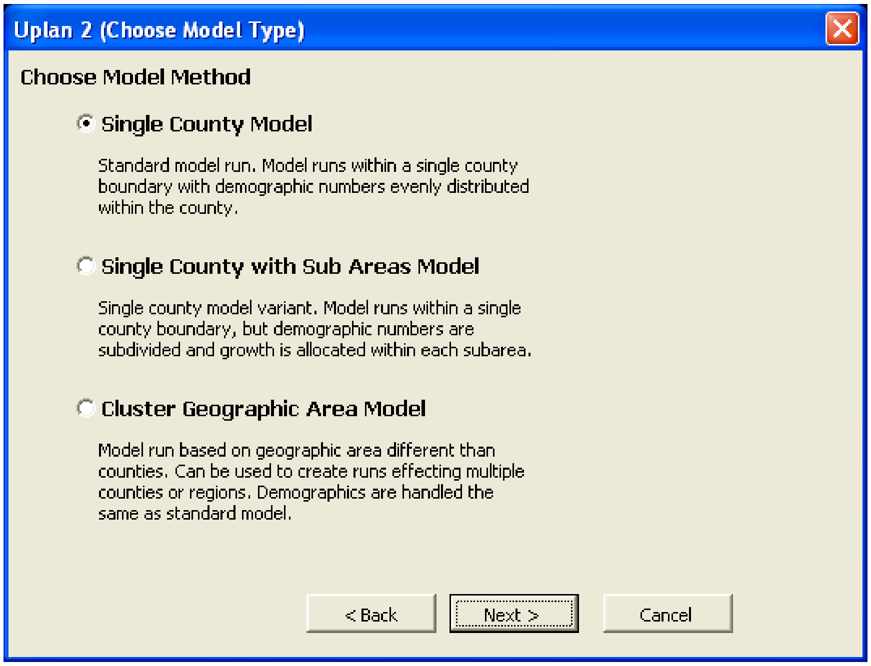


The UPlan Process

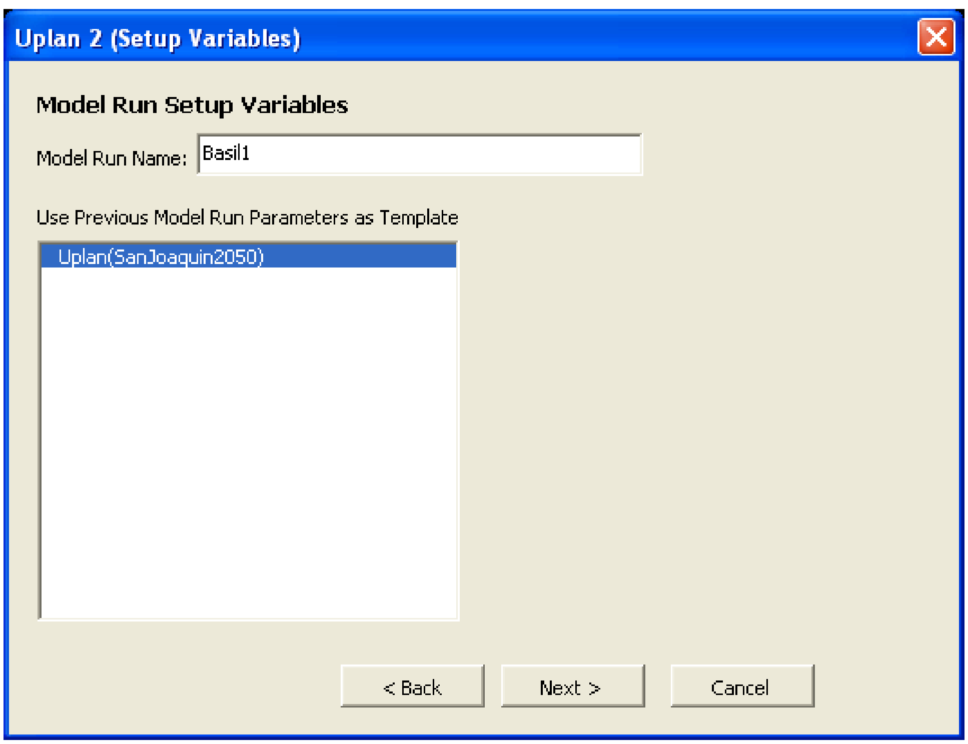
1) Start UPlan through the UPlan model interface, and click “next” to choose the sub-model interface.



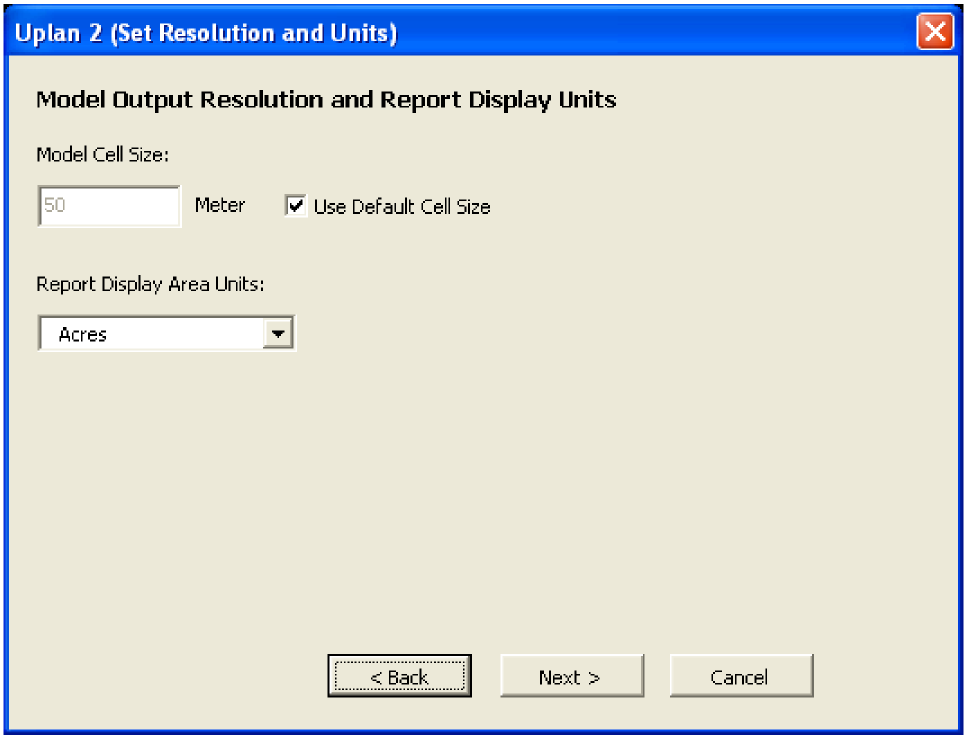
2) Choose model type (Single County Model, Single County with Sub Area Model, Cluster Geographic Area Model).



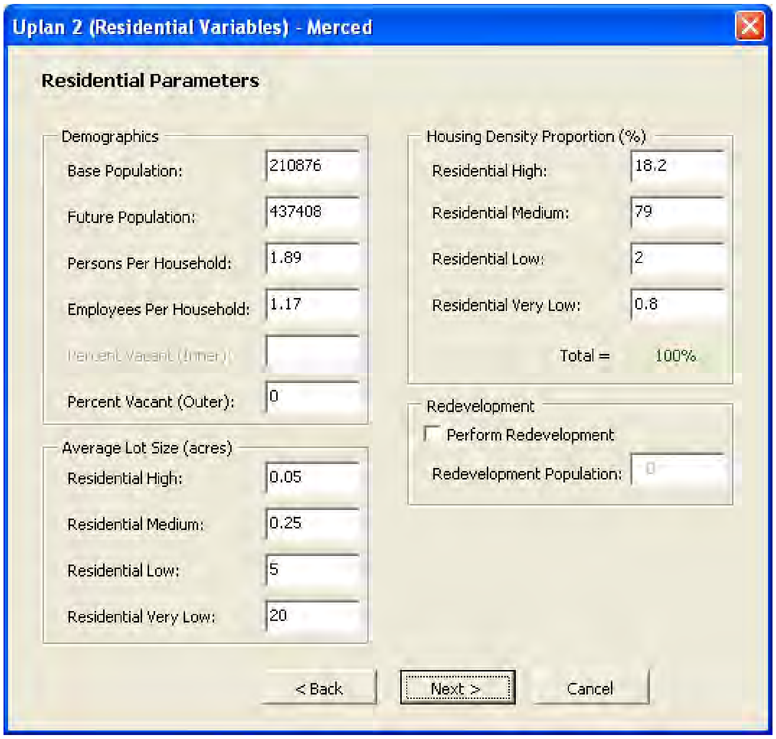
3) Naming the model run (any symbols can be used to run the model, and there are no limitations in length of the run).



4) Selecting county and setting resolution and units.



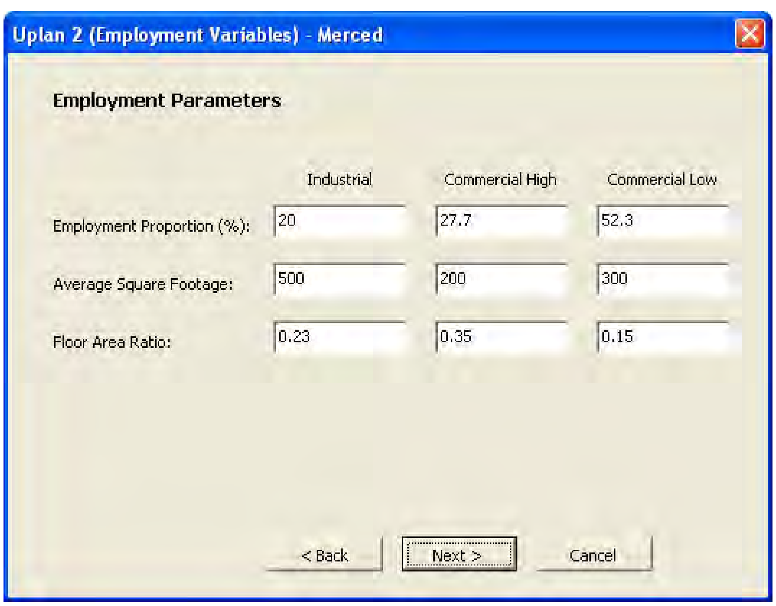
5) Model Input data: Residential Inputs
   %Res. High, *%Res. Med.*, *%Res. Low*, *% Res. Very Low*—These are used to specify the proportion of households in each of the four density categories:
   - High Density (Attached multifamily dwellings)
   - Medium Density (Mostly detached single-family dwellings)
   - Low Density (Rural dwellings with own well and septic)
   - Very low density (Rural dwellings with own well and septic)
   - The total should be 100%.



6) Model Data: Employment Inputs

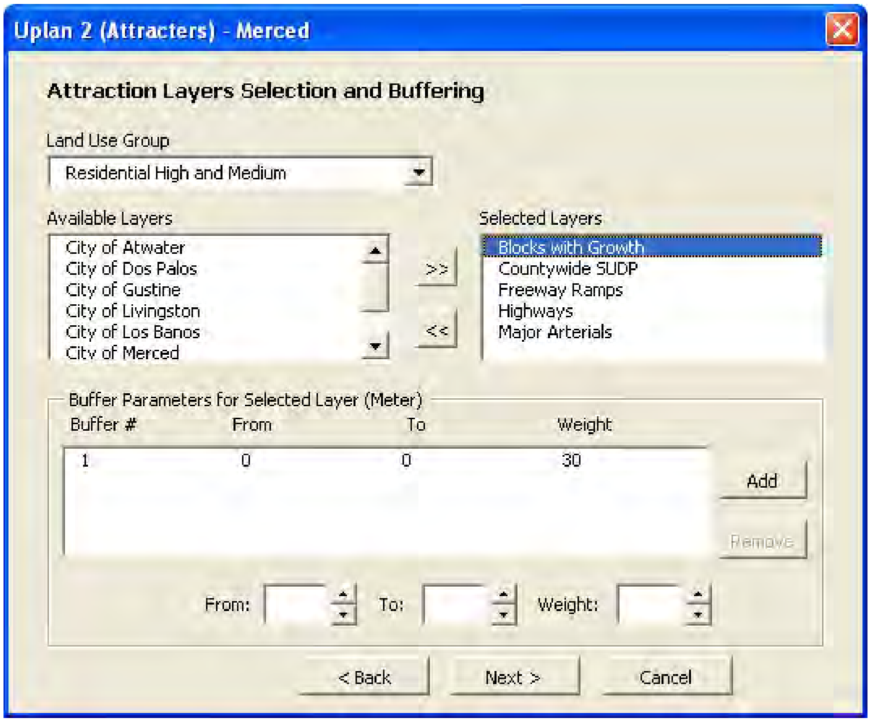
% Empl in Ind, *% Empl in Comm. High*, *% Empl in Comm. Low*—These parameters are used to set the proportion of employees in each of the broad economic sectors.

- Industry, light and heavy industry
- High-density Commercial, FAR > 2.0 or defined by the user
- Low-density Commercial, FAR ≤ 2.0 or defined by the user
- The total ≤ 100

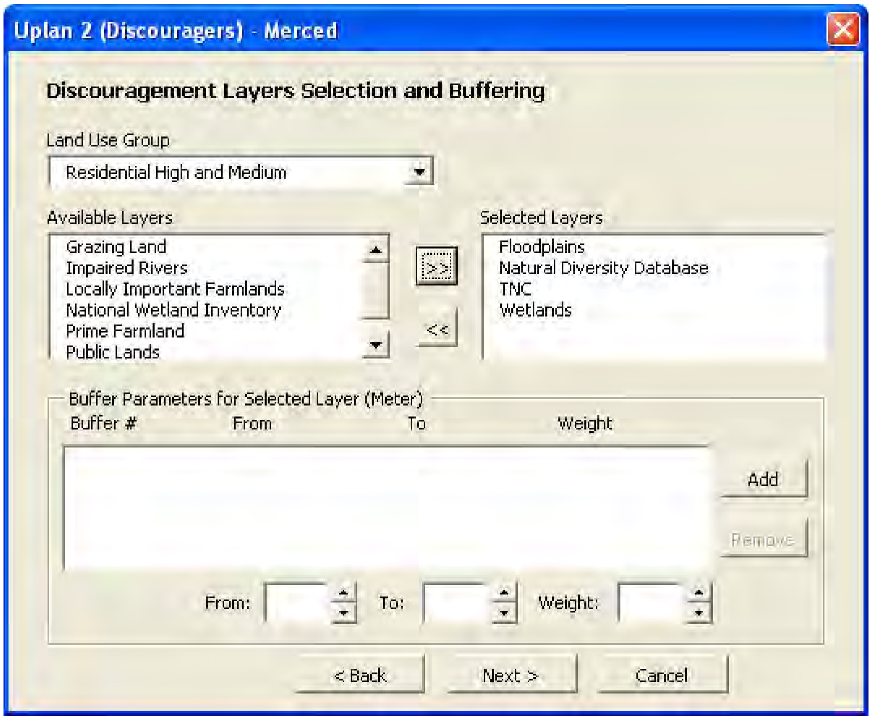


7) Choosing attractions and discouragements

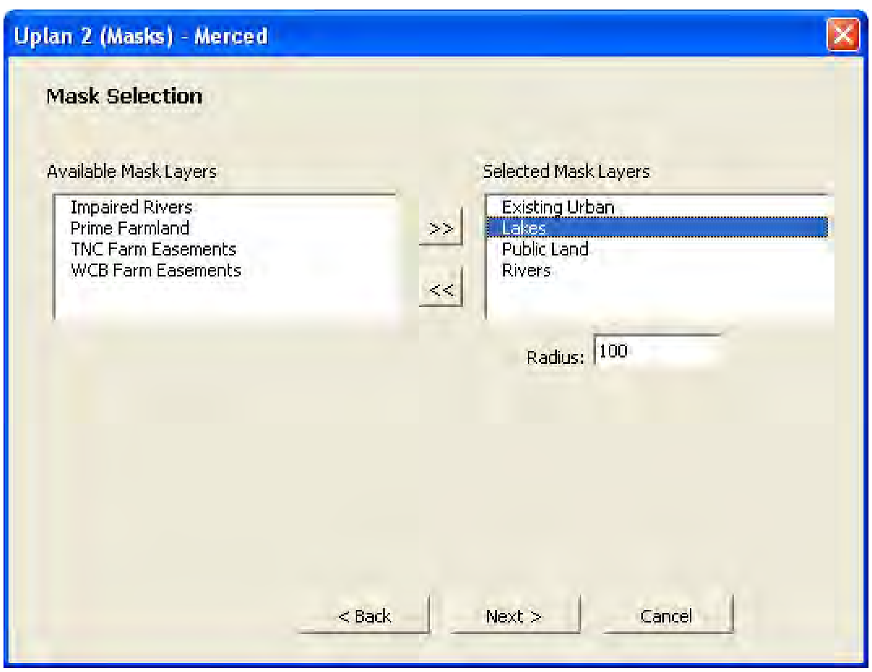
Attractions (such as freeways, other transport facilities, and city growth boundaries) can be been split into many different attraction types and each type assigned a weight.



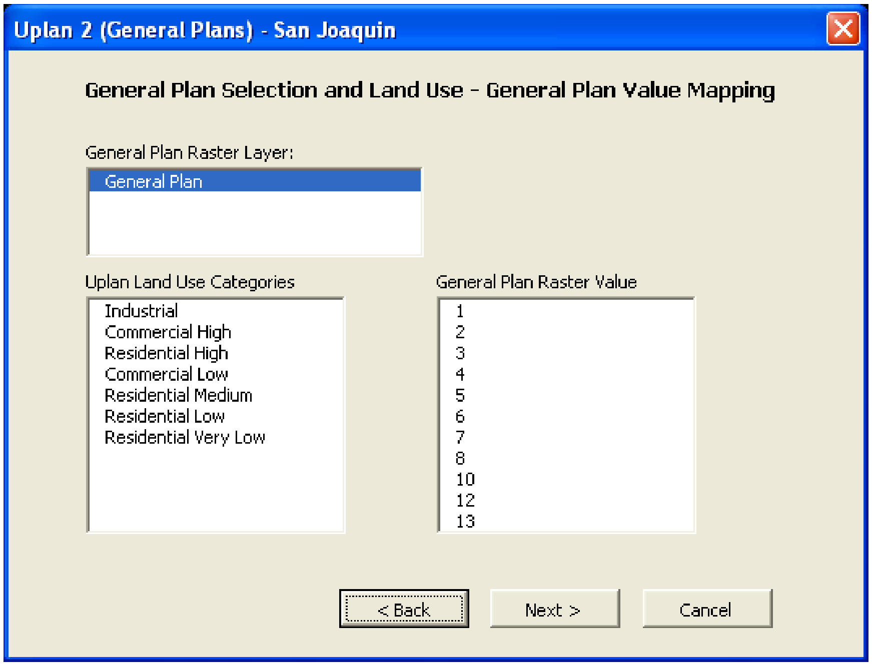
8) Discouragements (e.g., parks, watershed management areas) cannot be used for new development.



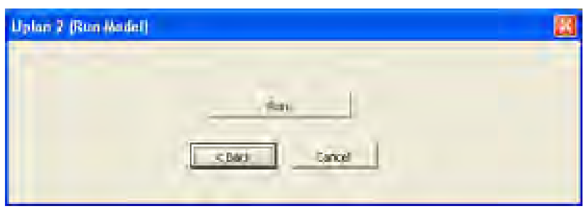
9) Mask layers (e.g., lakes and streams) are not developable (for jobs or housing).



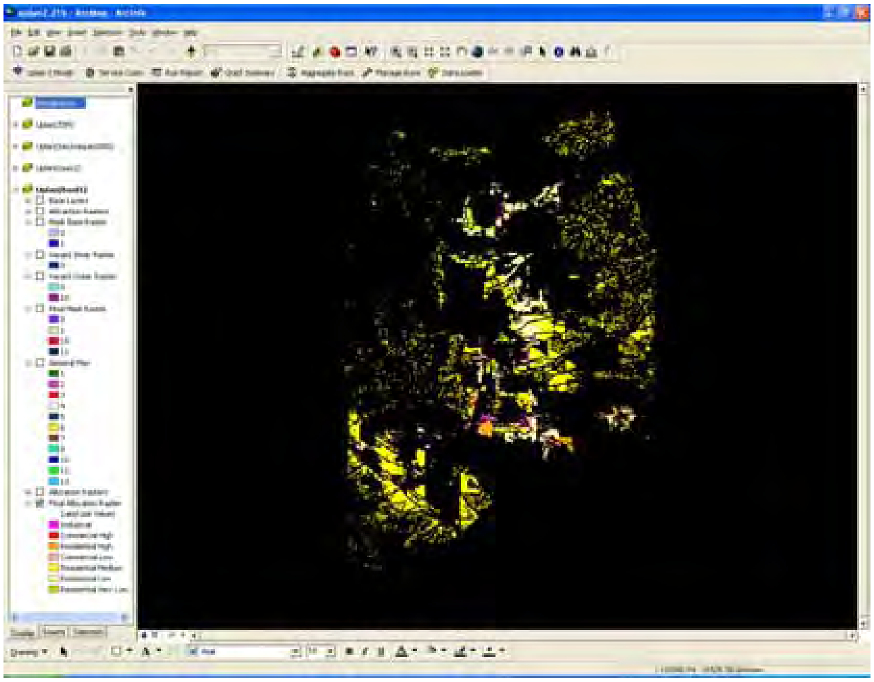
10) General plans and farmland are reclassified in the UPlan model and are represented by codes. (for example, Unclassified = 0, agriculture = 1, and industry =2).



11) Run model.



12) Final suitability map.



The Process of Land Suitability Analysis using the Anjomani Method (AM)

The Anjomani Method can utilize most any of the GIS software and models already described, in the initials stages. The first five ArcGIS stages (presented above), for example, are part of an Anjomani application and, therefore, will not be repeated here. Those steps continue as follows:

1) Specify factor ratings based on the internal characteristics of each suitability factor depending upon their contribution to the study area in terms of rates (Table D1).

Table D1: Rating of Each Suitability Factor from 0 to 10 for Each Land Use (least suitable and best suitable, respectively).

Flood Zones	OS	LR	MR	HR	LC	HC	IND	RD
100 Year Flood	10	10	10	10	10	10	10	10
Outside Flood	6	6	4	2	3	2	1	5

Water Lines	OS	LR	MR	HR	LC	HC	IND	RD
Within 100 Meters	4	10	10	10	10	10	10	7
Btwn 100 & 500 Meters	6	8	8	7	8	7	7	7
Btwn 500 & 1000 Meters	7	6	5	4	5	5	5	7
Btwn 1000 & 3000 Meters	10	4	3	2	4	2	2	7
Over 3000 Meters	10	3	3	1	3	1	1	7

Sewer Lines	OS	LR	MR	HR	LC	HC	IND	RD
Within 100 Meters	6	9	10	10	10	10	10	7
Btwn 100 & 500 Meters	6	7	8	7	8	7	7	7
Btwn 500 & 1000 Meters	8	6	6	5	6	5	5	7
Btwn 1000 & 3000 Meters	10	5	5	2	5	2	2	7
Over 3000 Meters	10	4	3	1	3	1	1	7

Tree Cover	OS	LR	MR	HR	LC	HC	IND	RD
Woodlands, Generalized	10	8	7	5	4	3	2	4
No Tree Cover	9	8	8	8	8	8	8	10
Urban Areas	6	6	7	8	7	8	8	7

Slope	OS	LR	MR	HR	LC	HC	IND	RD
Slight slope	10	10	10	10	10	10	10	10
Moderate slope	6	4	4	4	3	3	3	3
Steep Slope	4	2	1	1	1	1	1	1

Existing Land Use	OS	LR	MR	HR	LC	HC	IND	RD
Light Density Residential	10	10	8	3	5	2	2	6
Medium Density Residential	8	8	10	6	6	4	2	6
High Density Residential	7	2	5	10	8	6	5	6
Light Intensity Commercial	6	6	6	8	10	8	6	8
High Intensity Commercial	2	4	4	5	7	10	7	8
Industrial Areas	2	2	2	4	6	6	10	8
Vacant / Open Areas	10	10	10	10	10	10	10	10
Roads	4	5	6	6	8	8	8	10

Major Roads	OS	LR	MR	HR	LC	HC	IND	RD
Within 50 Meters	4	4	4	6	10	10	9	7
Btwn 50 & 200 Meters	5	7	7	8	9	9	9	7
Btwn 200 & 500 Meters	6	8	8	8	8	7	9	7
Btwn 500 & 1000 Meters	6	9	9	7	7	5	7	7
Btwn 1000 & 2000 Meters	8	8	7	6	5	3	5	6
Over 2000 Meters	8	7	6	5	4	2	3	3

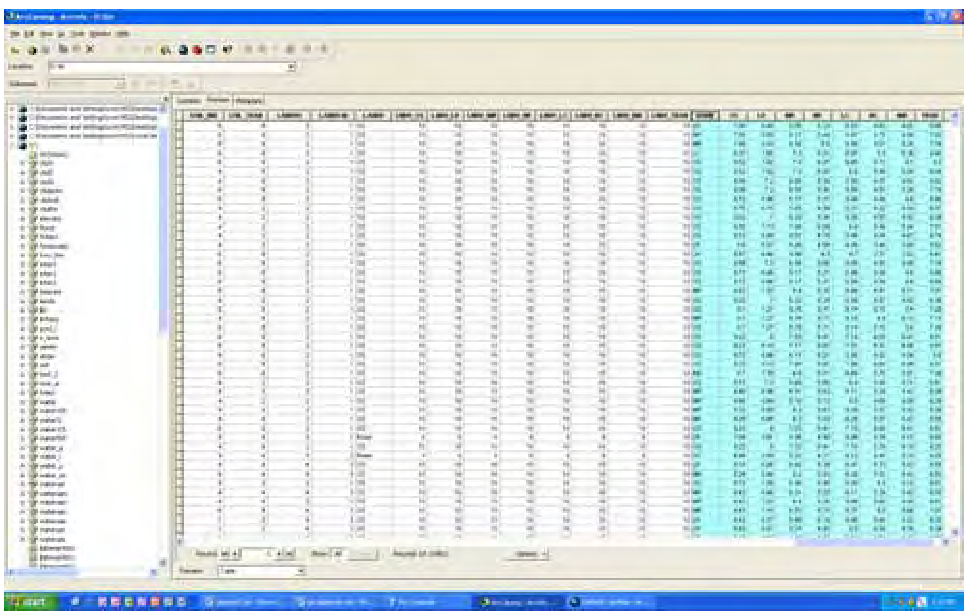
2) Similar to ArcGIS's Buffer command, this table is used to rate each cell in the study area.

3) Specify factor weights for each suitability factor, in terms of its relative importance for land use under study. The Anjomani Method suggests using a Delphi technique (Lindstrom and Turoff, 1977) or multi-criteria evaluation techniques such as the Analytical Hierarchy Process (AHP) (Saaty 2002) to derive the factor weights. Table D2 shows the results of applying a Delphi technique with a group of experts to derive these weights. (Note that the Delphi technique is characterized by three features: anonymity of experts, iteration with controlled feedback, and tallied responses over several sequential rounds (by the survey "director") (Dicky and Watts, 1978).

Table D2: Weightings for Each Land Use Suitability Factor from 0 to 10.

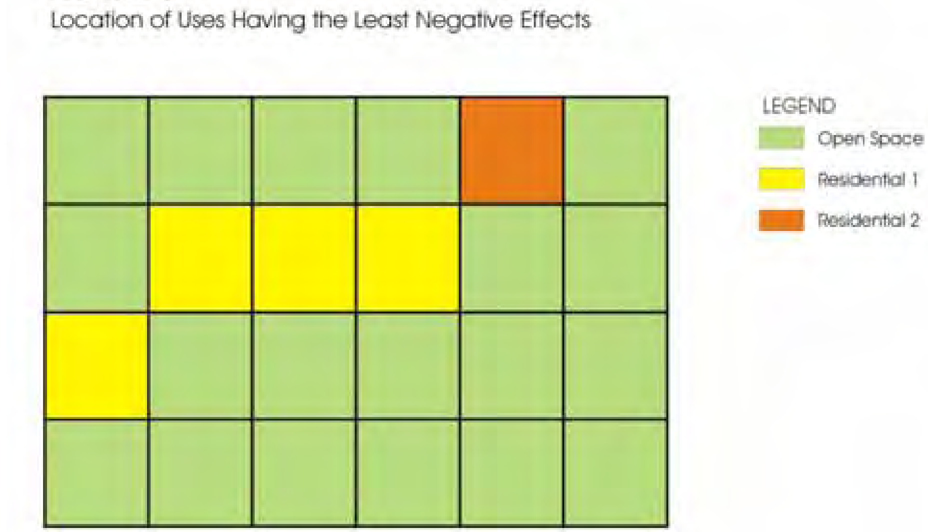
	Flood	Water	Sewer	Trees	Slope	Land Use	Roads	Soil type
OS	2.4	1.1	0.9	6.5	3.0	4.7	3.3	0.7
LR	7.8	4.1	4.2	5.6	5.1	8.2	3.2	4.1
MR	8.1	5.6	5.5	3.3	5.7	7.2	5.2	5
HR	7.9	6.3	6.6	3.9	5.9	6.7	7.1	5.6
LC	7.5	5.4	5.6	1.7	6.6	6.5	6.8	5.8
HC	7.5	7.1	7	1.9	6.8	7.6	8.8	6.9
IND	6.3	6.3	6.5	1.3	7	7.1	7.5	6.4
RD	5.8	1.6	1.6	2	6	7.5	8.1	4.9

4) Suitability for each land use at each site is the sum of the ratings for all the factors multiplied by their respective weights.

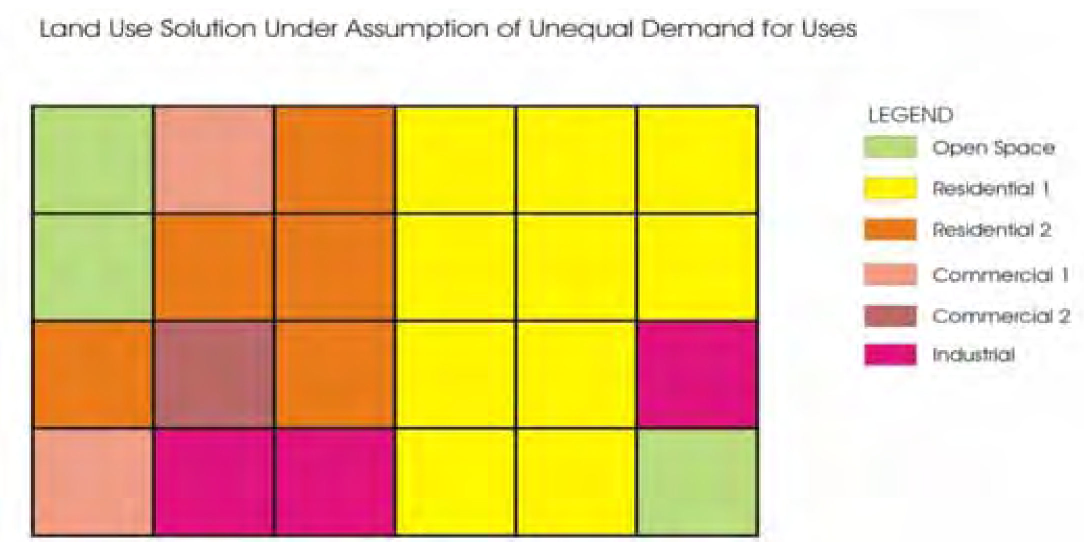


5) The Anjomani method explores different techniques for deriving the composite land use plan and provides guidelines and suggestions. For example, if the suitability scores are sorted to represent least cost (or maximum benefits), these may represent rational market decisions for

future development of the area. The following figure illustrates such a result for a small example area.



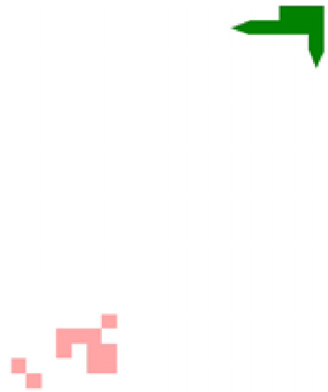
6) In deriving socially optimal results for an entire study area, the AM requires use of constrained optimization techniques. The following figure shows such a result for the same (above) example area and data set. One advantage of this technique is its consideration of demand for different land uses (or types of developments) in the study area, thus avoiding over- or under-assignment of various land uses.



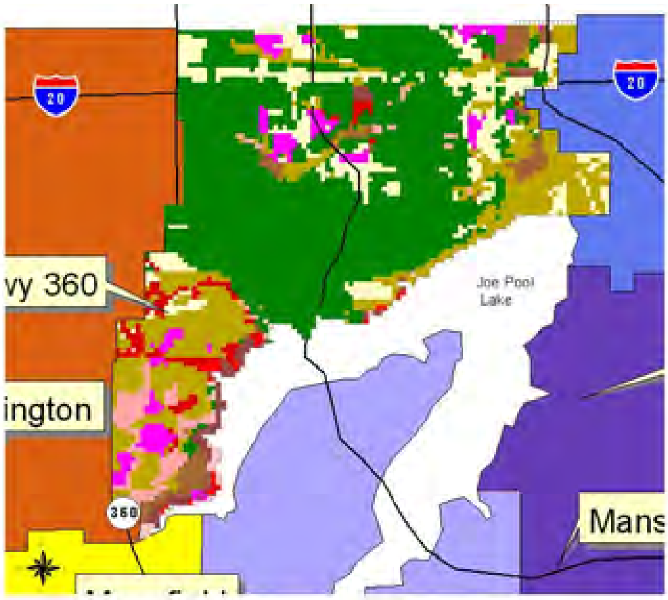
7) Anjomani's Method also suggests use of regression analysis if one seeks to derive results based on the continuation of past land development behaviors.

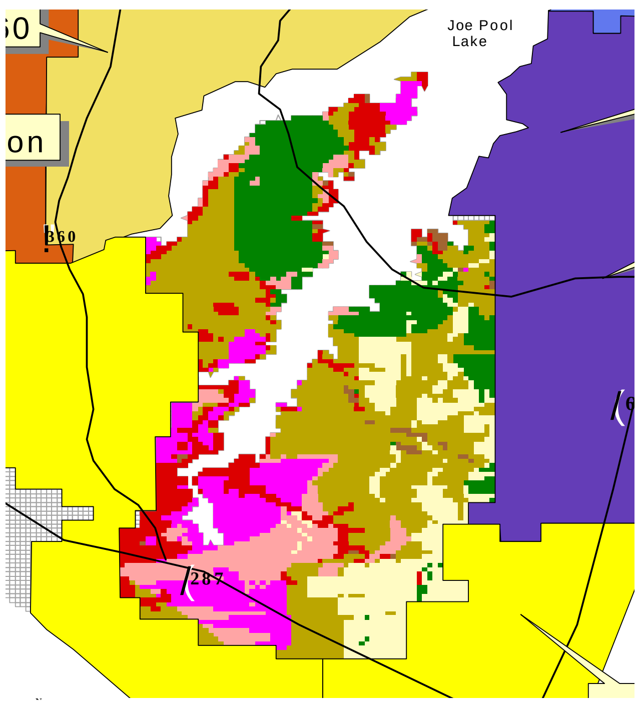
8) In order to fine tune the results, the AM proposes a Planning Support System process for iterative improvements, and explicit inclusion of design details for both land use and transport system plans.

9) The Anjomani Methods also recognizes two other important factors in deriving land use plans. These are the consideration of regional (economic) and jurisdictional (political) forces. The AM process goes through two additional rounds of analysis to improve the results based on consideration of these two factors. As an example, the following two figures illustrate attempts as a regional analysis round to improve an economic center and highway corridors, respectively.

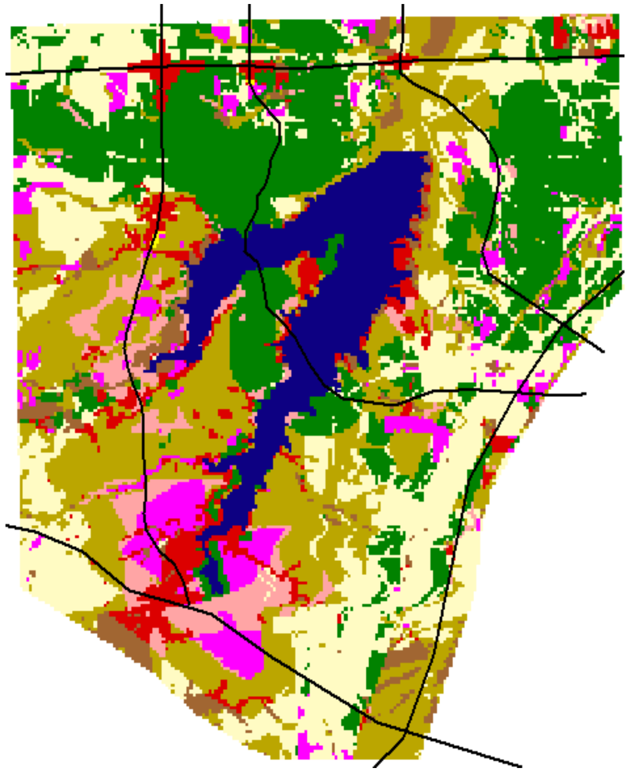


10) Similarly, the following two figures show, as part of the jurisdictional analysis round, attempts to improve major sections of a municipal area (one of the several political jurisdictions in the study area).





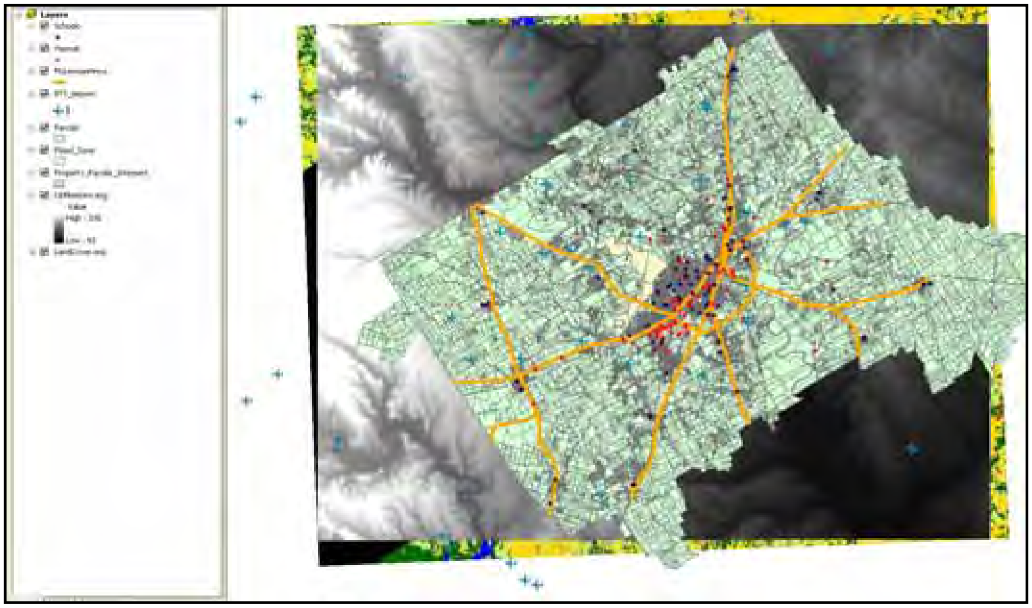
11) At the end of this planning process, AM brings together all the improved phases in a final land use plan which represents the results of an integrated and balanced approach, considerate of all the major factors.



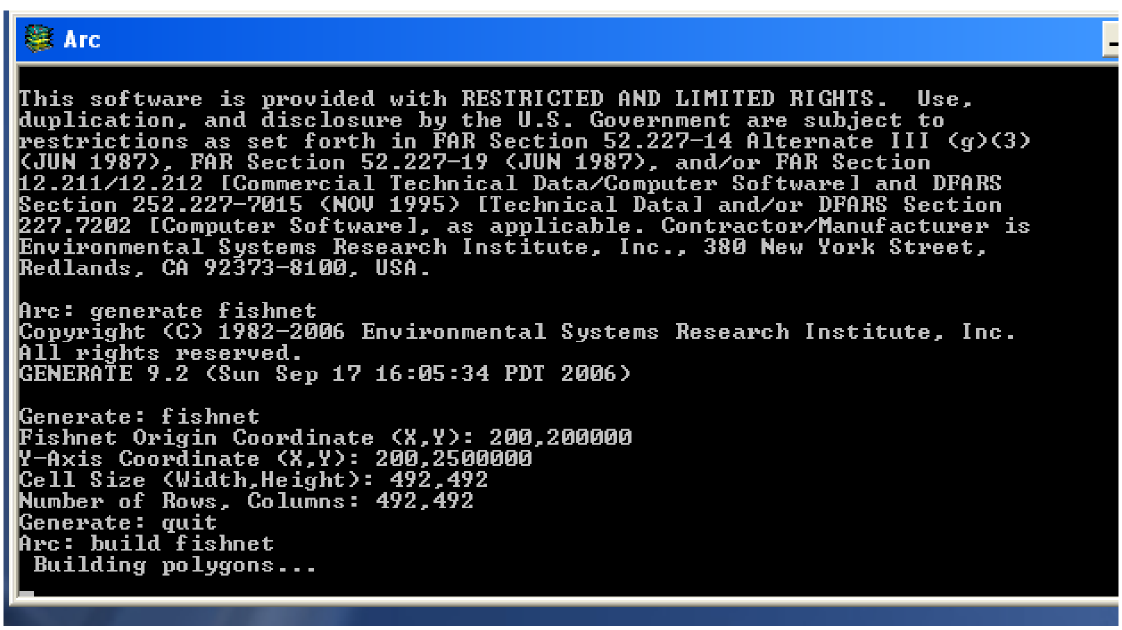
Appendix D2: Suitability Analysis in ArcGIS

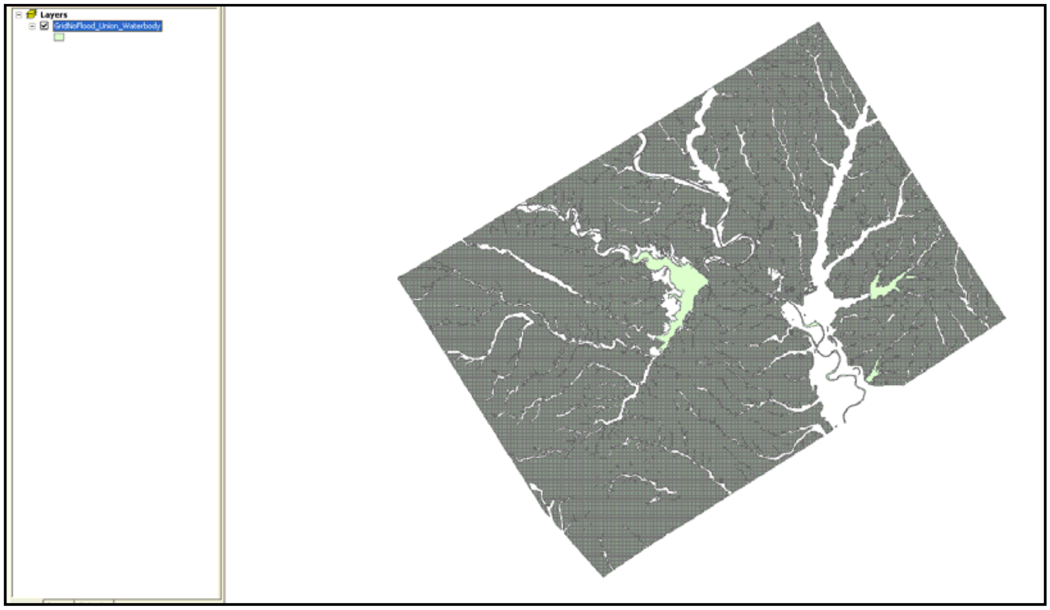
The process of land suitability analysis in ArcMap using ArcGIS

1) Map showing the shapefiles for the selected study area: Waco, McLennan County, Texas

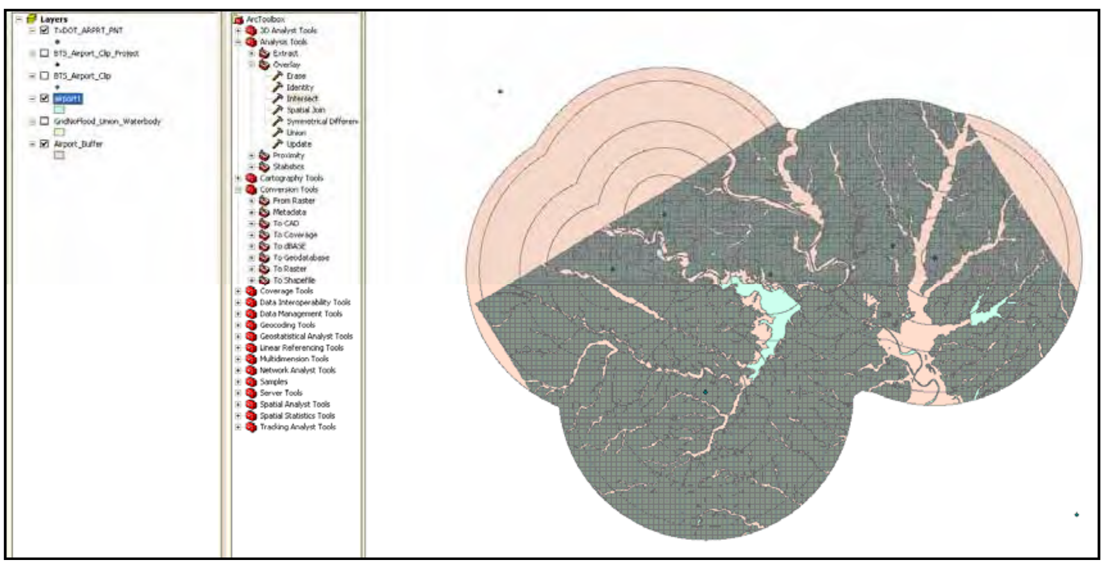


2) Grid/fishnet with a size of 150X150 meter was prepared and clipped to the study area (Waco) boundary

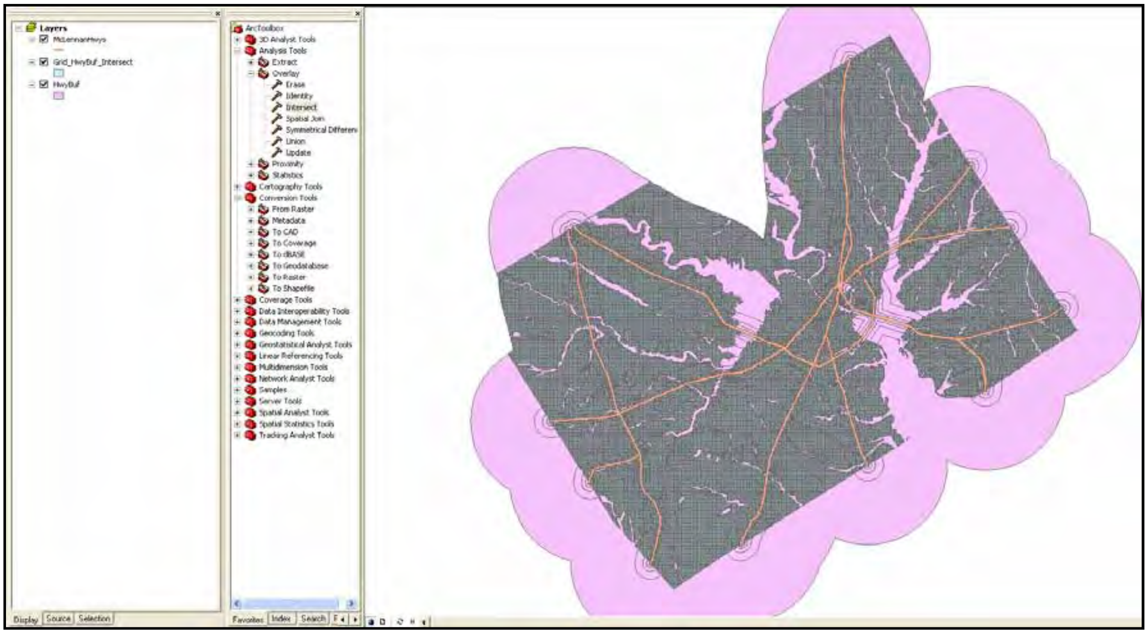




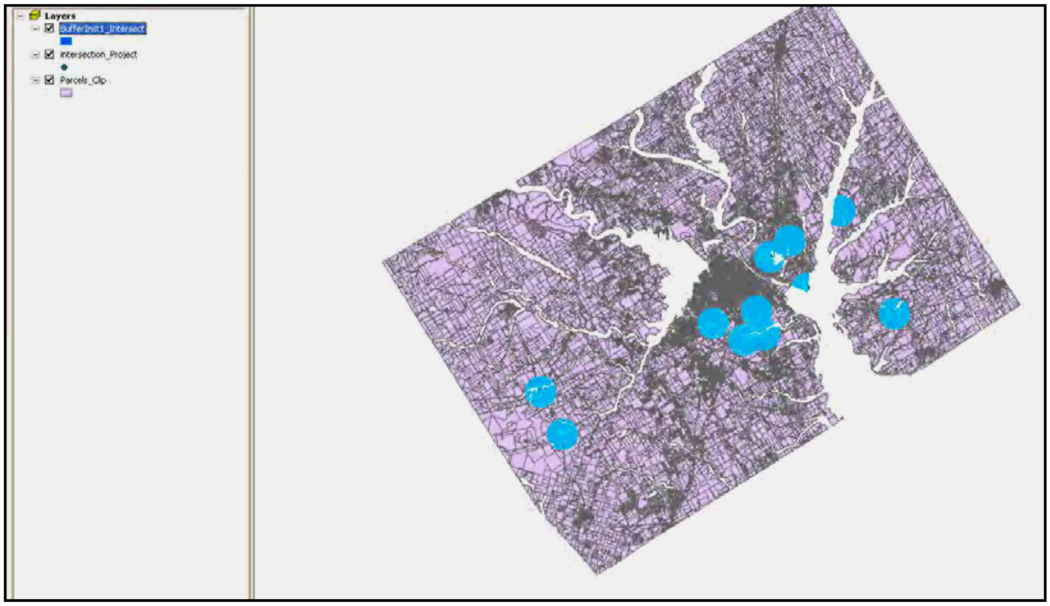
3) Map showing airport buffer (buffer distance: 1, 3, 5, 7, 10, and 11 miles)



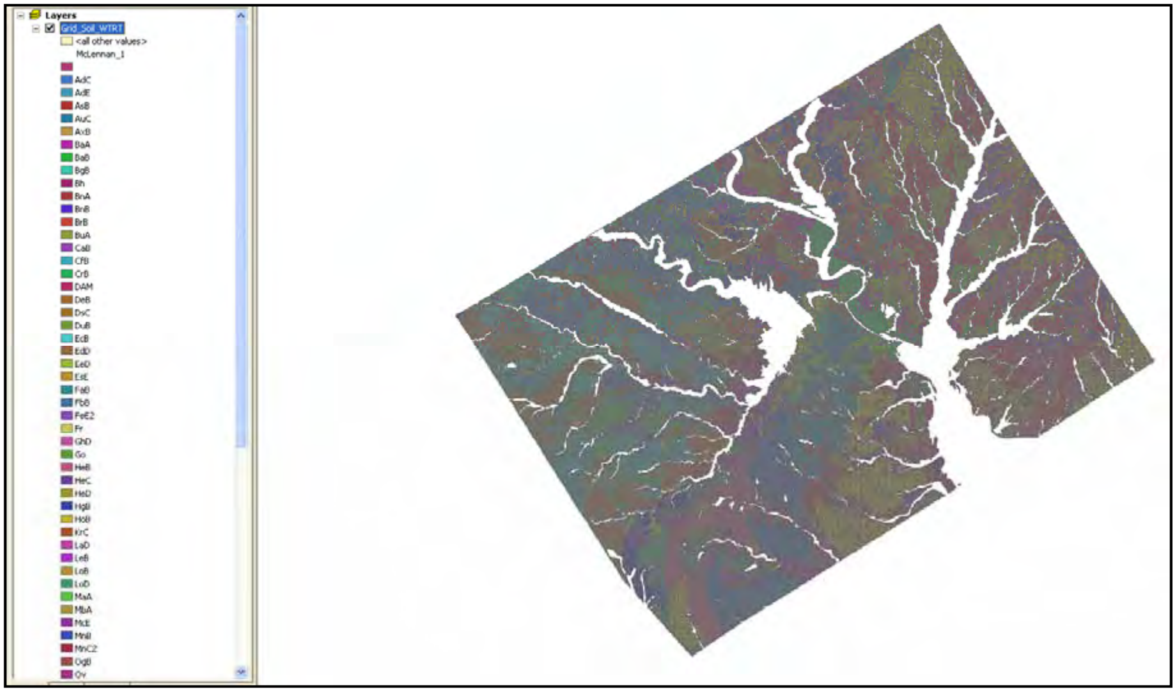
4) Map showing the Highway buffer with the study area boundary (buffer distance: 150, 600, 1500, 3000, and 6000 meters).



5) Map showing the major urban centers (buffer distance: 50, 200, 500, 1000, 2000, and 10000 meters).

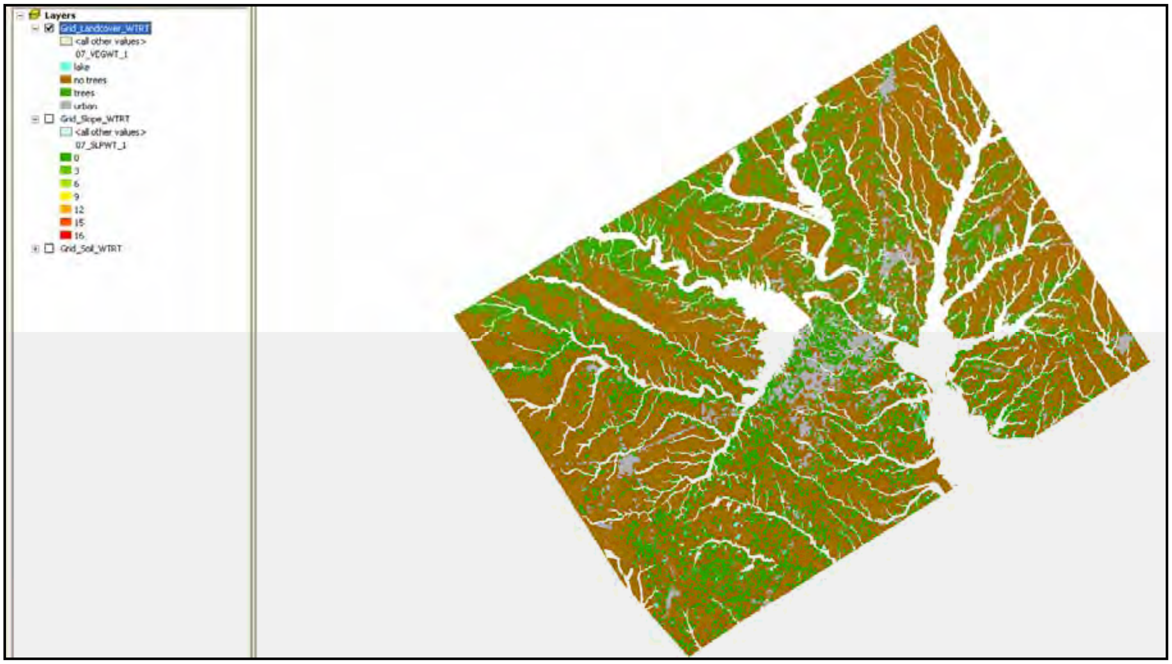


6) Map showing the soil shapefile with the study area boundary.

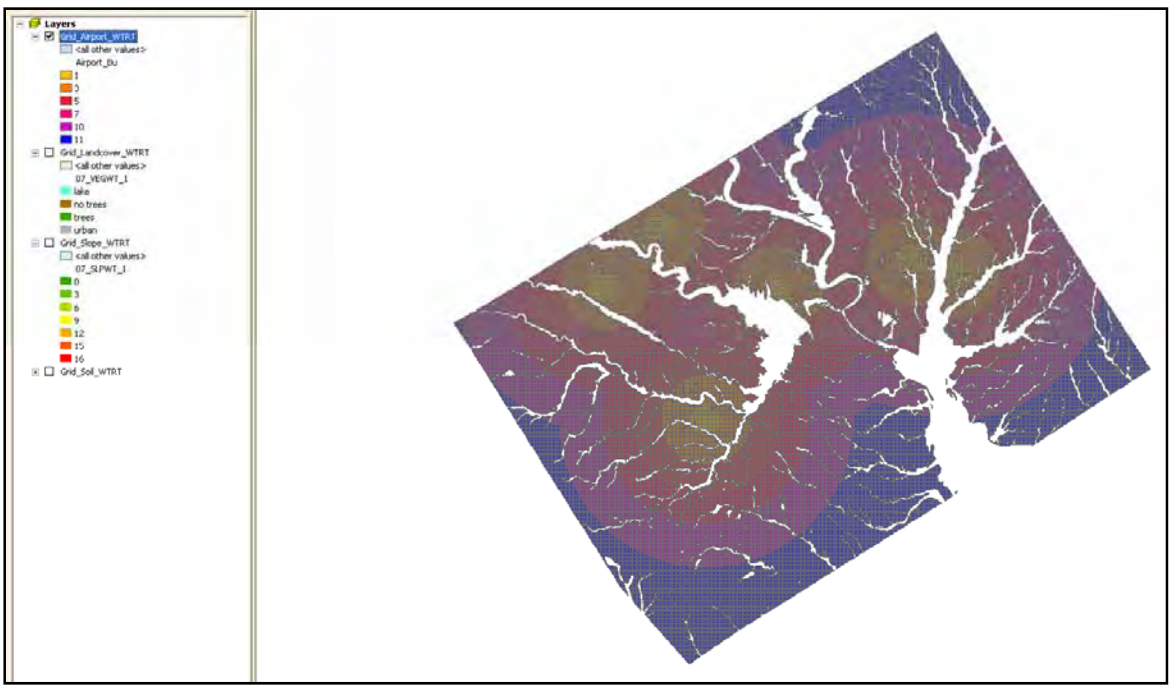


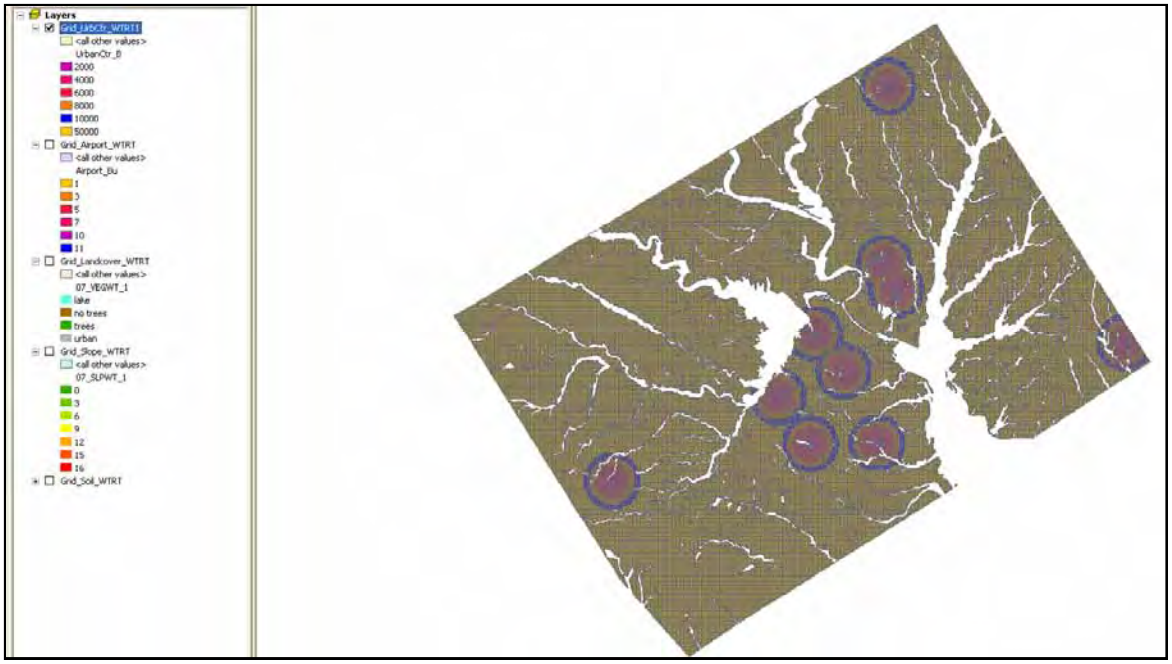
7) Map showing the slope shapefile with its respective weighted ratings.

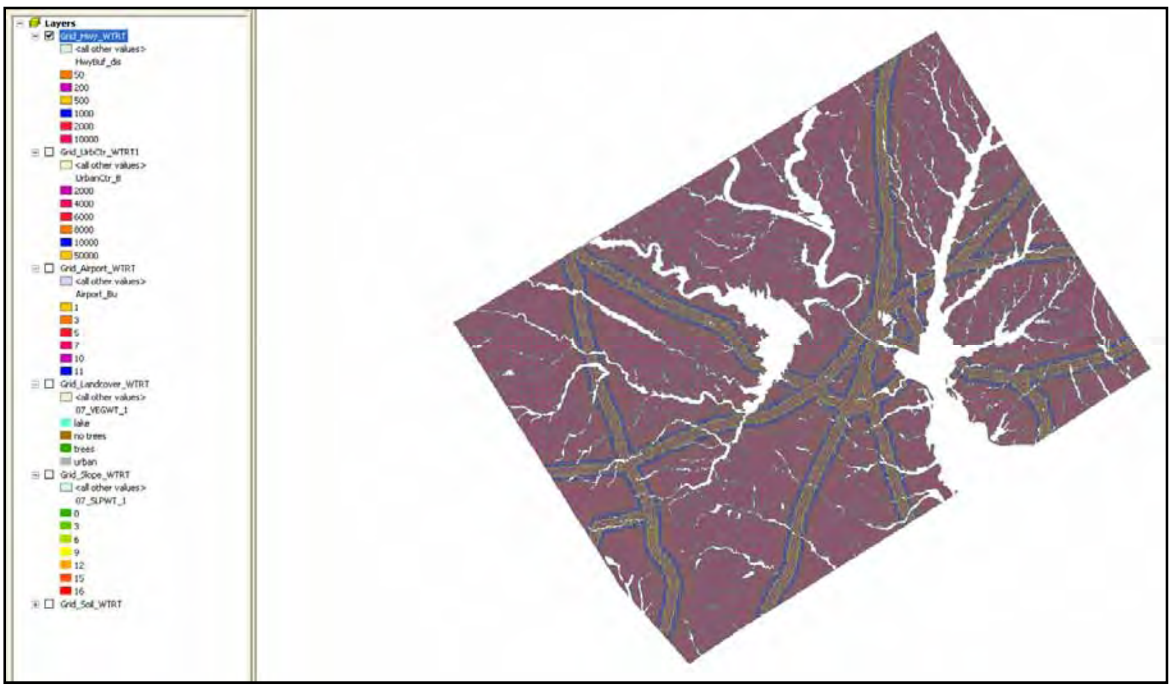
8) Map showing the land cover with its respective weighted ratings.



9) Map showing airports, urban centers, and highways with their respective weighted ratings.



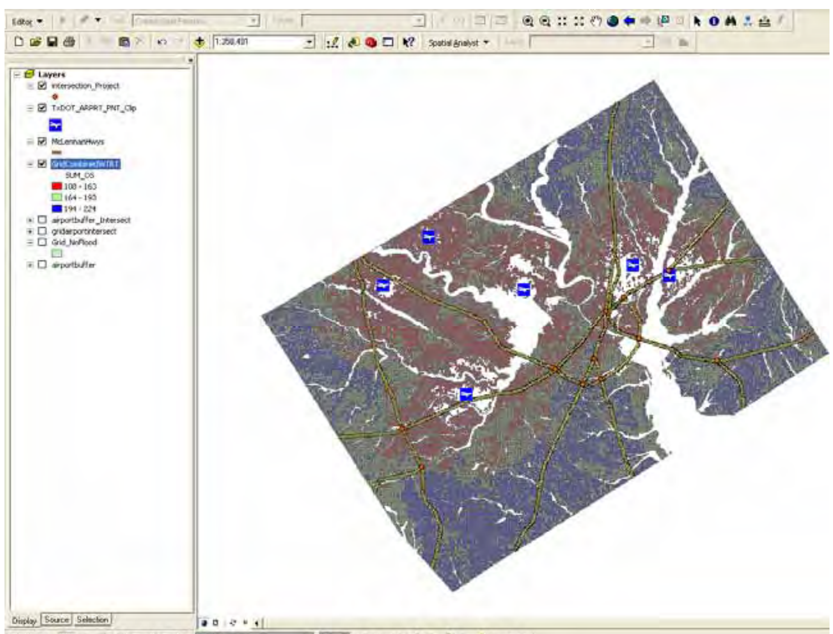


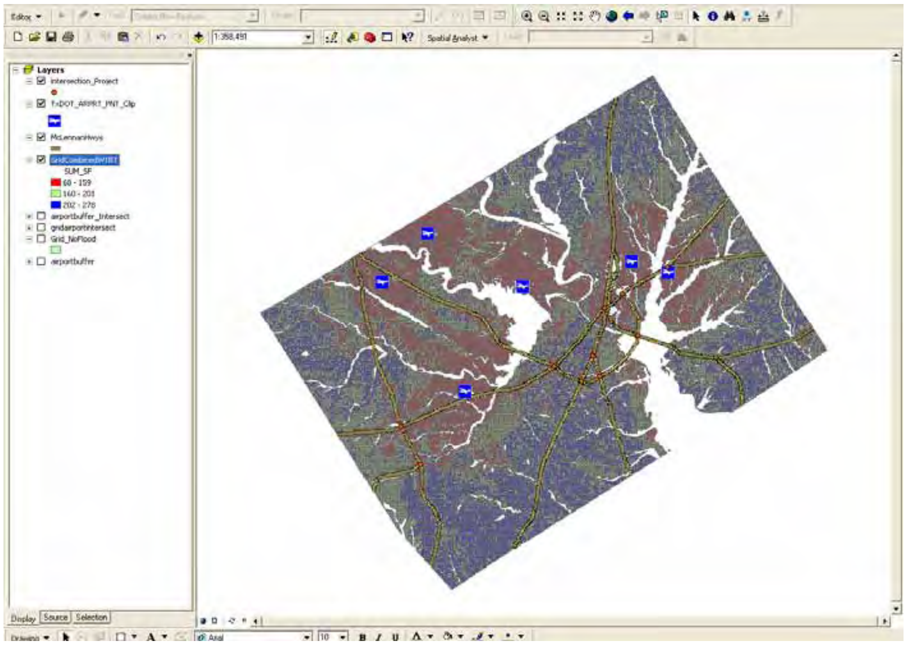


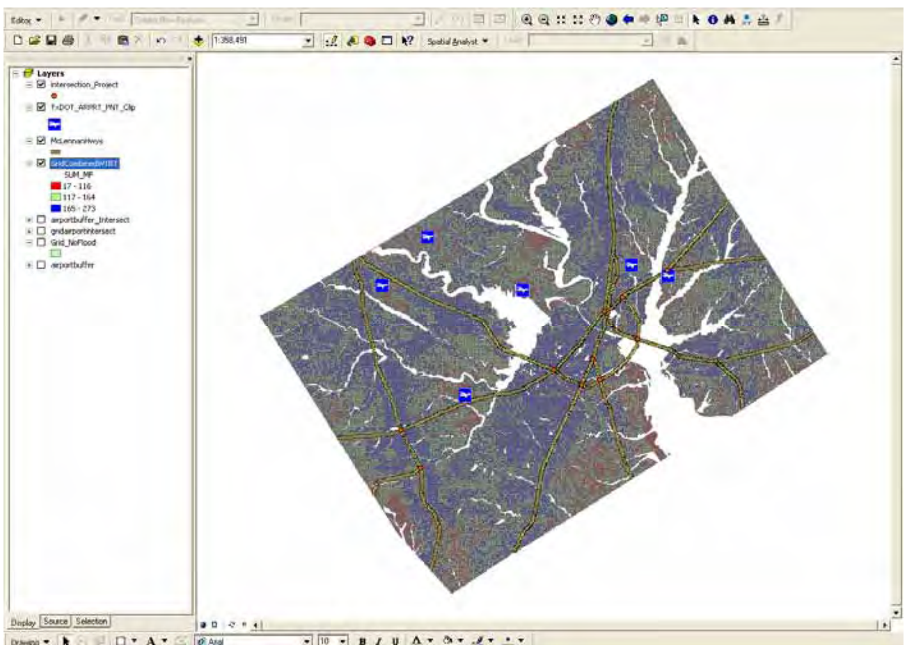
10) Attribute table showing all the factors with their respective weighted ratings and also the summed up values for all the selected land uses.

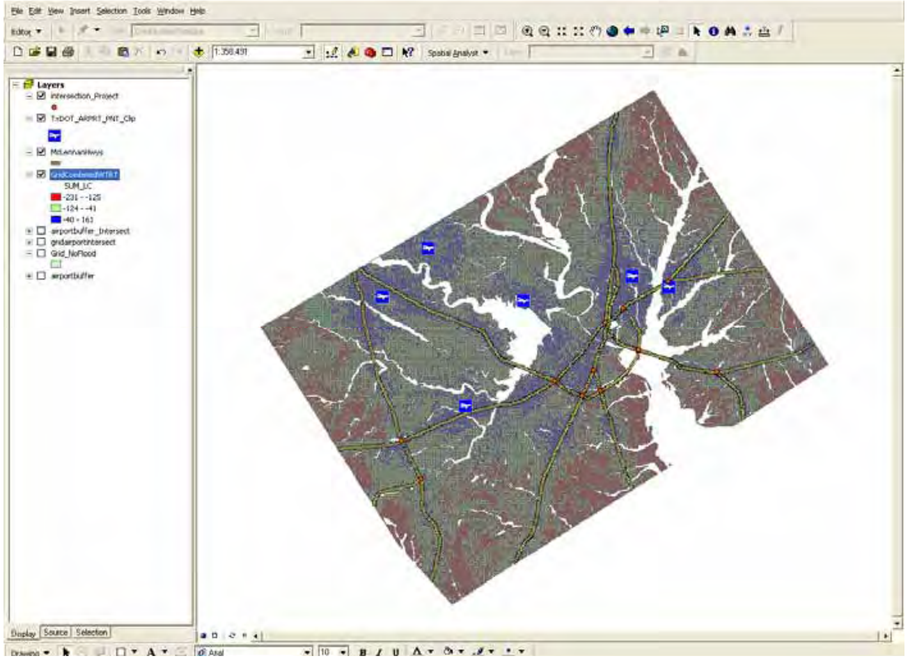
SOILW_5	07_SOILW_6	07_SOILW_7	HWY_OS	HWY_SF	HWY_MF	HWY_LC	HWY_HC	HWY_IND	SUM_OS	SUM_SF	SUM_MF	SUM_LC	SUM_HC	SUM_IND
28	29	47	0	0	0	0	0	0	63	3	66	210	238	258
55	58	59	0	0	0	0	0	0	64	8	90	232	245	277
55	58	59	0	0	0	0	0	0	64	8	90	232	245	277
55	58	59	0	0	0	0	0	0	64	8	90	232	245	277
55	58	59	0	0	0	0	0	0	64	23	65	218	221	249
55	58	59	0	0	0	0	0	0	64	23	65	218	221	249
55	58	59	0	0	0	0	0	0	64	8	90	232	245	277
55	58	59	0	0	0	0	0	0	64	8	90	232	245	277
55	58	59	0	0	0	0	0	0	64	23	65	218	221	249
55	58	59	0	0	0	0	0	0	64	8	90	232	245	277
17	29	18	0	0	0	0	0	0	65	58	55	-42	-91	-20
17	29	18	0	0	0	0	0	0	65	58	55	-42	-91	-20
17	29	18	0	0	0	0	0	0	65	58	55	-42	-91	-20
17	29	18	0	0	0	0	0	0	65	58	55	-42	-91	-20
17	29	18	0	0	0	0	0	0	65	58	55	-42	-91	-20
17	29	18	0	0	0	0	0	0	65	58	55	-42	-91	-20
17	29	18	0	0	0	0	0	0	65	58	55	-42	-91	-20
17	29	18	0	0	0	0	0	0	65	58	55	-42	-91	-20
55	58	59	26	32	48	-54	-44	0	66	19	113	197	241	277
55	58	59	26	32	48	-54	-44	0	66	19	113	197	241	277
-22	-23	-24	0	0	0	0	0	0	67	9	21	8	25	11
-22	-23	-24	0	0	0	0	0	0	67	9	21	8	25	11
-55	-58	-59	0	0	0	0	0	0	67	15	-11	-84	-166	-82
55	58	59	0	0	0	0	0	0	68	62	118	75	19	139
55	58	59	0	0	0	0	0	0	68	62	118	75	19	139
55	58	59	0	0	0	0	0	0	68	62	118	75	19	139
55	58	59	0	0	0	0	0	0	68	62	118	75	19	139
55	58	59	0	0	0	0	0	0	68	62	118	75	19	139
55	58	59	0	0	0	0	0	0	68	62	118	75	19	139
55	58	59	0	0	0	0	0	0	68	62	118	75	19	139
55	58	59	0	0	0	0	0	0	68	62	118	75	19	139
55	58	59	0	0	0	0	0	0	68	62	118	75	19	139
55	58	59	0	0	0	0	0	0	68	62	118	75	19	139
55	58	59	0	0	0	0	0	0	68	62	118	75	19	139
55	58	59	0	0	0	0	0	0	68	62	118	75	19	139
28	29	47	0	0	0	0	0	0	68	20	47	154	79	150
55	58	59	0	0	0	0	0	0	68	24	120	218	227	270
55	58	59	0	0	0	0	0	0	68	39	95	204	203	242
55	58	59	0	0	0	0	0	0	68	39	95	204	203	242
55	58	59	0	0	0	0	0	0	68	24	120	218	227	270
11	12	12	0	0	0	0	0	0	68	66	63	-36	-76	9
0	0	0	0	0	0	0	0	0	69	3	30	-8	-28	32
55	58	59	0	0	0	0	0	0	69	37	115	200	233	242
55	58	59	0	0	0	0	0	0	69	37	115	200	233	242
17	17	18	0	0	0	0	0	0	69	64	134	52	85	149
-11	-12	-12	0	0	0	0	0	0	70	39	25	42	-17	33
-11	-12	-12	0	0	0	0	0	0	70	39	25	42	-17	33
55	58	59	0	0	0	0	0	0	70	38	123	317	353	347
55	58	59	0	0	0	0	0	0	70	8	79	225	238	270

11) Maps showing suitability maps for each land use such as open space, single-family, multi-family, low density commercial, high density commercial, and industrial.

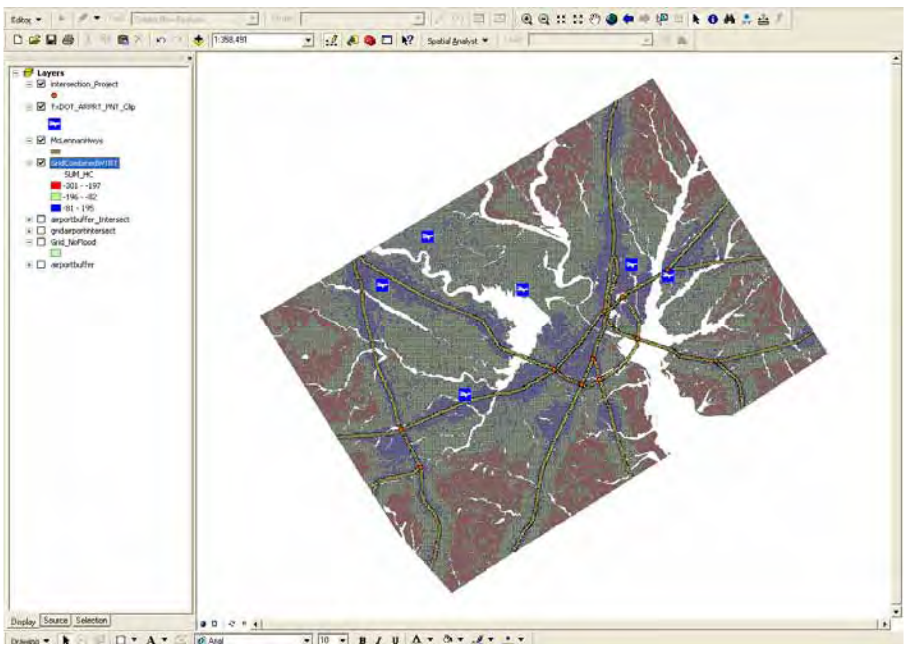



Layers
Intersection_Project
TxDOT_ARPRT_PNT_Clip
McLennanHwys
GridCombinedWTRT
SUM_SF
60 - 159
160 - 201
202 - 270
airportbuffer_Intersect
gridairportintersect
Grid_NoFlood
airportbuffer
Spatial Analyst
1:358,491
Display
Source
Selection

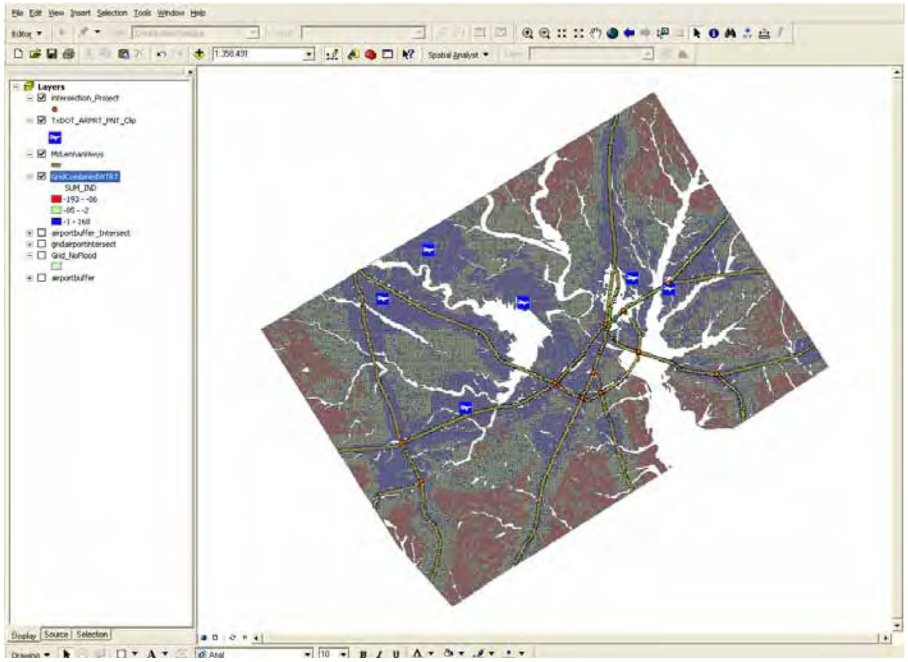

Layers
Intersection_Project
TxDOT_ARPRT_PNT_Clip
McLennanHwys
GridCombinedWTRT
SUM_MF
17 - 116
117 - 164
165 - 273
airportbuffer_Intersect
gridairportintersect
Grid_NoFlood
airportbuffer
Spatial Analyst
1:358,491
Display
Source
Selection



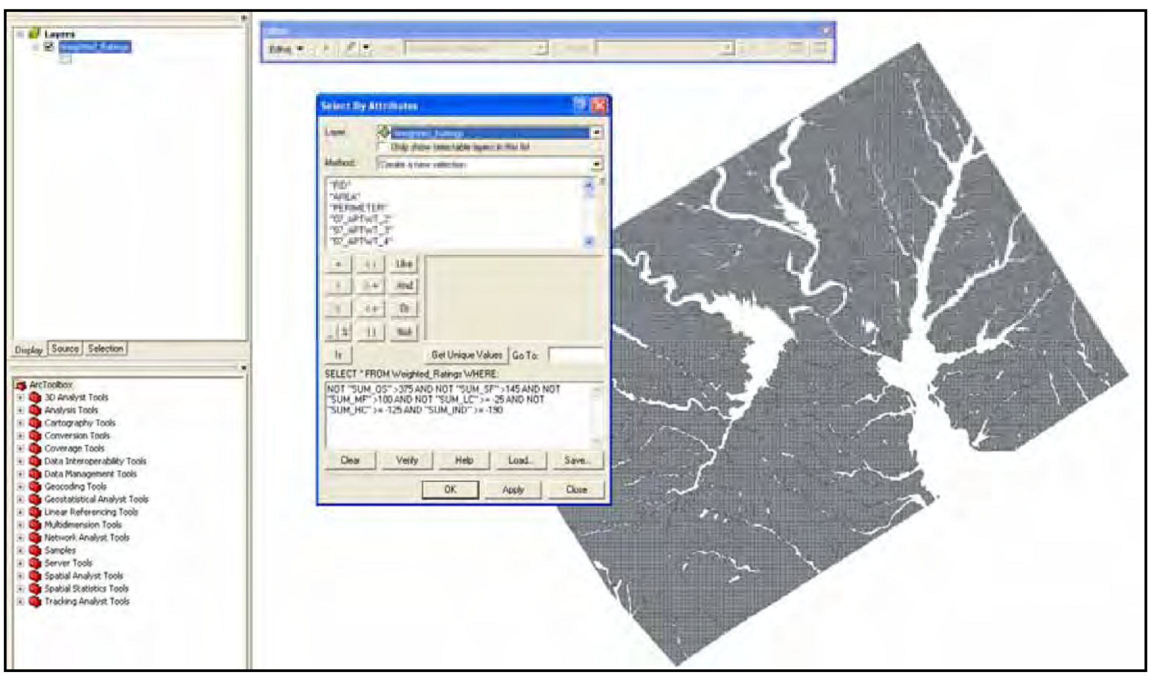
File Edit View Insert Selection Tools Window Help
Editor
Spatial Analyst
Layers
intersection_Project
TxDOT_ARPRT_PNT_Clip
McLennanHwys
SUM_LC
-231 - -125
-124 - -41
-40 - 161
airportbuffer_Intersect
gridairportintersect
Grid_NoFlood
airportbuffer
Display
Source
Selection



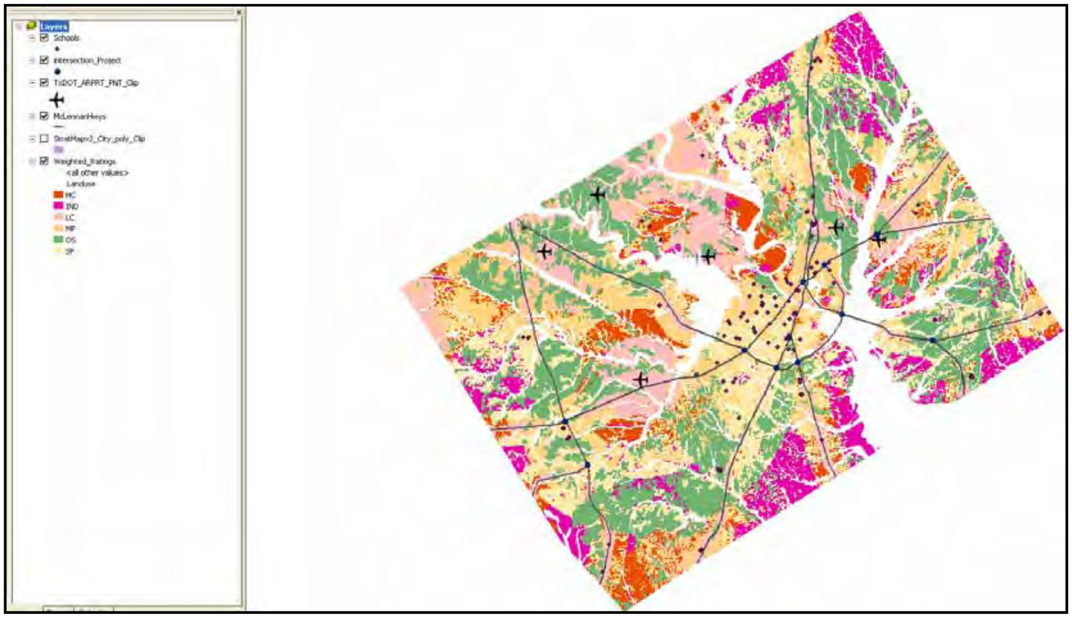
Editor
Spatial Analyst
Layers
intersection_Project
TxDOT_ARPRT_PNT_Clip
McLennanHwys
SUM_HC
-301 - -197
-196 - -82
-81 - 195
airportbuffer_Intersect
gridairportintersect
Grid_NoFlood
airportbuffer
Display
Source
Selection



12) Rule-based combination method was used in order to develop a composite suitability map. 'Select by attributes' command was used to assign the different land uses.



Open space—27%, Single family—22%, Multi-family—20%, Low commercial—15%, High commercial—7%, and Industrial—9%

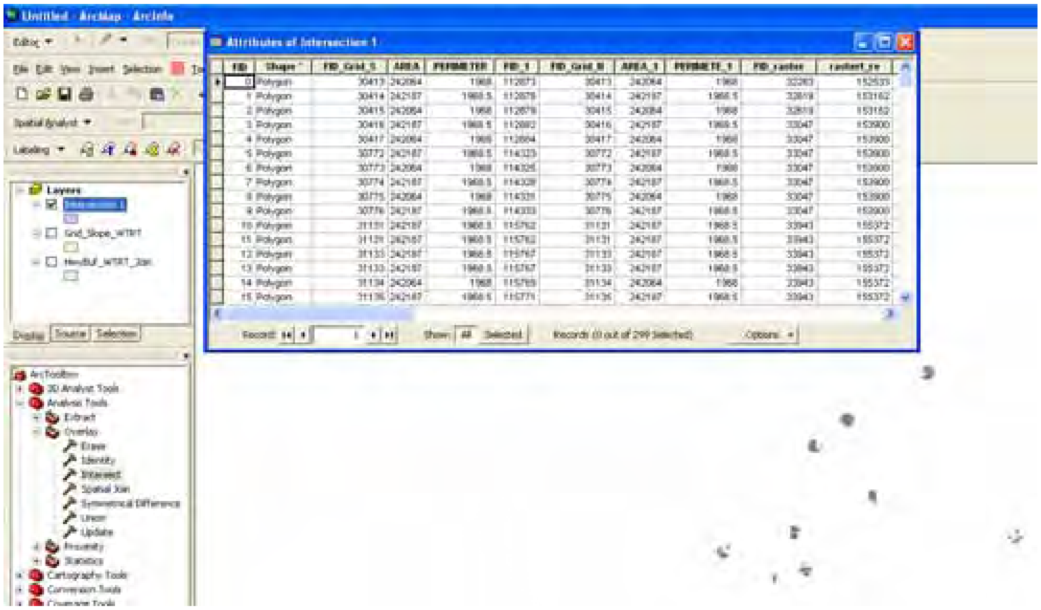


1. Residential Medium (RM)

Data sets considered:

a) Access to Highways—between 600 and 1,500 meters
b) Slope—between 0% and 3%
c) Distances to Urban Centers—between 1,000 and 2,000 meters

The total number of cells selected after performing the query command is 297.
Anticipated total growth is 6,380 households.
Average lot size is 0.25 acres.
The size of a single grid is 5.5 acres.
One grid can hold 22 households (5.5/0.25 = 22 HH).
The total number of grids required for allocating 6,380 HH is 6,380/ 22 = 290 grids.

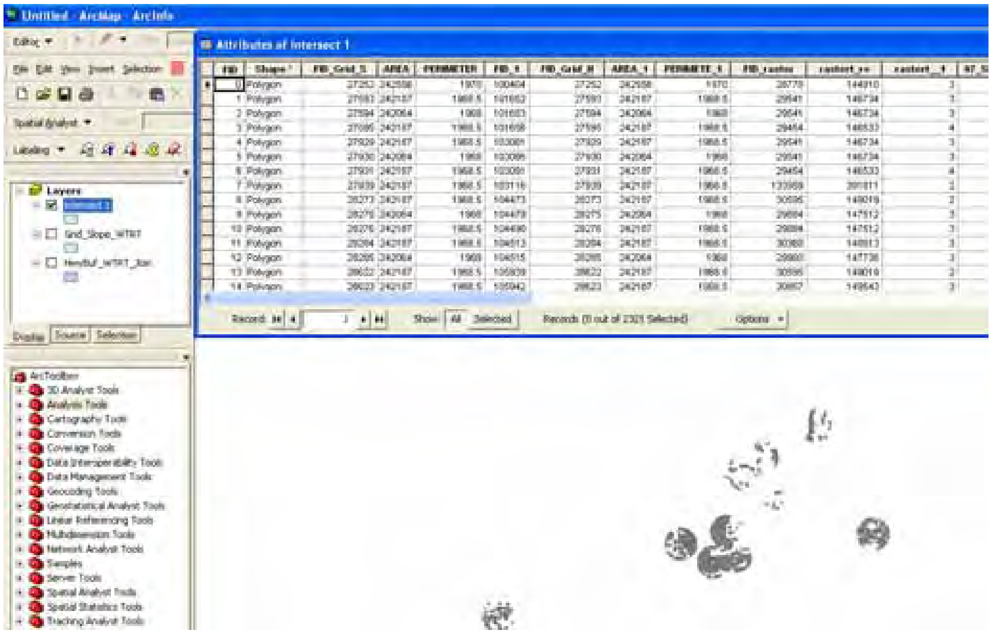


2) Residential Low (RL)

Data sets considered:

a) Access to Highways—greater than 1,500 meters
b) Slope—between 3% and 9%
c) Distances to Urban Centers—between 2,000 and 10,000 meters.

The total number of cells selected after performing the query command is 2,321.
Anticipated total growth is 1,250 households.
Average lot size is 5 acres.
The size of a single cell is 5.5 acres.
One grid can hold 1.1 household (5.5/5 = 1.1 HH).
The total number of cells required for allocating 1,250 HHs is 1,250/1.1 = 1,136 cells.

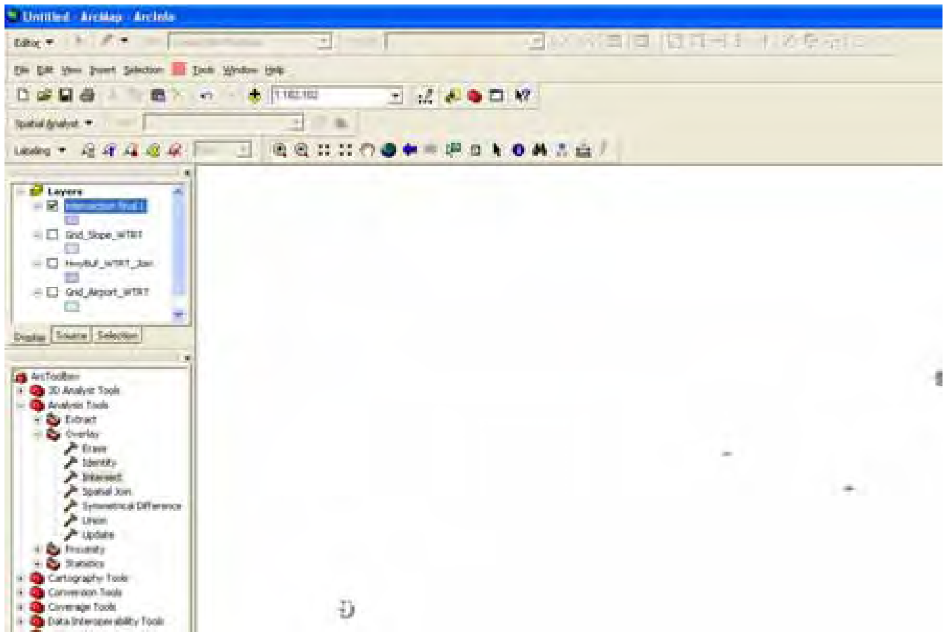


3) Low Commercial (LC)

Data sets considered:

a) Access to Highways—between 600 and 1500 meters
b) Distances to Major Airports—between 7 and 11 miles
c) Slope—between 0% and 3%
d) Distances to Urban Centers—between 500 and 1,000 meters.

The total number of cells selected after performing the query command is 63.
Anticipated total growth is 6,105 employees.
Average sq. footage is 300, which is 0.0069 acres.
The size of a single grid is 5.5 acres.
One grid can hold 797 employees (5.5/0.0069 = 797 employees).
The total number of grids required for allocating 6,105 employees is 6,105/797 = 8 grids.

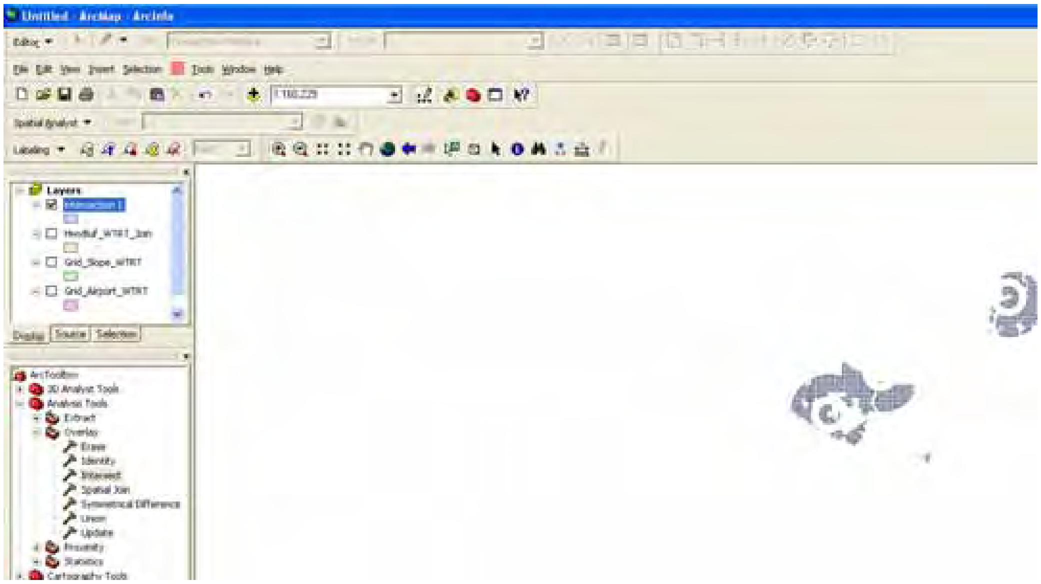


4) Industrial (IND)

Data sets considered:

a) Access to Highways—greater than 1500 meters
b) Railway Line—1,000-ft buffer
c) Distances to Major Airports—between 1 and 5 miles
d) Slope—0%.

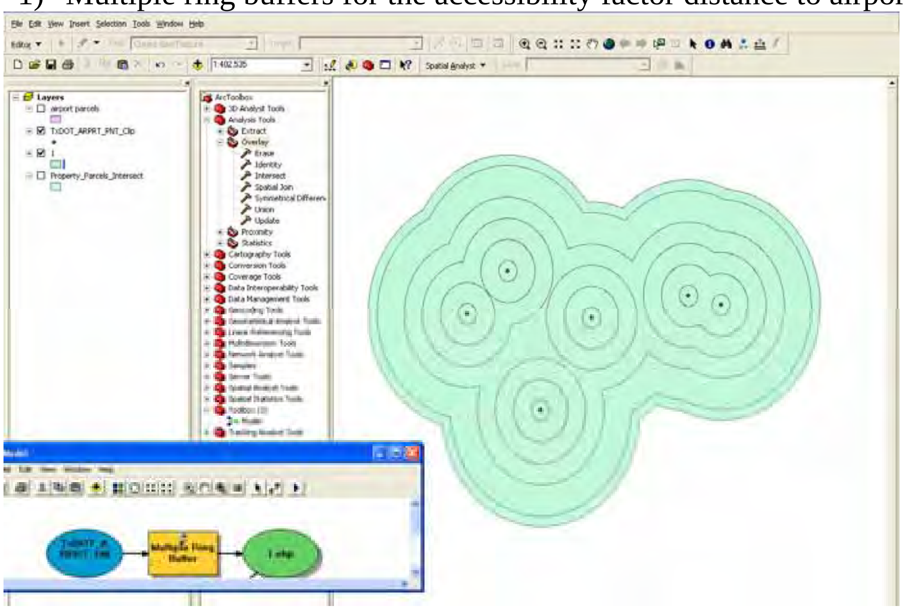
The total number of grids selected after performing the query command is 580.
Anticipated total growth is 4,508 employees.
Average sq. footage is 500, which is 0.0115 acres.
The size of a single grid is 5.5 acres.
One grid can hold 478 employees (5.5/0.0115 = 478 employees).
The total number of grids required for allocating 478 employees is 4,508/478 = 10 grids.



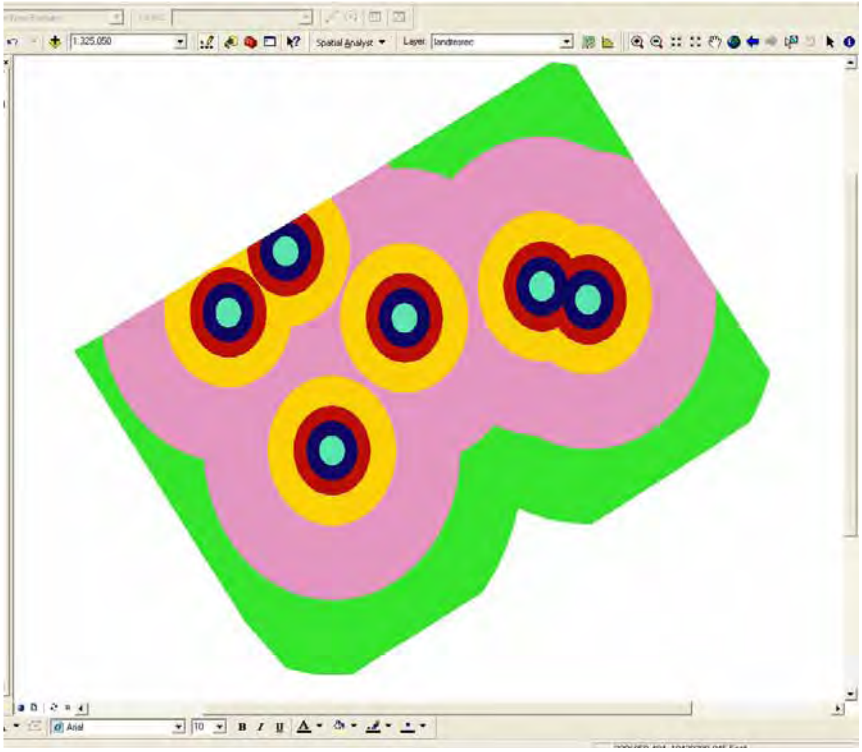
Map showing the multiple ring buffers for the factor- distances to airports.

ArcView Spatial Analyst

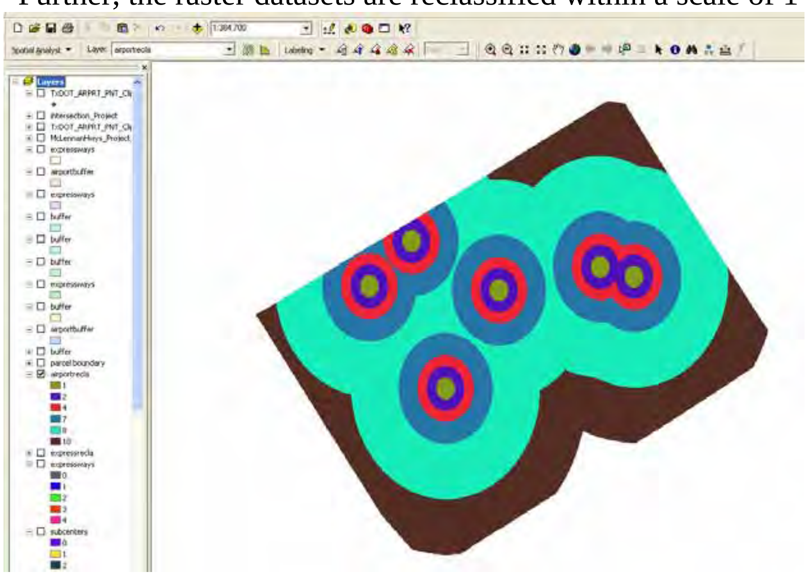
1) Multiple ring buffers for the accessibility factor distance to airports



2) After producing the multiple ring buffers for the distance related factors, the shapefiles are converted to raster datasets to perform further analysis.



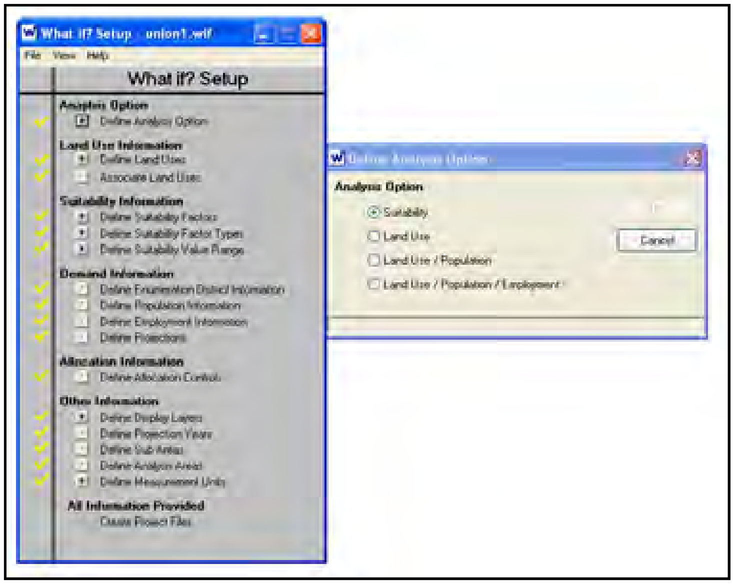
3) Further, the raster datasets are reclassified within a scale of 1 to 10.



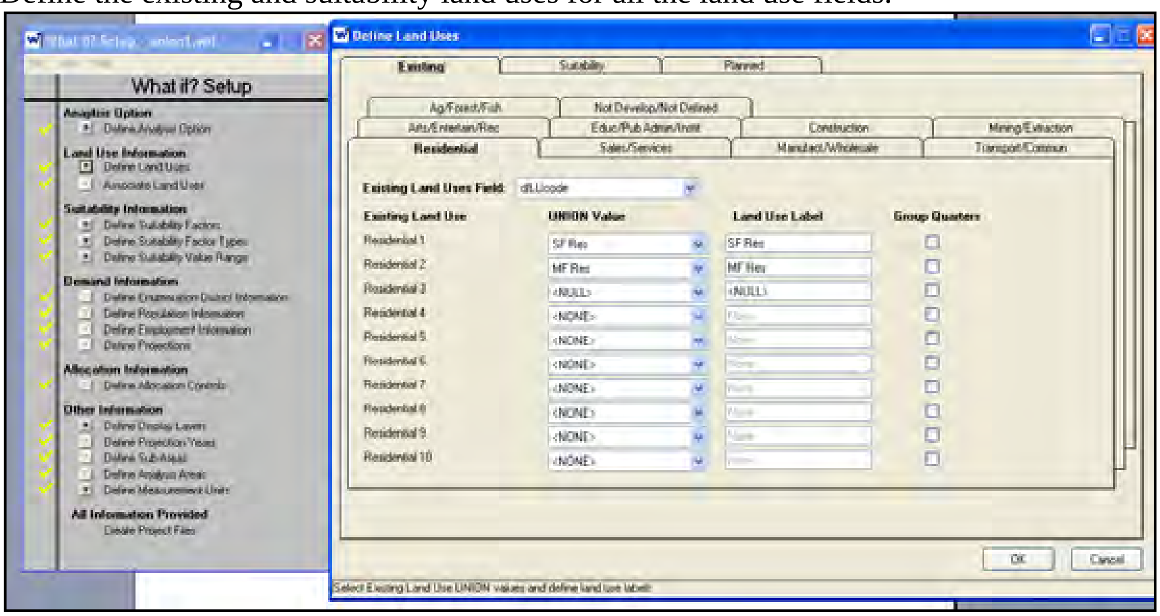
The process of land suitability analysis in *What if?*-

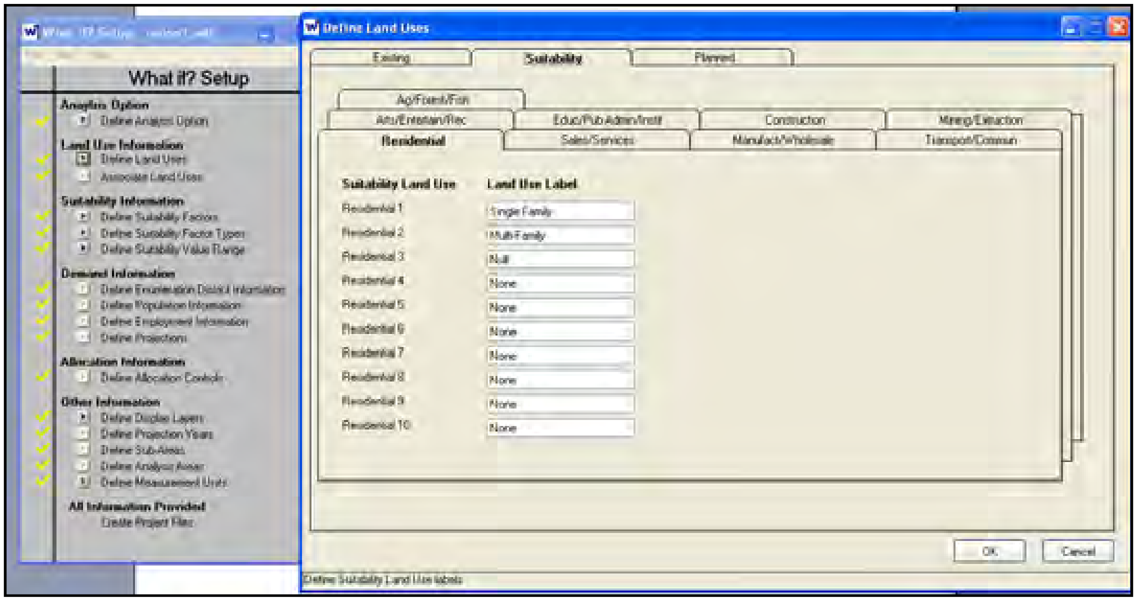
After importing the UNION file from ArcGIS to *What if?* Setup option, the steps to be followed are:

1) Define the Analysis Option- Suitability option.

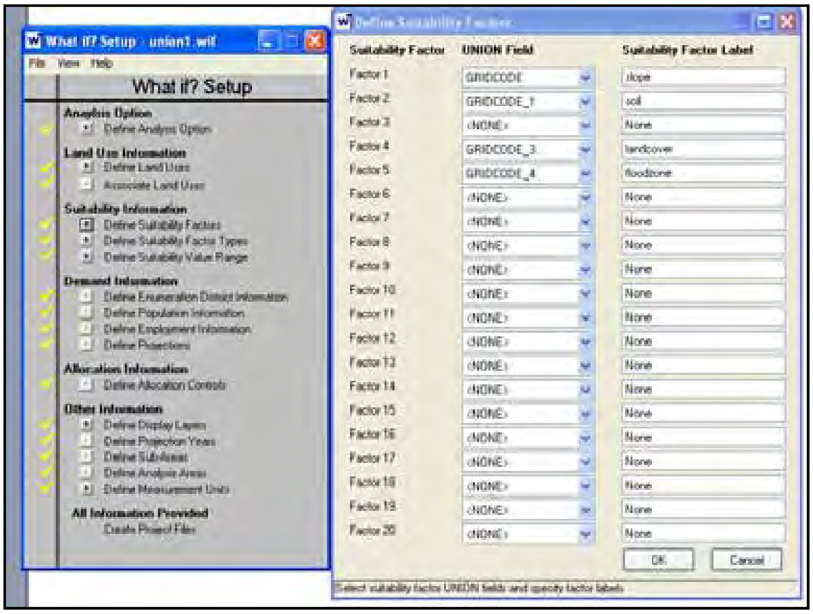


2) Define the existing and suitability land uses for all the land use fields.

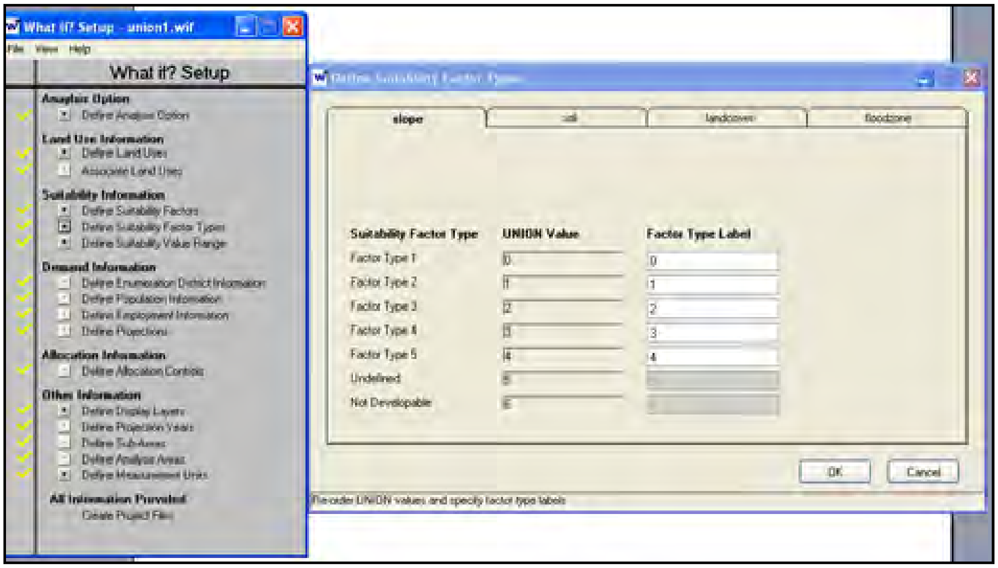




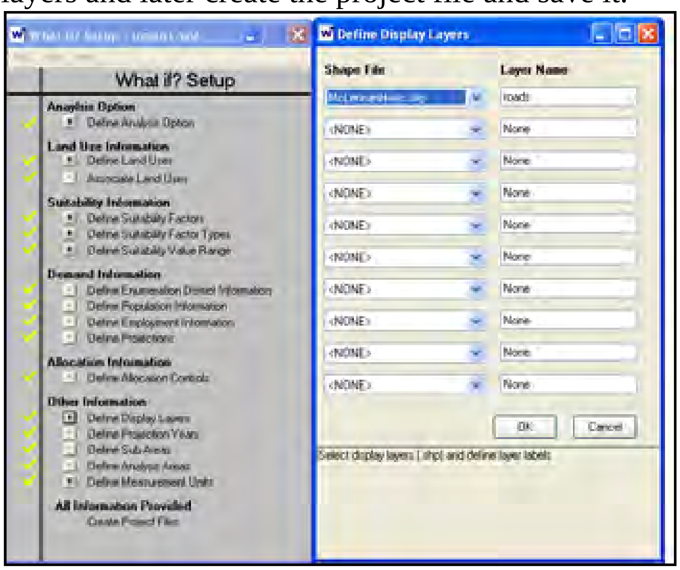
3)Define the Suitability Factors.

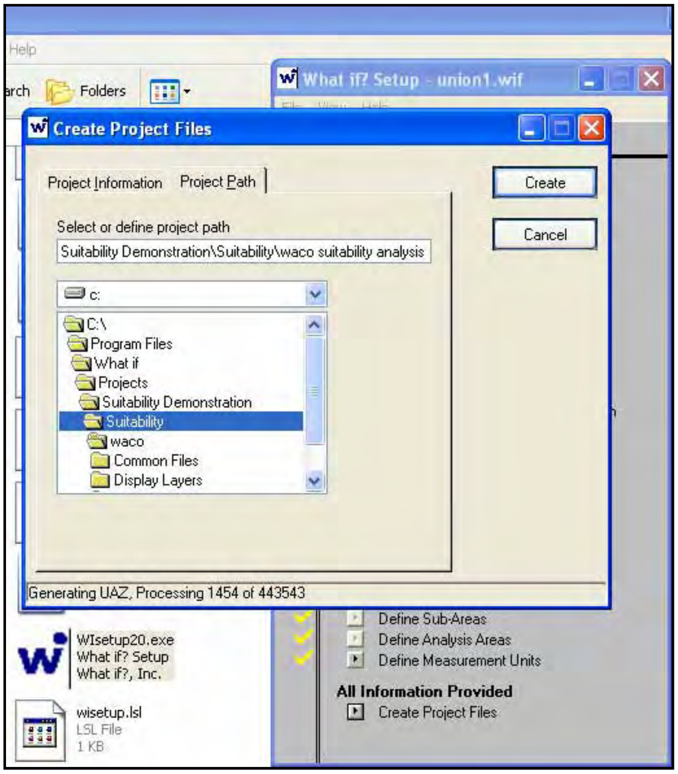


4) Define the Suitability Factor Types.

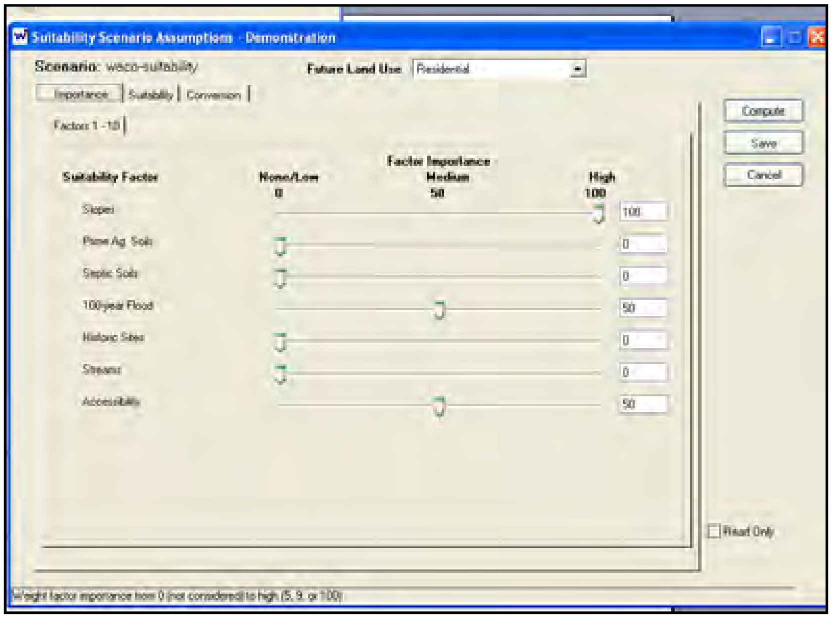


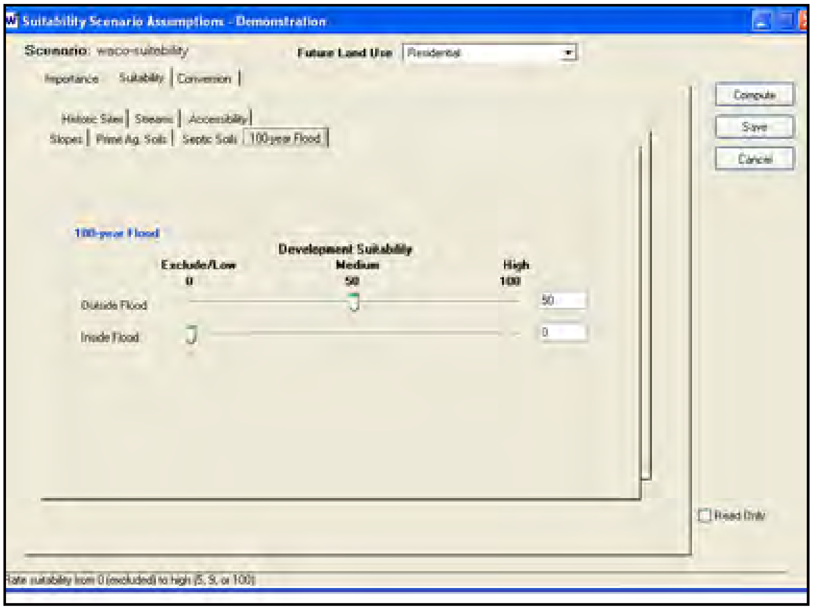
5) Define the Display layers and later create the project file and save it.



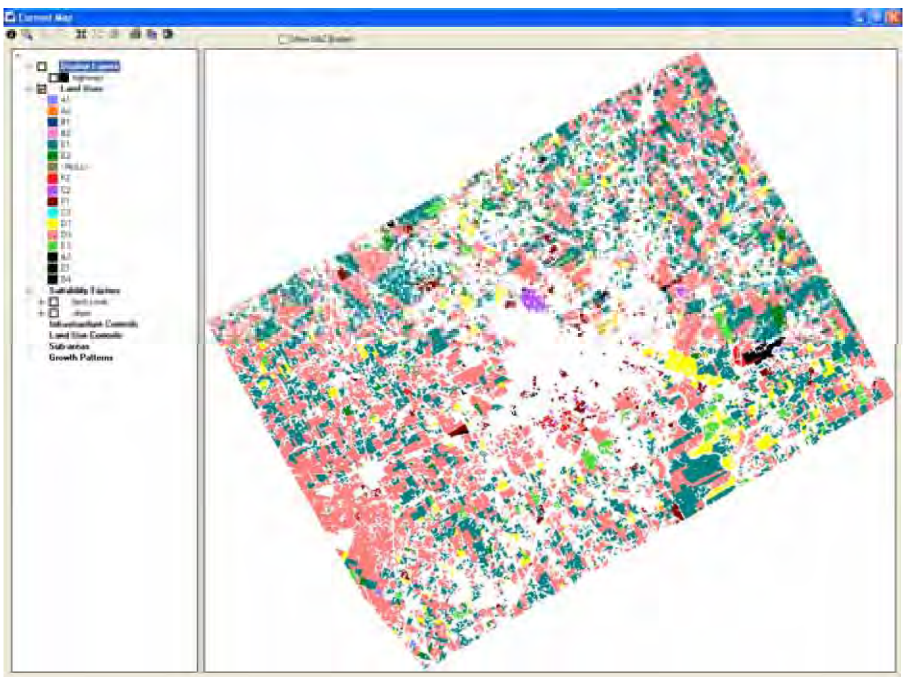


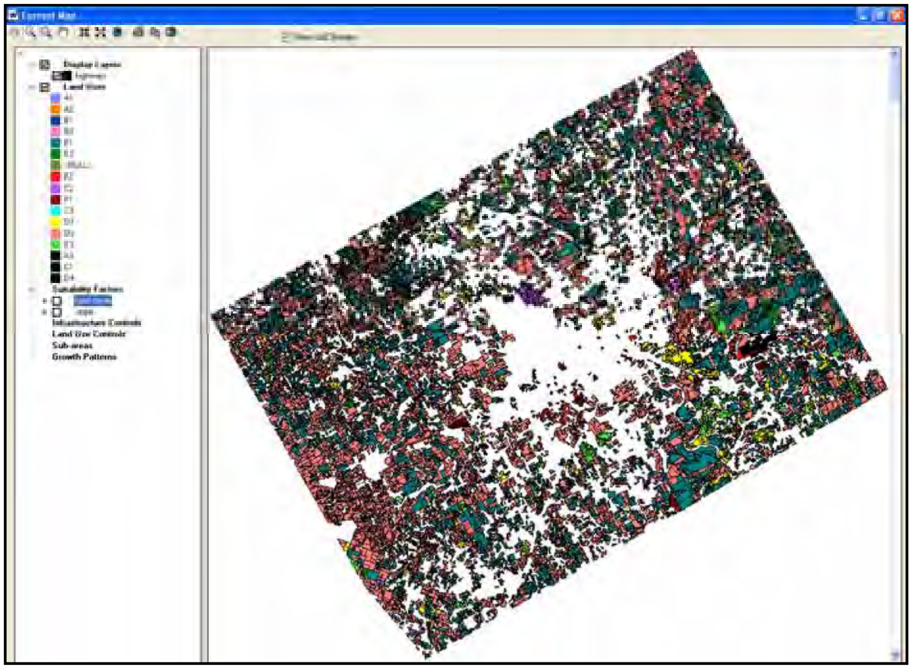
6) Open the saved project file in *What if?* Project option. Select the suitability factors and modify the values if needed.

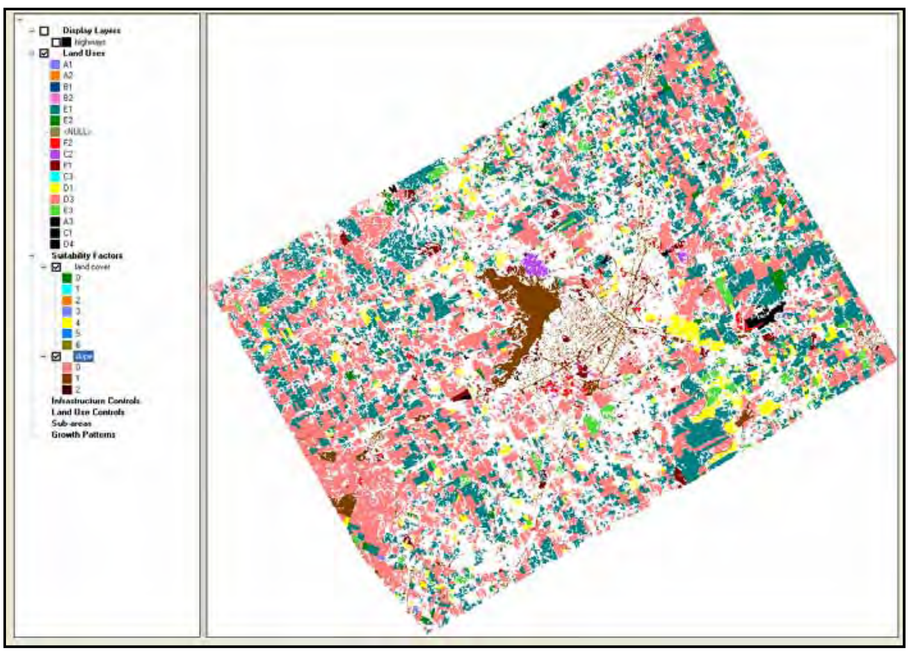




7) Maps showing the existing land use for the study area.

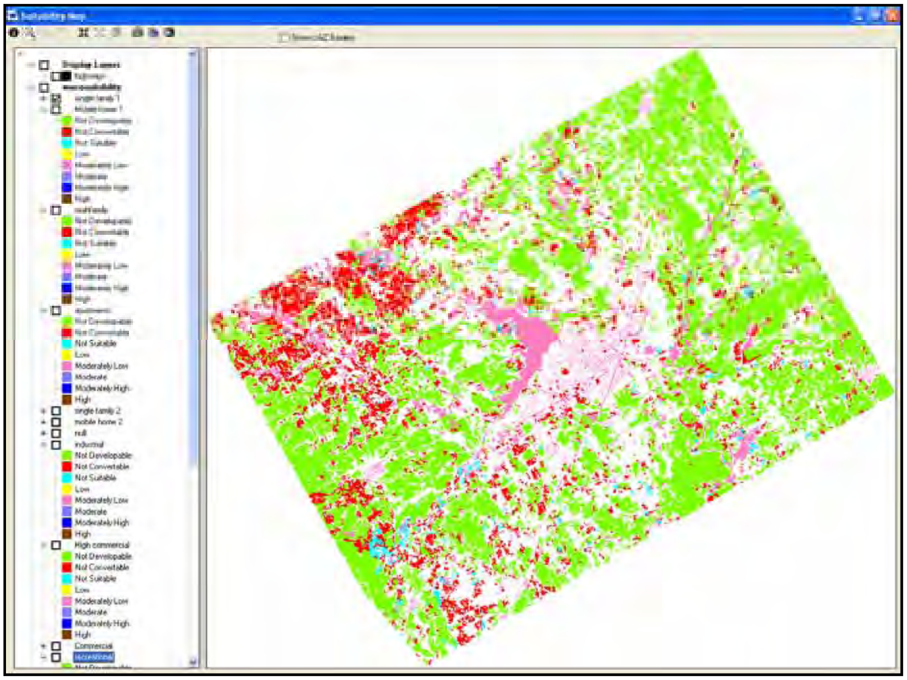




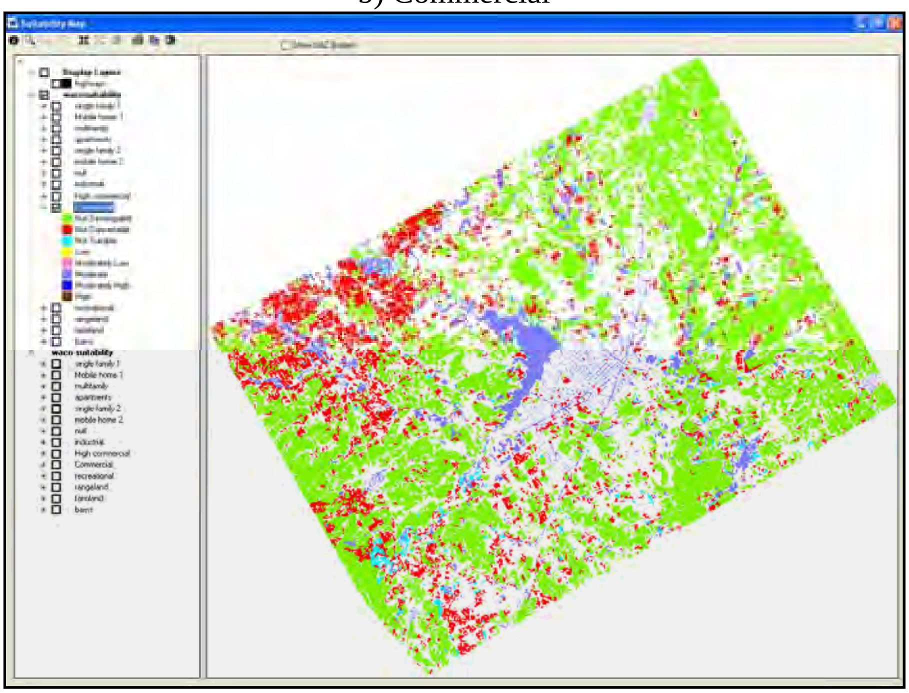


8) Suitability maps for each individual land use.

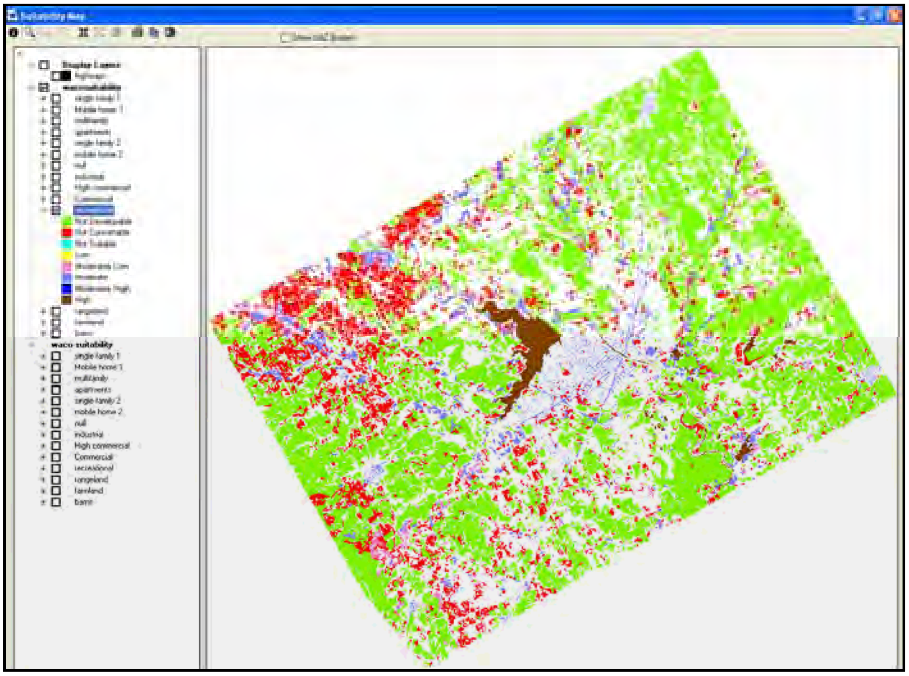
a) Single-family



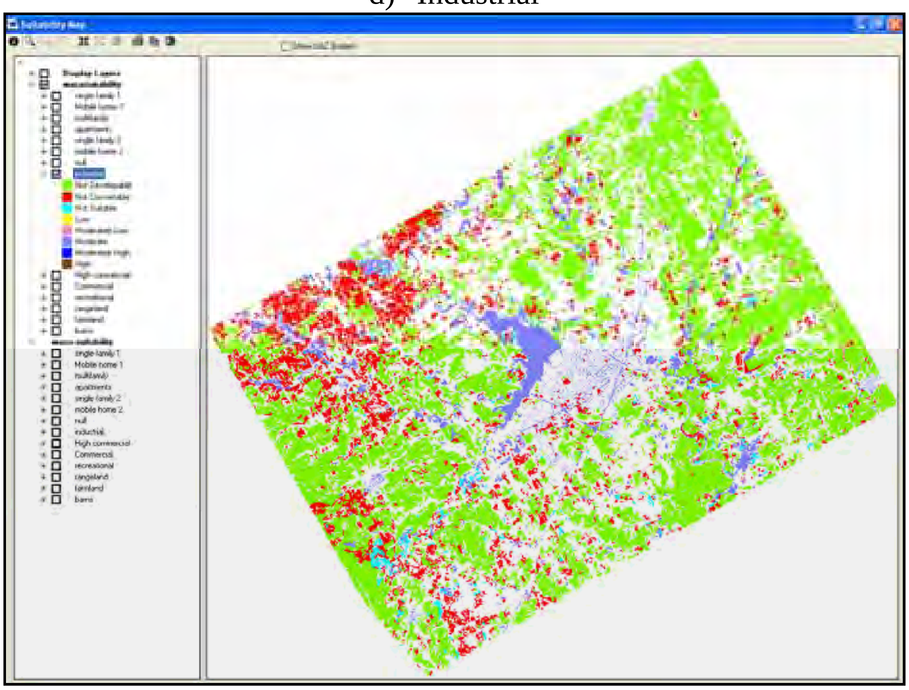
b) Commercial



c) Recreational



d) Industrial



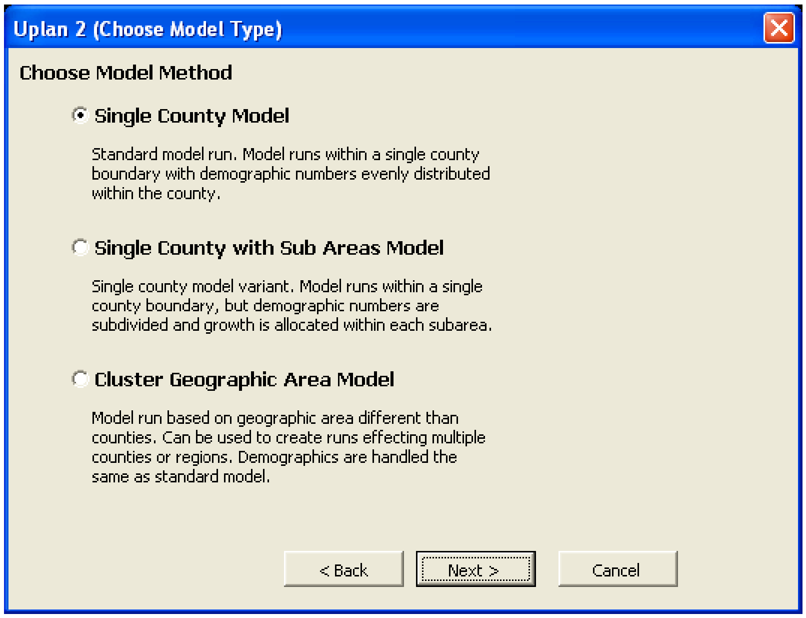
The process of land suitability analysis in UPlan-

UPlan: Data loader

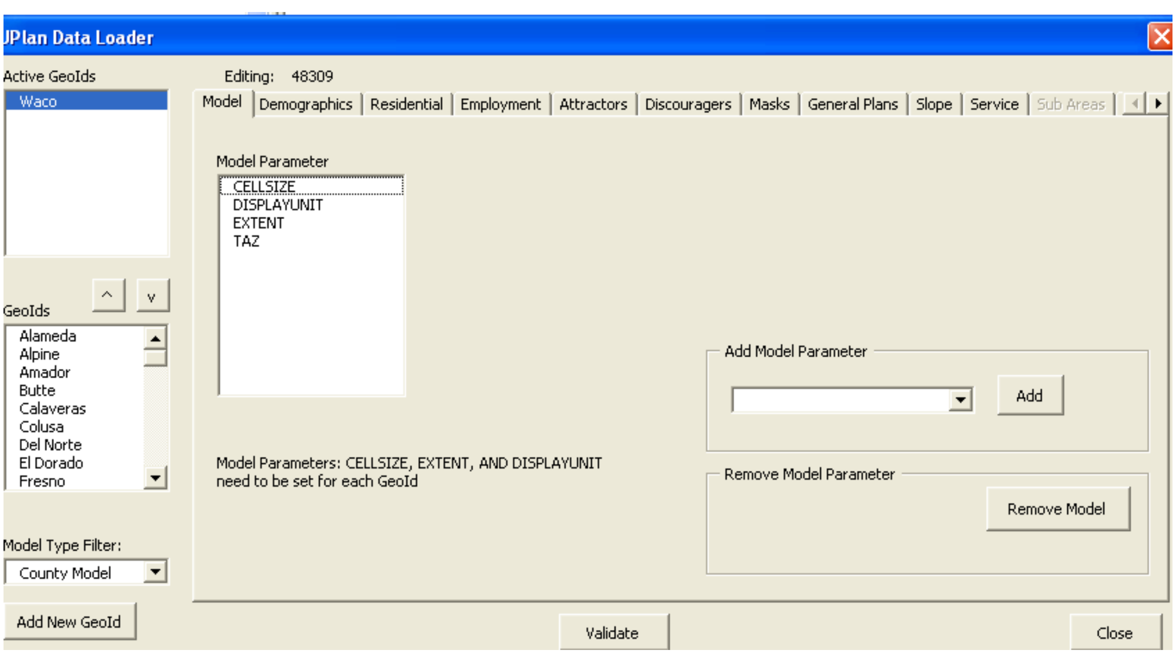
New data is uploaded into UPlan through data loader

13) Start UPlan through UPlan model interface and click next to choose sub model interface.

14) Choose model type (Single County Model, Single County with Sub Area Model, Cluster Geographic Area Model).
15) Naming the model run (any symbols can be used to run the model run, and there are no limitations in length of the run.

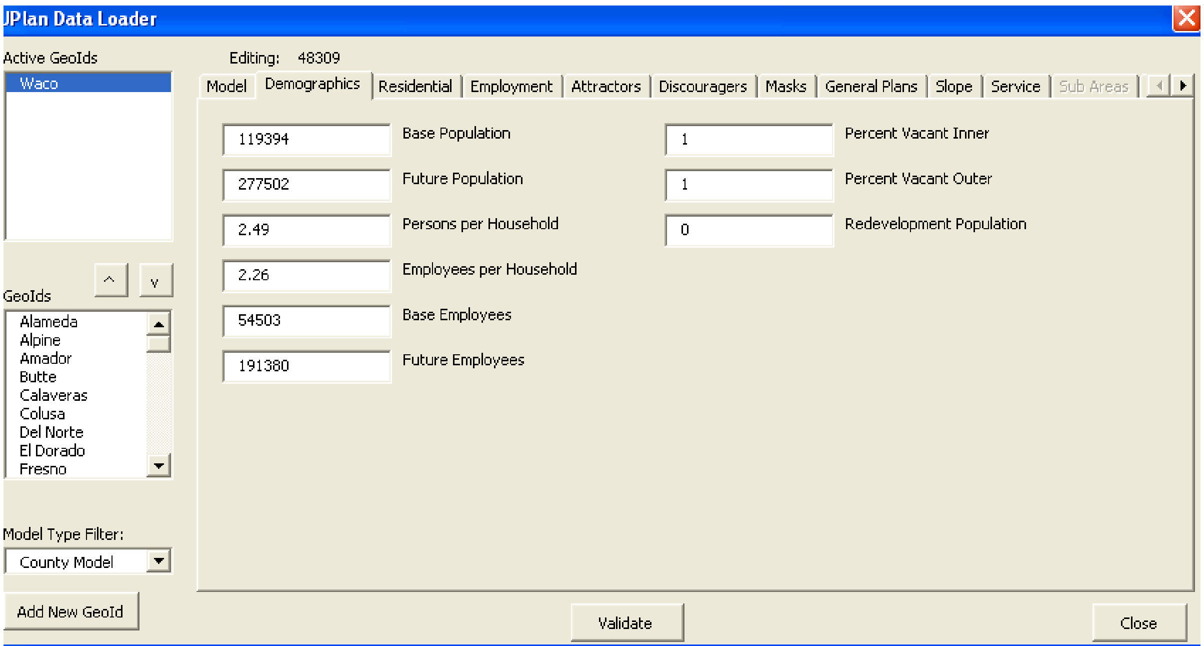


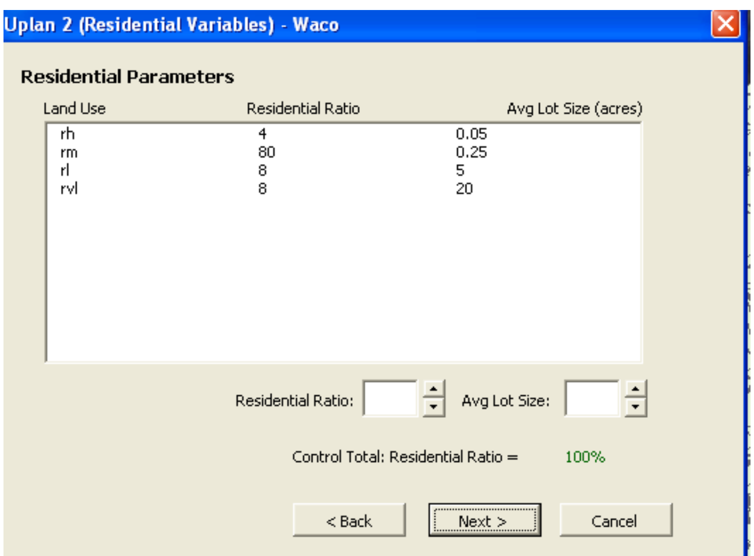
16) Selecting county and setting resolution and units.



17) Model Input data: Residential Inputs:
Residential Inputs:
%Res. High, %Res. Med., %Res. Low, % Res. Very Low—These are used to specify the proportion of households in each of the four density categories:

- High Density (Attached multifamily dwellings)
- Medium Density (Mostly detached single-family dwellings)
- Low Density (Rural dwellings with own well and septic)
- Very low density (Rural dwellings with own well and septic)
- The total should be 100%.

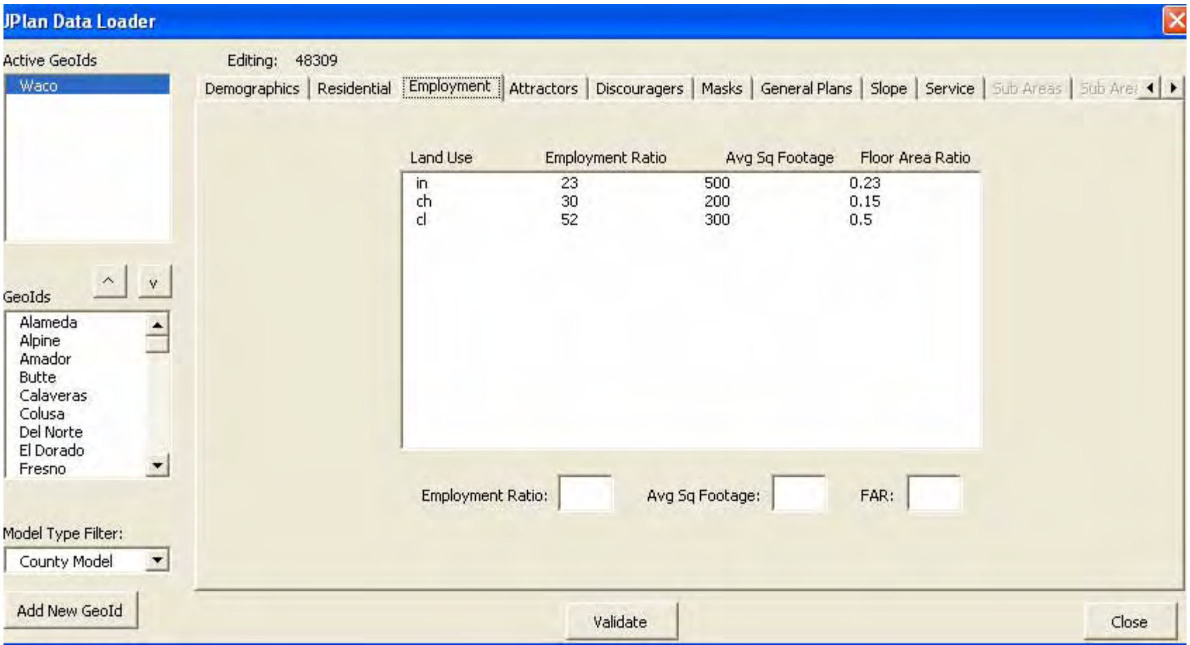




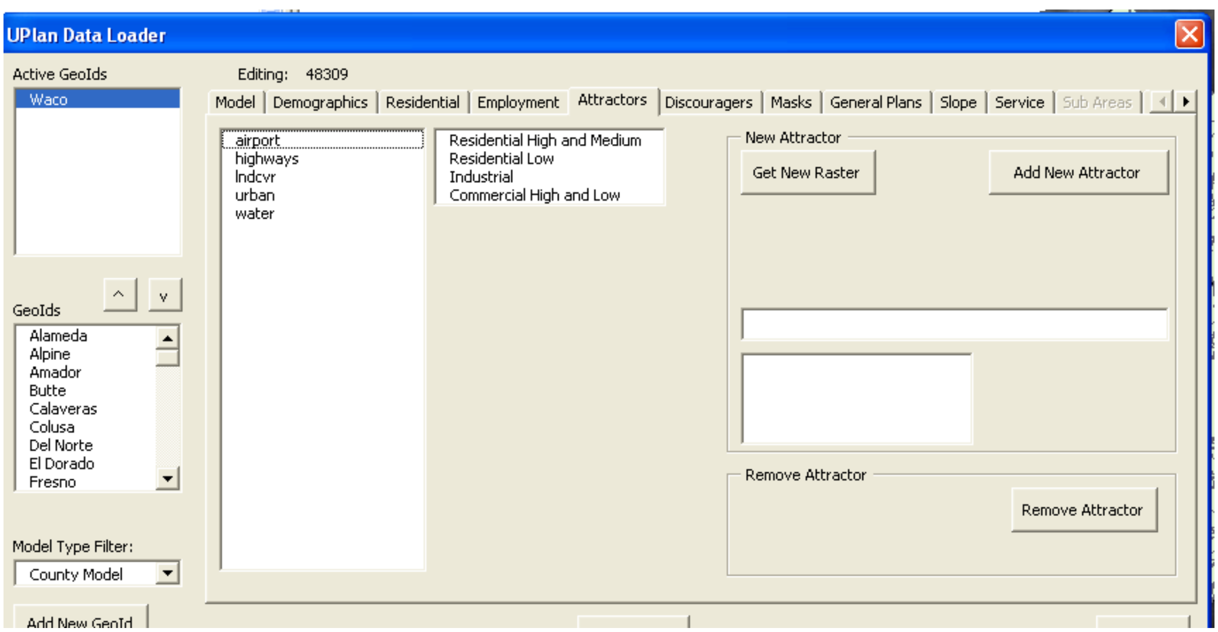
18) Employment Inputs:
% Empl in Ind, % Empl in Comm. High, % Empl in Comm. Low—These parameters are used to set the proportion of employees in each of the broad economic sectors.

- Industry, light and heavy industry
- High-density Commercial, FAR > 2.0 or defined by the user

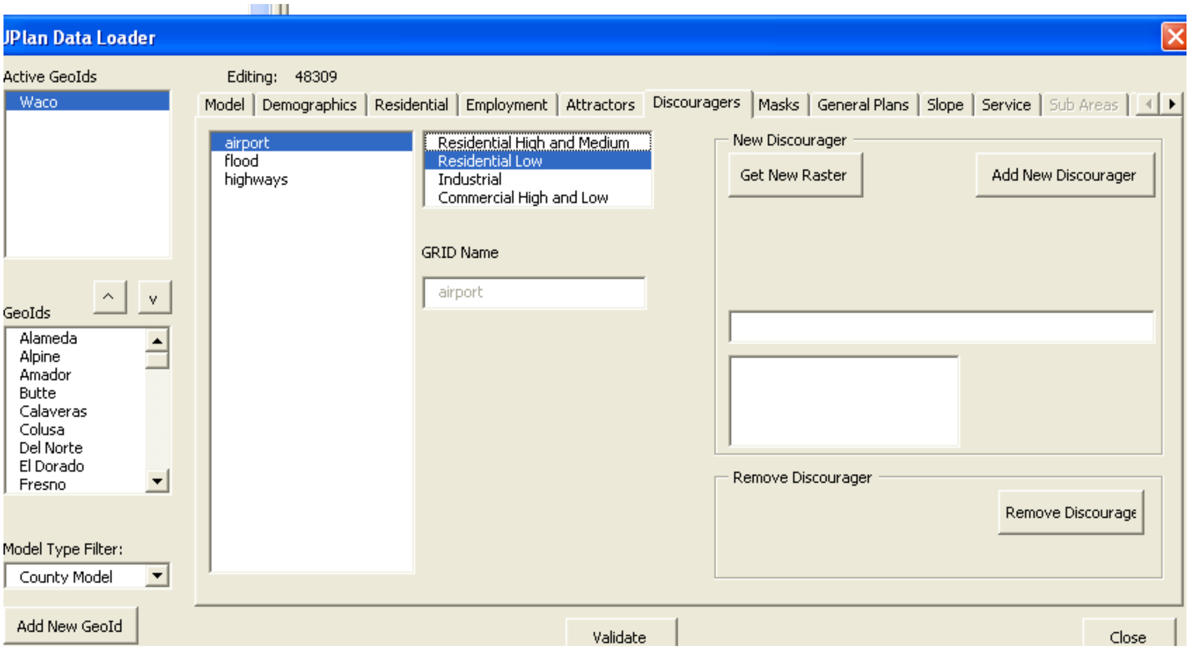
- Low-density Commercial, FAR <= 2.0 or defined by the user
- The total <= 100



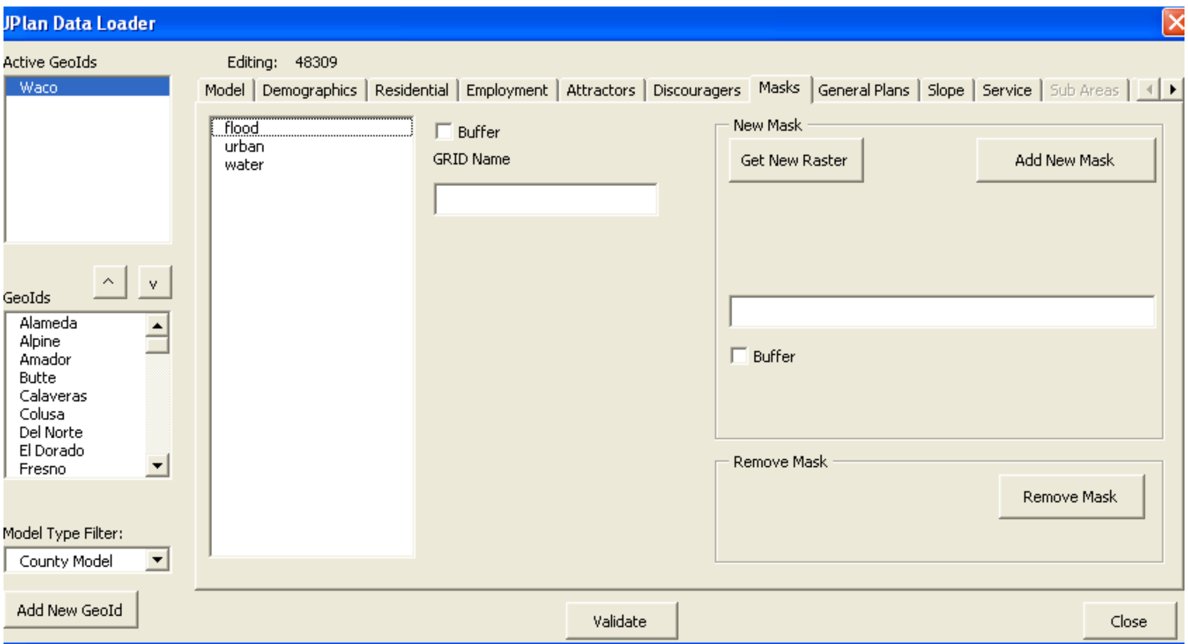
19) Attraction Grids (highways, open spaces).



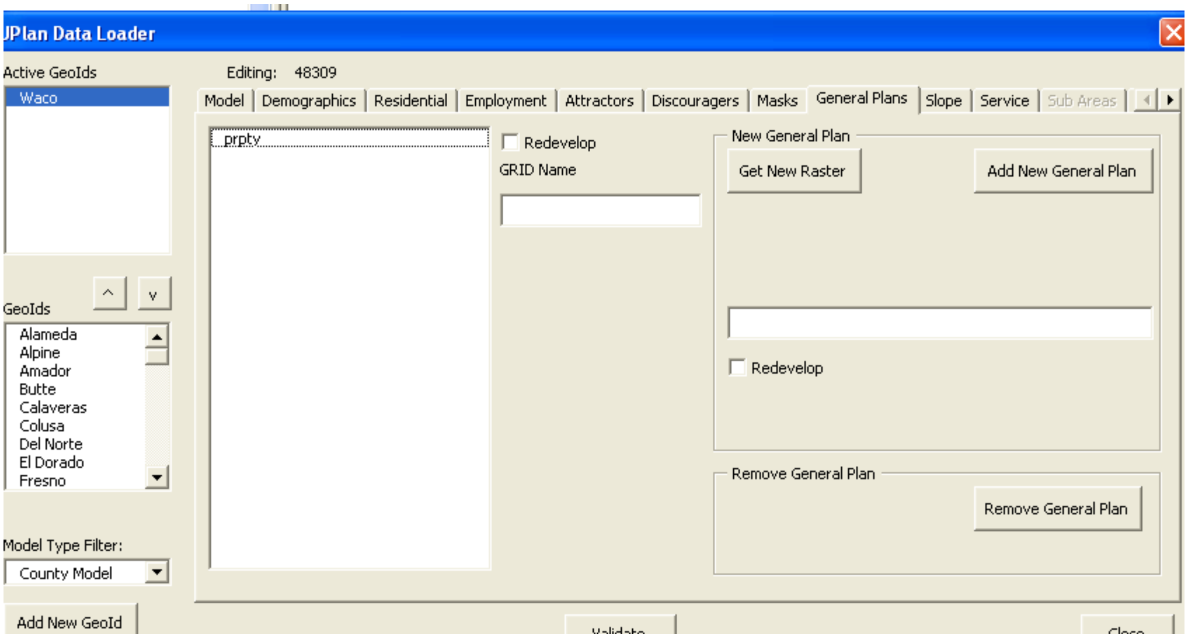
20) Discouragement features



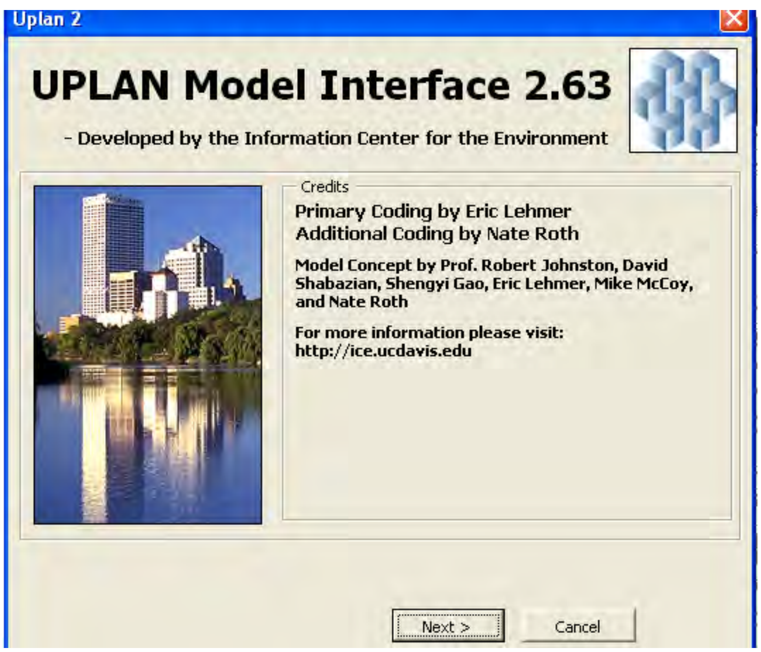
21) Mask (Lakes and streams) are not developable in terms of employment and housing.



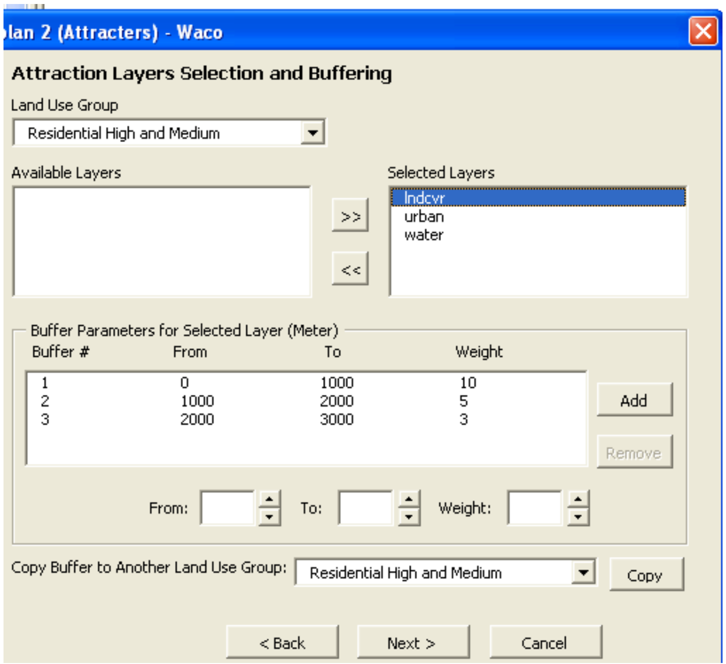
22) General plans and farmland are reclassified in UPlan model and are represented by codes.(for example, Unclassified-0, agriculture-1, industry-2).



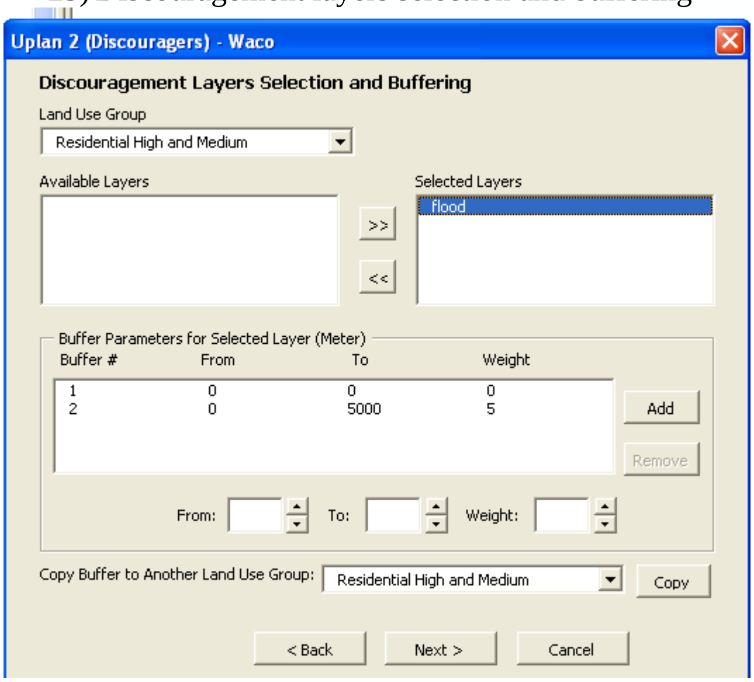
23) Run model



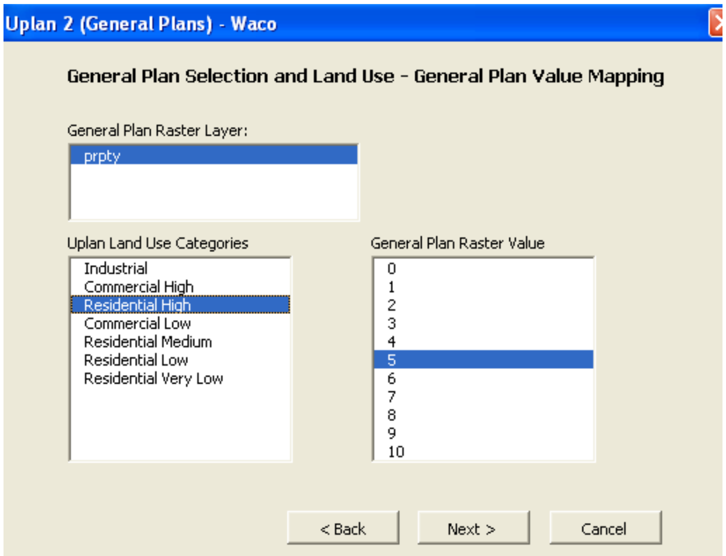
24) Attraction layers selection and buffering



25) Discouragement layers selection and buffering



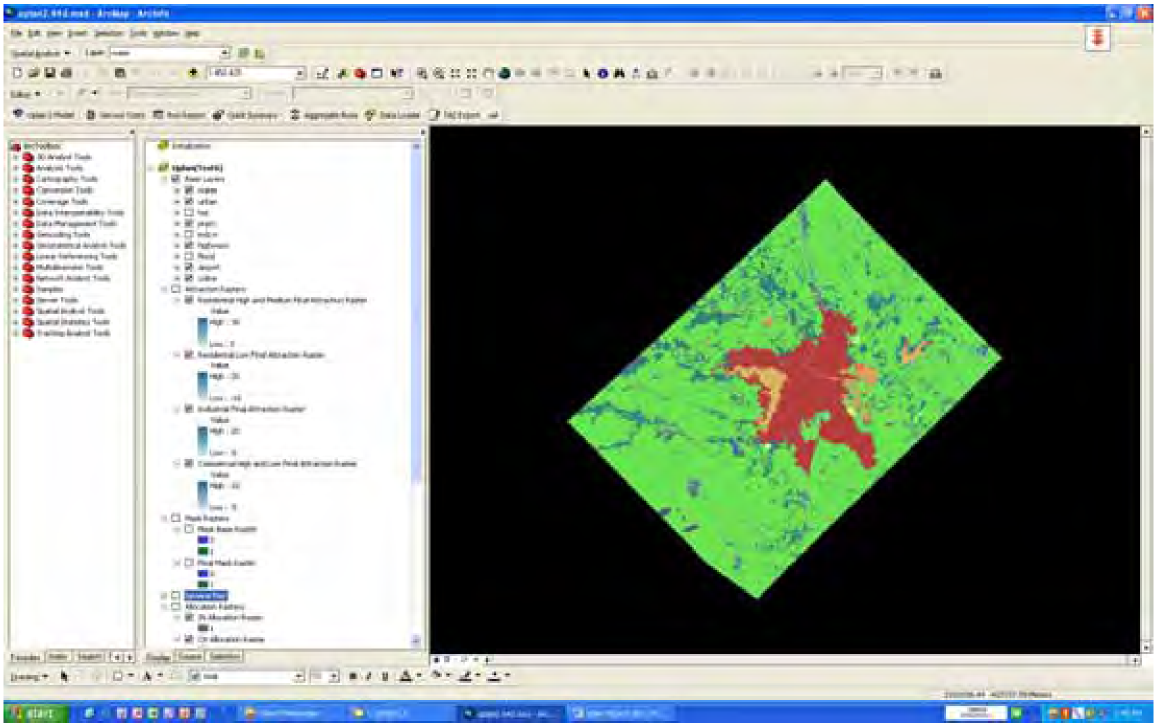
26) General plan selection and land use



27) Mask selection

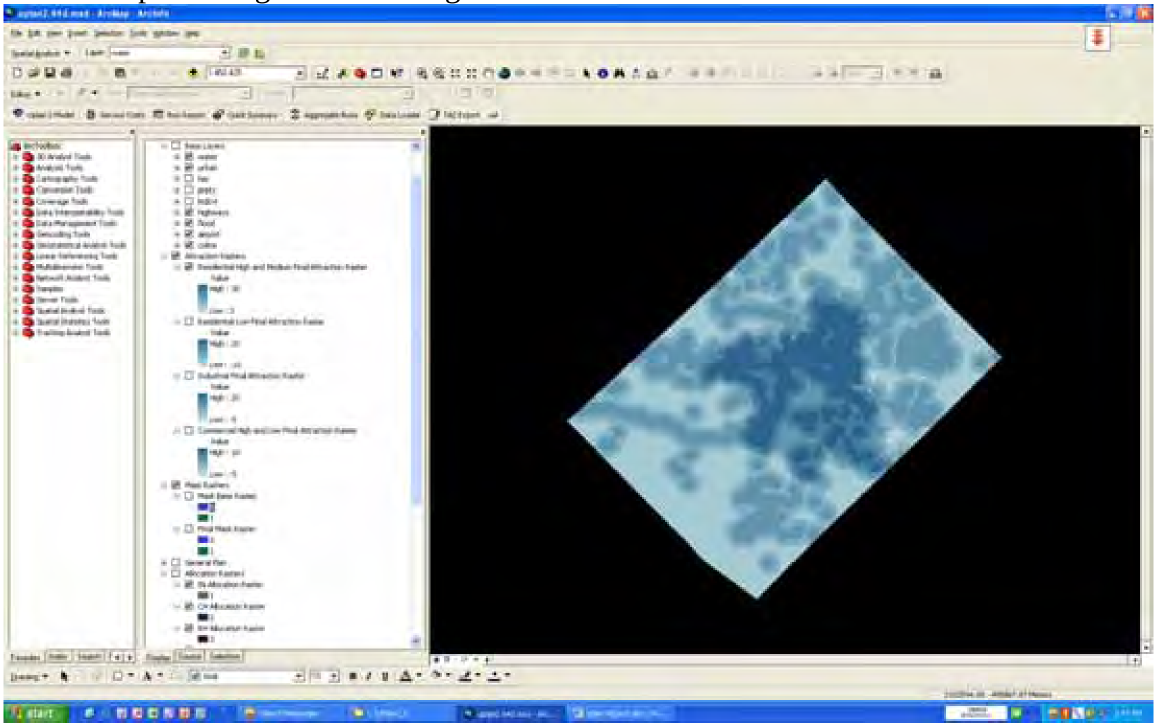
28) Run model

17) Ouput maps

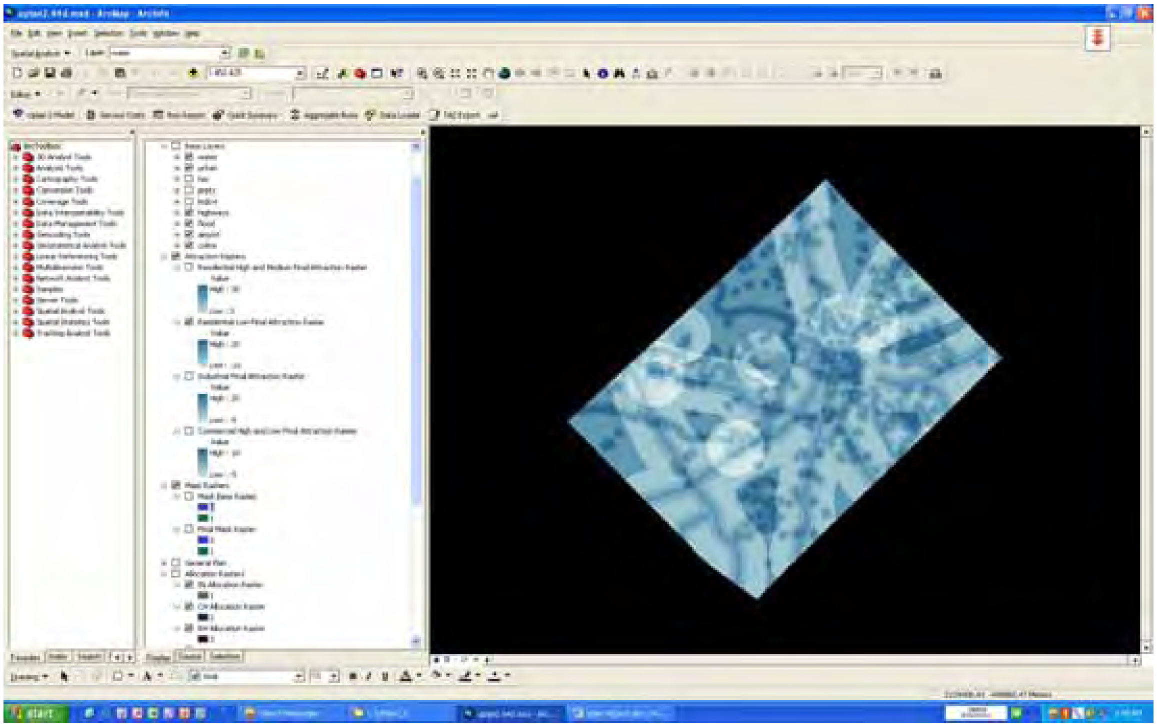


18) Allocation Raster

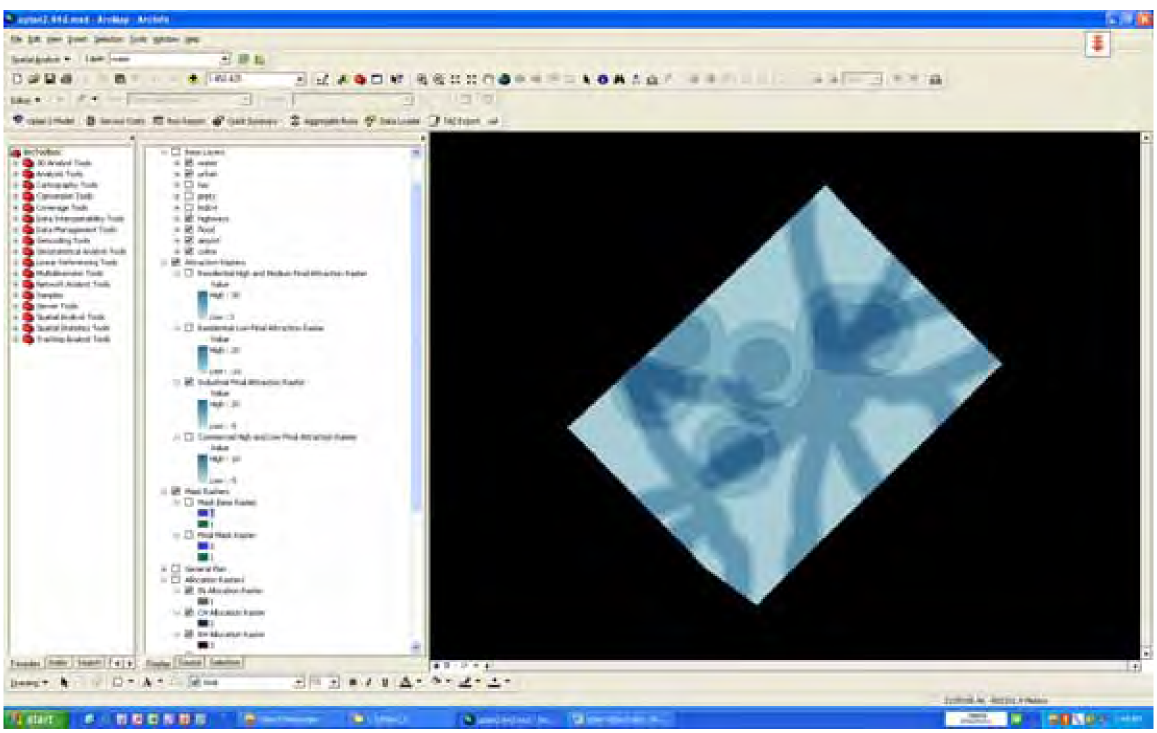
Map showing residential high and medium final raster



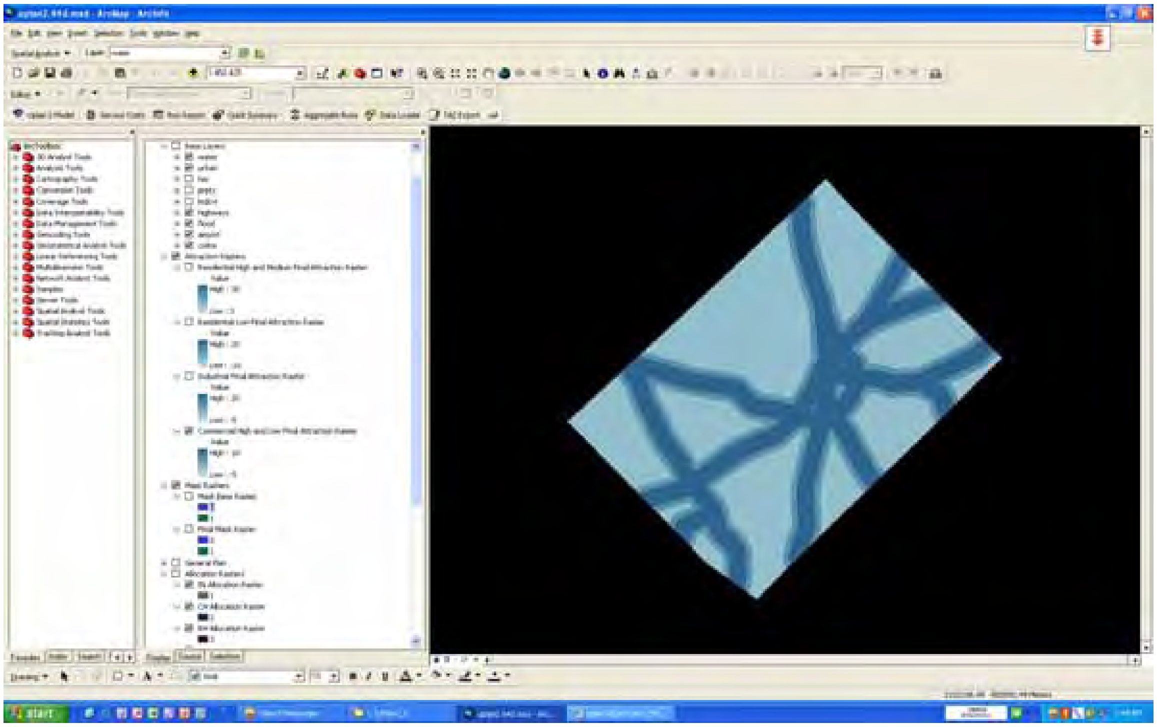
19) Map showing residential low final raster



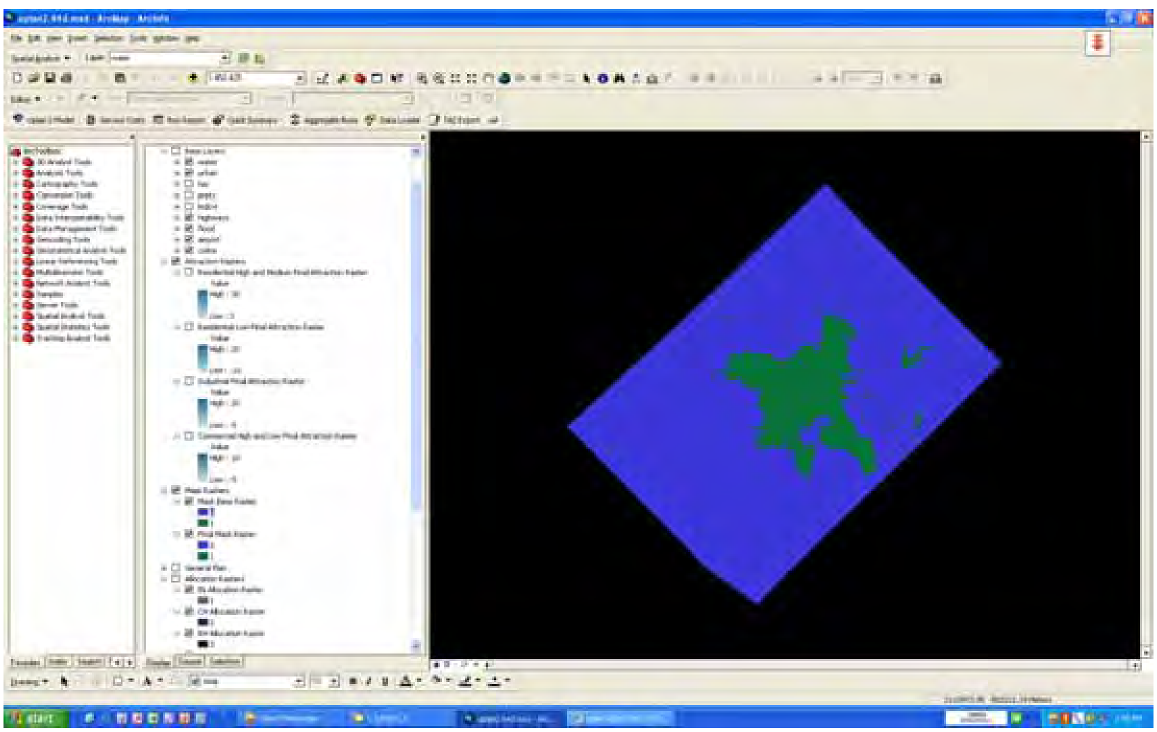
20) Map showing industrial final raster



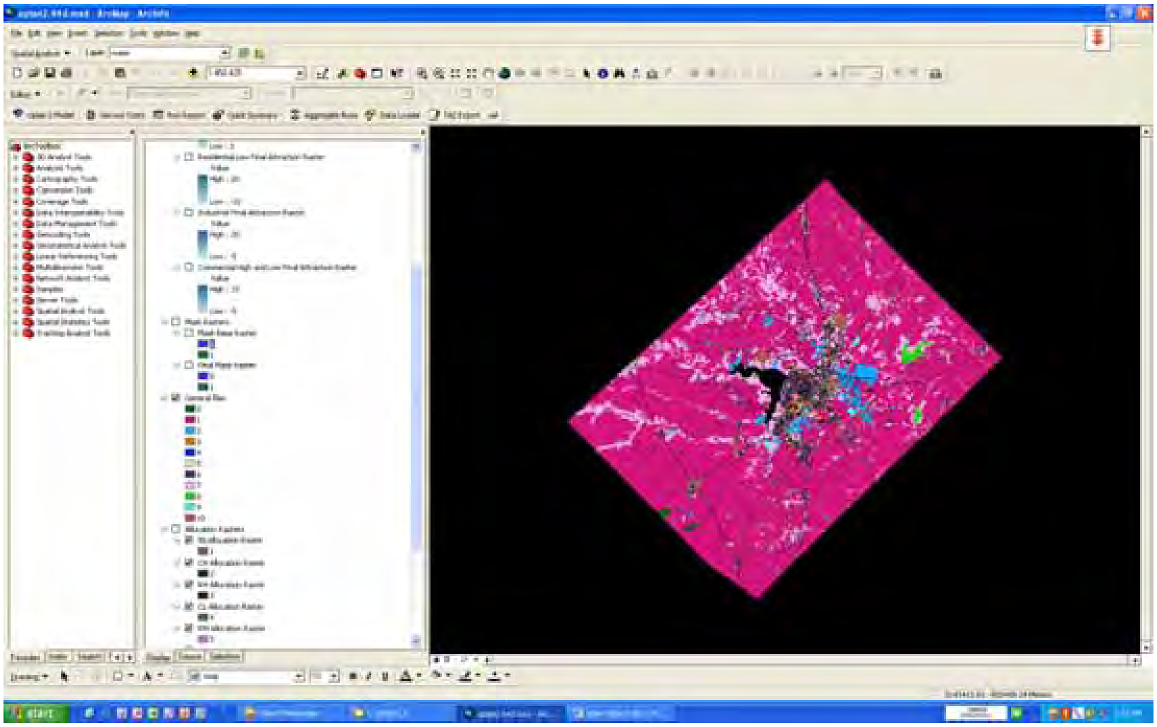
21) Map showing commercial final raster



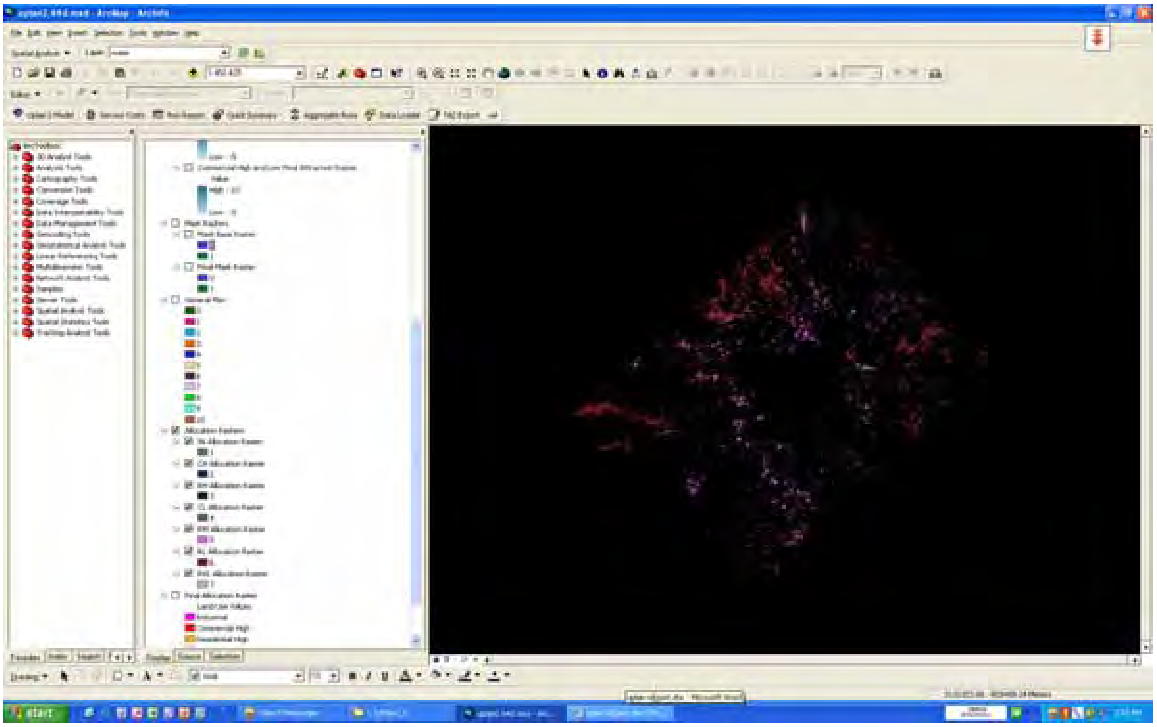
22) Mask Map



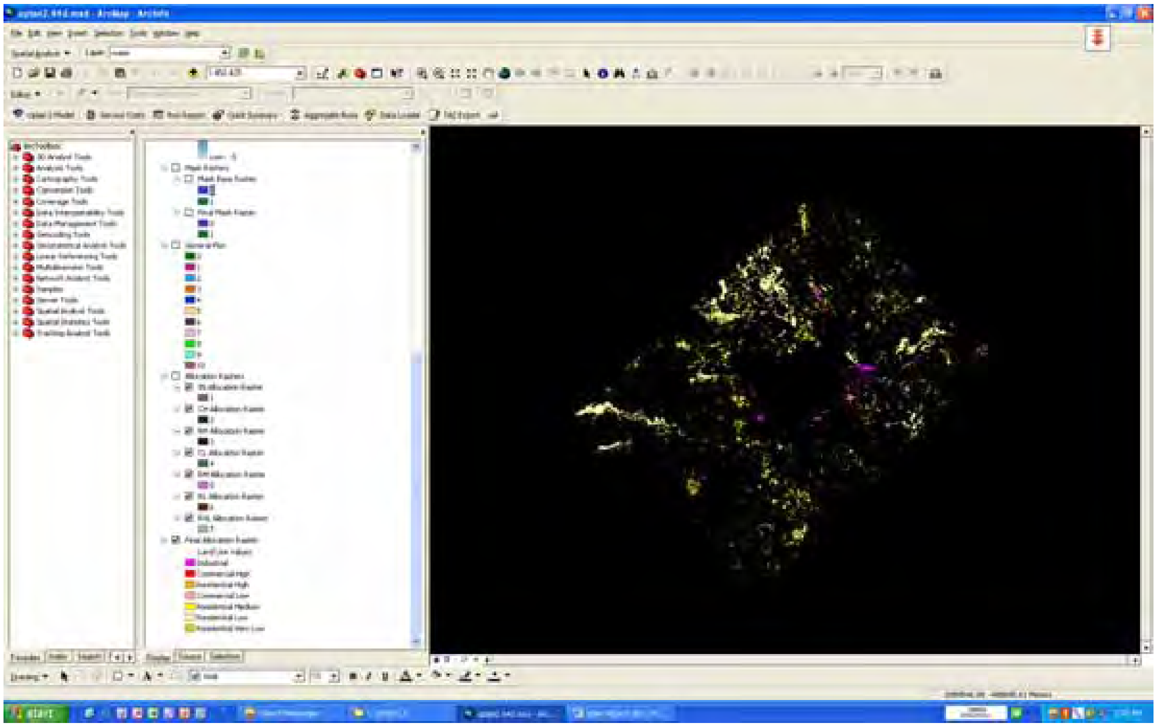
23) General plan



24) Final allocation raster



25) Final allocation of land use data.



Appendix E: Model Estimation Results

Table E1: Residential Location Choice Model Estimation Results

Coefficients	Estimates	t-statistic
ln(Housing Cost)	-0.191	-19.43
ln(Basic Sector employment within walking distance)	0.0321	3.96
ln(Residential Units)	-0.779	-54.32
ln(Residential units within walking distance)	0.32	19.69
Cost to income ratio	0.0524	2.66
Percentage of high-income households within walking distance if high income	0.0237	19.49
Percentage of low-income households within walking distance if low income	0.0476	19.76
Percentage of mid income households within walking distance if mid income	0.0259	24.36
Percentage of minority households within walking distance if minority	0.0239	5.33
Percentage of minority households within walking distance if not minority	-0.0267	-5.76
Log-likelihood is:	-16750.072	
Null Log-likelihood is:	-20553.436	
Likelihood ratio index:	0.185	
Adj. likelihood ratio index:	0.1845	

Notes: 1) All the variables are at the gridcell level. 2) Walking distance is assumed as 600 meters.

Table E2(1): Industrial Employment Location Choice Model Estimation Results

Coefficients	Estimates	t-statistics
ln(Industrial sqft)	0.117	4.52
ln(Industrial sqft within walking distance)	-0.0601	-1.11
Non-Residential Improvement value per sqft	0.0604	2.96
ln(Total employment within walking distance)	-0.358	-1.34
ln(Total Non-Residential sqft within walking distance)	0.341	1.16
Log-likelihood is:	-5693.37	
Null Log-likelihood is:	-5727.616	
Likelihood ratio index:	0.00598	
Adj. likelihood ratio index:	0.00511	

Table E2(2): Commercial Employment Location Choice Model Estimation Results

Coefficients	Estimates	t-statistics
ln(Total non-residential improvement value per sqft)	-0.0482	-2.41
ln(Industrial sqft.)	0.01625	4.69
ln(Retail sector employment within walking distance)	-0.1661	-13.73
ln(Service sector employment within walking distance)	0.26663	16.88
ln(Total Land Value)	-0.0888	-2.48
ln(Total Value)	0.07806	2.36
Log-likelihood is:	-9237.97	
Null Log-likelihood is:	-9506.35	
Likelihood ratio index:	0.028	
Adj. likelihood ratio index:	0.0276	

Table E2(3): Home-Based Employment Location Choice Model Estimation Results

Coefficients	Estimates	t-statistics
ln(Total non-residential improvement value per sqft)	0.379	2.44
Average income per residential unit	2.31E-05	2.15
Building Age	0.135	3.54
ln(Average land value per acre within walking distance)	0.243	2.15
ln(Total population within walking distance)	-0.638	-4.36
ln(Same sector employment within walking distance)	0.308	4.51
Log-likelihood is:	-199.058	
Null Log-likelihood is:	-248.287	
Likelihood ratio index:	0.198	
Adj. likelihood ratio index:	0.174	

Table E3: Land Price Model Estimation Results

Coefficients	Estimates	t-statistics
Constant	34.916	13.97
Total Improvement Value	0.0246	39.01
Total Residential value per residential unit within walking distance	0.0301	50.45
Year Built	0.0121	48.17
Is near Highway?	-0.891	-7.5
ln(Basic sector employment within walking distance)	-0.0592	-13.88
ln(Commercial sqft.)	0.0228	17.09
ln(Commercial sqft. within walking distance)	0.00555	4.7
ln (Distance to highway)	-0.151	-77.81
ln (Home access to employment)	-3.453	-18.87
ln(Industrial sqft.)	0.00591	3.53
ln(Industrial sqft. within walking distance)	0.0112	8.1
ln(residential units)	0.246	125.39
ln(residential units within walking distance)	0.0875	57.58
ln(Retail sector employment within walking distance)	0.0122	5.01
ln(Service sector employment within walking distance)	0.0875	27.01
ln(Total employment within walking distance)	0.0821	16.13
ln(Total Non-residential sqft. within walking distance)	-0.00426	-2.67
Percent commercial within walking distance	0.00426	14.21
Percent developed within walking distance	0.00474	27.48
Percent low-income households within walking distance	-0.00191	-21.06
Percent open space within walking distance	-0.013	-55.48
Percent residential	-0.0049	-42.4
Percent residential within walking distance	-0.0079	-67.52
R-Squared:	0.63014	
Adjusted R-Squared:	0.63011	

Table E4: Residential Land Share Model Estimation Results

Coefficients	Estimates	t-statistic
Constant	56.0323	2.33
ln(Basic sector employment within walking distance)	-0.023	-3.24
ln(Home access employment within walking distance)	-4.0594	-2.26
ln(Residential units)	0.1853	27.71
ln(Residential units within walking distance)	-0.017	-2.56
ln(Retail sector employment within walking distance)	-0.0241	-3.16
ln(Service Sector employment within walking distance)	-0.0525	-6.05
Travel time to CBD	0.00377	2.79
R-Squared:	0.0476	
Adjusted R-Squared:	0.0472	

Appendix F: Model Description

Gravity-based Land Use Model (G-LUM)

The residential land use allocation module, RESLOC, uses Equation F1 and the employment land use allocation module, EMPLOC, uses Equation F2. The following indices are used: *n*—household type (low income, medium-low income, medium-high income, high income), *i* and *j*—travel zones, *k*—employment sector (basic, retail, service), and *t*—time period. Note that for each household type, the following equation holds:

$$Q_{j,t}^{n} = \sum_{k} a_{k,n} \frac{E_{j,t}^{k}}{1-u_{k}}$$

Where $a_{k,n}$ is the regional average ratio of type n households per type k employees, u is the unemployment rate, and E is employment. Other notation needed is c—travel cost between zones, x—the percentage of developable land already developed, L^r—residential land, and N—number of households. Employment by sector and households by type are constrained by regional control totals and are scaled accordingly if a control total is not met.

$$N_{i,t}^{n} = \eta^{n} \sum_{j} Q_{j,t}^{n} \exp\left(\rho^{n} c_{i,j,t-1}\right) \frac{\left(1+x_{i,t-1}\right)^{r^{n}} \left(L_{i,t-1}^{r}\right)^{q^{n}} \left(1+\dfrac{N_{i,t-1}^{n}}{\sum_{n'} N_{i,t-1}^{n'}}\right)^{b^{n}}}{\sum_{i} \left(1+x_{i,t-1}\right)^{r^{n}} \left(L_{i,t-1}^{r}\right)^{q^{n}} \left(1+\dfrac{N_{i,t-1}^{n}}{\sum_{n'} N_{i,t-1}^{n'}}\right)^{b^{n}} \exp\left(\rho^{n} c_{i,j,t-1}\right)} + \left(1-\eta^{n}\right) N_{i,t-1}^{n} \quad \text{(F1)}$$

$$E_{j,t}^{k} = \eta^{k} \sum_{i} \frac{N_{T,i,t-1} \left(E_{j,t-1}^{k}\right)^{\omega^{k}} \exp\left(\rho^{k} c_{i,j,t-1}\right)}{\sum_{j'} \left(E_{j',t-1}^{k}\right)^{\omega^{k}} \exp\left(\rho^{k} c_{i,j,t-1}\right)} + \left(1-\eta^{k}\right) E_{j,t-1}^{k} \quad \text{(F2)}$$

The land allocation module, LUDENSITY uses Equation F3, which allows for an exponential decrease in the amount of land allocated to each use as the number of uses in a zone increases. The index l denotes the use (basic employment, commercial employment [retail and service], residential [all income levels]). U denotes the amount of use, β are parameters determined through calibration, and δ is a constant determined based on L_0 - an initial condition for the amount of land devoted to a use if the zone is empty. If the amount of allocated land exceeds zone size, then it is assumed that multi-story development is taking place.

$$L^{l}_{j,t} = \frac{U^{l}_{j,t} - U^{l}_{j,t-1}}{\beta_1\left(\left(\sum_{l'} U^{l'}_{j,t}\right) + \delta^{l}\right)^{\beta_2}} + L^{l}_{j,t-1} \qquad (F3)$$

Travel Demand Model Description

1. Trip Generation

The trip generation step determines the number of trips produced from and attracted to each zone. A look-up table is used (Table F1) based on Martin (1998, p. 25) to obtain trip production rates by purpose (home-based work, home-based other, and non-home-based) and by income level.

Table F1: Daily Person Trip Production Rates by Purpose and Income Levels

Income Level	Average Trips per Household	% Average Trips by Purpose HBW	HBO	NHB
Low	6	16	60	24
Medium	9.3	21	56	23
High	12.7	20	55	25

Trip attractions by trip type (table F2) are determined using modified versions of the equations provided on page 28 of Martin (1998).

Table F2: Trip Attractions

Trip Type	Trip Attraction per Zone
Home-based Work	*1.45*TOTAL_EMPLOYMENT*
Home-based Non-work	*9.00*RETAIL+ 0.5*BASIC+1.7*SERVICE+0.5*TOTAL_HH*
Non-home-based Work	*4.10*RETAIL+ 0.5*BASIC+1.2*SERVICE+0.5*TOTAL_HH*

Once trip productions and attractions have been calculated, balancing must be done to ensure that the sum of all productions equals the sum of all attractions. Assuming that trip productions are more reliable and that all trips are internal trips, the number of trips attracted to each zone for each purpose should be scaled by the total number of productions divided by the total number of attractions.

2. Trip Distribution

The trip distribution step uses a gravity model to connect the trip productions and attractions and determine a matrix of flows from origin zones to destination zones. For each of the four trip types, the following equation was used:

$$T_{ij} = \frac{P_i A_j e^{-\theta c_{ij}}}{\sum_{j'} A_{j'} e^{-\theta c_{ij'}}} \qquad (F4)$$

where, i is the index for origin zones and j is the index for destination zones, T_{ij} is the flow from i to j, P_i is the number of trips produced from zone i, A_j is the number of trips attracted to zone j, and c_{ij} is the cost of travel from i to j. The mean value of the parameter θ is 0.123 for home-based work trips, 0.094 for non-home-based work trips, and 0.1 for non-work trips. Intrazonal travel times were assumed to be as follows (Martin, 1989, page 39):

$$c_{ii} = \frac{0.5 * 3600\left[\frac{\sec}{hr}\right] * \sqrt{ZoneArea\left[mi^2\right]}}{15\left[mph\right]} \qquad (F5)$$

3. Mode Choice

A multinomial logit choice model is used to determine the share of trips that use transit, shared ride, and drive alone modes. The utility equation for each mode is given in table F3 and the coefficients are based on recommendations from Martin (1998, p. 66). All travel times, in-vehicle travel time (*IVTT*) and out-of-vehicle travel time (*OVTT*) are in minutes and *COST* is in units of cents.

Table F3: Mode Utilities

Mode	Utility
Transit	$-0.025*IVTT_{TR}-0.050*OVTT_{TR}-\beta_2*COST_{TR}$
Shared ride	$-0.025*IVTT_{SR}$
Drive alone	$-0.025*IVTT_{DA}$

Here, β_2 is a function of the origin zone's income distribution and an assumed value of time as percentage of income. The table on page 66 of Martin (1998) was used to determine appropriate values of this coefficient. $IVTT_{TR}$ was assumed to be twice $IVTT_{DA}$ plus twelve minutes (Lemp, XX) and $IVTT_{SR}$ is assumed to be 150% of $IVTT_{DA}$. The share of trips by each mode, m, is calculated using Equation A6, where U_m denotes the utility for mode m.

$$\text{Share of mode } m = \frac{e^{U_m}}{\sum_m e^{U_m}} \qquad (F6)$$

4. Route Choice

Person flows are converted to vehicle flows using the following equivalents: 7.1 for transit, 2 for shared ride, and 1 for drive alone. Additionally, bus flows are converted to passenger car flows by multiplying by a factor of 3 because of the additional space buses consume on roadways and their increased impact on capacity. The total passenger car units are assigned to the road network assuming user equilibrium (UE) behavior; no user can unilaterally switch routes without incurring a higher cost. The solution method is as follows:

1) Find the shortest path between each origin-destination pair using Dijkstra's algorithm (1959).
2) Average the current path flows with previous path flows using the golden section method.
3) Calculate the new path costs.
4) If not converged, return to Step 1.

See Sheffi (1985) for a detailed description of the steps.

Appendix G: Latin Hypercube Sampling

Latin Hypercube Sampling (McKay et al., 1979) involves partitioning the range of the random parameters into sections of equal probability and selecting samples from each section in a manner which ensures that the entire sample space is covered. The number of sections is fixed to be equal to the number of sample realizations needed. Then for each dimension of the uncertain parameter, random realizations are generated from each of the sections. In this task, the realizations for each region or dimension are combined in a random manner to form $|\Omega|$ realization vectors.

Note that the realizations obtained using Latin Hypercube, $\xi^{\omega} \forall \omega \in \Omega$, are not IID. In order to use the standard formula for obtaining a sample estimate of the variance, M batches of realizations were generated (Stein 1987): $\xi^{\omega m} \forall \omega \in \Omega, m \in \{1..M\}$. The average of the function $F(y, \xi^{\omega m})$ for each batch m is calculated as $F^m = \frac{1}{|\Omega_m|} \sum_{\omega \in \Omega_m} F(y, \xi^{\omega m})$ where $|\Omega_m|$ denotes the number of realizations in each batch.

The estimate of $E[F^m]$ or $\hat{F}_M$ is calculated as $\hat{F}_M = \frac{1}{M} \sum_{m=1}^{M} F^m$. By assuming the average of the function for each batch, F^m, to be IID, the variance of the estimate is obtained as the variance of F^m across the M batches. When a sufficiently large number of batches is generated (M > 30), the Central Limit Theorem states that F^m can be assumed to be normally distributed allowing confidence intervals to be determined.

Appendix H: Base Year Data

Table H1: Base Year Employment and Households

	Employment (by sector)			Households (by income)			
Zone	**Basic**	**Retail**	**Service**	**Low**	**Med-Low**	**Med-High**	**High**
1	602	125	475	1	4	4	1
2	5	100	478	8	33	29	5
3	5	101	472	6	30	28	8
4	5	102	470	7	35	30	8
5	599	124	474	1	5	3	1
6	5	95	479	10	30	30	9
7	600	96	832	12.5	37.5	56.25	18.75
8	5	542	838	31	91	134	44
9	610	508	831	10	36	55	20
10	5	101	460	6	35	30	10
11	5	105	464	7	33	37	9
12	5	553	831	27	88	130	41
13	5	653	837	35	91	136	47
14	5	542	822	31	91	134	43
15	5	104	471	7	30	31	7
16	5	98	476	8	32	32	8
17	5	554	832	12	36	55	19
18	150	550	833	31	94	130	46
19	5	540	819	11	38	56	19
20	5	110	470	6	31	30	9
21	591	113	472	2	5	4	1
22	5	111	478	9	30	30	9
23	5	95	465	7	31	30	10
24	5	98	464	7	30	29	6
25	589	124	459	1	4	6	1

Table H2: Base Year Land Allocation

	Land (by use) [acres]				
Zone	Basic	Commercial	Residential	Streets	Undevelopable
1	120.4	86.4	16	0.01	0
2	1.44	115.022	193.5	0.01	0
3	1.55	114.6	158.4	0.01	0
4	1.475	114.4	224	0.01	0
5	125.79	85.514	17	0.01	10
6	1.45	113.652	205.4	0.01	0
7	114	120.64	188.75	0.01	0
8	0.9	175.26	414	0.01	0
9	91.5	133.9	121	0.01	0
10	1.5	131.835	218.7	0.01	0
11	1.385	110.386	210.7	0.01	0
12	0.975	178.536	386.1	0.01	0
13	0.905	178.8	386.25	0.01	0
14	0.995	178.684	418.6	0.01	0
15	1.5	120.75	213.75	0.01	0
16	1.4	111.93	200	0.01	0
17	1.005	184.338	183	0.01	0
18	26.85	164.577	388.29	0.01	0
19	1.055	183.465	192.2	0.01	0
20	1.395	111.36	188.48	0.01	0
21	135.93	84.825	22.8	0.01	0
22	1.38	112.499	190.32	0.01	0
23	1.65	134.4	226.2	0.01	0
24	1.7	140.5	216	0.01	0
25	147.25	87.45	24	0.01	0

Appendix I: TELUM Model

i) Employment Allocation Model (EMPAL):

The following equations were used to allocate the employment in a zone (Putman 2005):

$$E_{j,t}^{k} = \lambda^{k} \sum_{i} P_{i,t-1} A_{i,t-1}^{k} W_{j,t-1}^{k} c_{i,j,t}^{\alpha^{k}} \exp\left(\beta^{k} c_{i,j,t}\right) + \left(1-\lambda^{k}\right) E_{j,t-1}^{k} \tag{I1}$$

where

$$A_{i,t-1}^{k} = \left[\sum_{l} \left(E_{l,t-1}^{k}\right)^{a^{k}} \left(L_{l}\right)^{b^{k}} c_{i,l,t}^{\alpha^{k}} \exp(\beta^{k} c_{i,l,t}) \right]^{-1} \tag{I2}$$

$$W_{j,t-1}^{k} = \left(E_{j,t-1}^{k}\right)^{a^{k}} \left(L_{j}\right)^{b^{k}} \tag{I3}$$

$E_{j,t}^{k}$ = Employment (place-of-work) of type k in zone i at time t
$P_{i,t}$ = Total number of households in zone i at time t
$c_{i,j,t}$ = Impedance (travel time or cost) between zones i and j at time t
L_{j} = Total area of zone i

λ^{k} , α^{k} , β^{k} , a^{k} , b^{k} are empirically derived parameters

ii) Disaggregate Residential Allocation Model (DRAM®):

The allocation of households in a zone was based on the following equations (Putman S.H, 2005):

$$N_{i}^{n} = \eta^{n} \sum_{j} Q_{j}^{n} B_{j}^{n} W_{i}^{n} c_{i,j}^{a^{n}} \exp\left(\beta^{n} c_{i,j}\right) + (1-\eta^{n}) N_{i,t-1}^{n} \tag{I4}$$

$$Q_{j}^{n} = \sum_{k} a_{k,n} E_{j}^{k} \tag{I5}$$

$$B_{j}^{n} = \left[\sum_{i} W_{i}^{n} c_{i,j}^{a^{n}} \exp\left(\beta^{n} c_{i,j}\right) \right]^{-1} \tag{I6}$$

$$W_{i}^{n} = \left(L_{i}^{v}\right)^{q^{n}} \left(x_{i}\right)^{r^{n}} \left(L_{i}^{r}\right)^{s^{n}} \prod_{n'} \left[\left(1 + \frac{N_{i}^{n'}}{\sum_{n} N_{i}^{n}} \right)^{b_{n'}^{n}} \right] \tag{I7}$$

N_i^n = Households of type n residing in zone i
$c_{i,j}$ = Impedance (travel time or cost) between zones i and j
$a_{k,n}$ = (Regional) number of type n households per type k employee
E_j^k = Employment of type k (place-of-work) in zone j
L_i^v = Vacant developable land in zone i
x_i = 1.0 plus the proportion of developable land already developed in zone i
L_i^r = Residential land in zone i

α^n, q^n, r^n, s^n, $b_{n'}^n$ are empirically derived parameters

iii) Land Consumption Model (LANCON):

The land consumption for different purposes in a zone at a time period is computed by the following equations. The TELUM user manual does not mention the LANCON equation used in it. The LANCON equation given (Putman S.H, 1991) requires the land use data of both lead and lag year for calibration. The LANCON equations given below were constructed by us based on the calibration report of TELUM.

The residential land-consumption equation is:

$$\frac{L_{r,i,t}}{N_{T,i,t}} = k_0 L_{D,i,t}^{k_1} \left(\frac{L_{d,i,t}}{L_{D,i,t}}\right)^{k_2} \left(\frac{E_{b,i,t}}{E_{T,i,t}}\right)^{k_3} \left(\frac{E_{c,i,t}}{E_{T,i,t}}\right)^{k_4} \left(\frac{N_{l,i,t}}{N_{T,i,t}}\right)^{k_5} \left(\frac{N_{h,i,t}}{N_{T,i,t}}\right)^{k_6} \qquad \text{(I8)}$$

where

$L_{r,i,t}$ = Amount of residential land use in zone i at time t

$N_{T,i,t}$ = Total number of households in zone i at time t

$L_{D,i,t}$ = Amount of developable (developed plus vacant) land use in zone i at time t

$L_{d,i,t}$ = Amount of "developed" land use in zone i at time t

$E_{T,i,t}$ = Total employment in zone i at time t

$E_{b,i,t}$ = Amount of "basic" employment in zone i at time t

$E_{c,i,t}$ = Amount of "commercial" employment in zone i at time t

$N_{l,i,t}$ = Number of low-income households in zone i at time t

$N_{h,i,t}$ = Number of high-income households in zone i at time t

$k^0 - k^6$ are empirically estimated parameters

The "basic" industry land-consumption equation is:

$$\frac{L_{b,i,t}}{E_{b,i,t}} = g_0\left(\frac{L_{d,i,t}}{L_{D,i,t}}\right)^{g_1}\left(\frac{E_{b,i,t}}{E_{T,i,t}}\right)^{g_3}\left(\frac{E_{c,i,t}}{E_{T,i,t}}\right)^{g_3}\left(\frac{N_{l,i,t}}{N_{T,i,t}}\right)^{g_4}\left(\frac{N_{h,i,t}}{N_{T,i,t}}\right)^{g_5} \tag{I9}$$

$g^0 - g^5$ are empirically estimated parameters

The "commercial" industry land-consumption equation is:

$$\frac{L_{c,i,t}}{E_{c,i,t}} = p_0\left(\frac{L_{d,i,t}}{L_{D,i,t}}\right)^{p_1}\left(\frac{E_{b,i,t}}{E_{T,i,t}}\right)^{p_3}\left(\frac{E_{c,i,t}}{E_{T,i,t}}\right)^{p_3}\left(\frac{N_{l,i,t}}{N_{T,i,t}}\right)^{p_4}\left(\frac{N_{h,i,t}}{N_{T,i,t}}\right)^{p_5} \tag{I10}$$

$p^0 - p^5$ are empirically estimated parameters