UrbanSim: Modeling Urban Development for Land Use, Transportation and Environmental Planning

Paul Waddell
Department of Urban Design and Planning and
Daniel J. Evans School of Public Affairs
University of Washington
Box 353055
Seattle, WA 98195
pwaddell@u.washington.edu

UrbanSim: Modeling Urban Development for Land Use, Transportation and Environmental Planning

Abstract

Metropolitan areas have come under intense pressure to respond to federal mandates to link planning of land use, transportation, and environmental quality; and from citizen concerns about managing the side effects of growth such as sprawl, congestion, housing affordability, and loss of open space. The planning models used by Metropolitan Planning Organizations (MPOs) were generally not designed to address these questions, creating a gap in the ability of planners to systematically assess these issues. UrbanSim is a new model system that has been developed to respond to these emerging requirements, and has now been applied in three metropolitan areas. This paper describes the model system and its application to Eugene-Springfield, Oregon.

Introduction

The relationships between land use, transportation, and the environment are at the heart of growth management. The emerging concern that construction of new suburban highways induces additional travel, vehicle emissions, and land development, making it implausible to 'build our way out' of congestion, has reshaped the policy context for metropolitan transportation planning (Downs, 1992). Recognizing the effects of transportation on land use and the environment, the Clean Air Act Amendments of 1990, the 1991 Intermodal Surface Transportation Efficiency Act (ISTEA) mandated that MPOs integrate metropolitan land use and transportation planning. These legislative actions have led to subsequent legal challenges to the traditional approach to transportation planning that ignores these feedback effects (Garret and Wachs, 1996). The passage of the Transportation Equity Act for the 21st Century (TEA21) in 1998, as the successor to ISTEA, softened these planning requirements somewhat, but significant pressure remains to better coordinate metropolitan planning of land use, transportation, and the environment.

Requirements for improved and better integrated land use and transportation models have emerged not only in response to this federal legislation, but also from state growth management programs that promote closer linkage of land use and transportation planning. Pressure for change has also come from the community of practicing and academic planners and advocates for the environment and alternative modes of transportation that have become frustrated with the state of the practice, as exemplified in the Portland LUTRAQ project (Blizzard, 1996). In response to the growing concern regarding the limitations of current land use and transportation models, the Travel model Improvement Project (TMIP) was formed as a collaborative effort by the Federal Highway Administration, the Federal Transit Administration and the Environmental Protection Agency.

In 1995, TMIP hosted an international conference on land use modeling to convene practitioners, researchers, and consultants, to assess the state of the practice and to make recommendations for new model development to address limitations in the current practice. Recommendations put forward at this conference included moving fairly quickly toward random utility-based models; using a clear behavioral basis describing the principal actors and choices involved in urban development and transportation; placing greater emphasis on the use of models for policy analysis, planning, and sensitivity testing; recognizing the varying temporal and geographic scales relevant to different processes in urban development; moving to disaggregate models and data; drawing on multiple disciplines; developing modular models; increased use of GIS and Remote Sensing; and testing the effects of transportation on land use (Weatherby, 1995).

Federal efforts to improve the state of the practice through the TMIP program have focused almost exclusively on long-term investment in a research and development effort for a new traffic microsimulation model, called TRANSIMS, and have not yet made any investment in new land use modeling approaches. As a result, the initiative for developing new land use models has been taken up at the state and local level. Efforts such as the Oregon Department of Transportation's Transportation and Land Use Model Integration Project (TLUMIP), the State of Utah's Quality Growth Enhancement Tools (QGET) and Envision Utah efforts, the Oahu Metropolitan Planning Organization's investments in new land use and transportation models, among others, are leading the way.

Within the Oregon growth management context, the Oregon Department of Transportation launched an ambitious effort in 1996 to develop new integrated models to evaluate the interactions between transportation and land use. The TLUMIP effort had two components. The first was the implementation of a statewide land use and transportation model, for which the TRANUS model (de la Barra, 1989) was adopted. The second component of TLUMIP was the development of UrbanSim, a new metropolitan-scale land use model for integration with transportation models. UrbanSim was designed specifically to address the policy analysis requirements of metropolitan growth management, with particular emphasis on land use and transportation interactions.

The Oregon TLUMIP effort extended the original UrbanSim design developed for Honolulu, Hawaii, and implemented a prototype version in the Eugene-Springfield metropolitan area. Testing of the current version of the model in the Eugene-Springfield area using data from 1980 to 1994 has provided a useful empirical validation of the model. UrbanSim has since been applied in Honolulu and Salt Lake City, and other metropolitan areas are beginning to apply it as well. The UrbanSim software is distributed as Open Source software under the GNU General Public License, which allows anyone to use, modify and redistribute the source code at no cost. It is available at www.urbansim.org.

The objectives of this paper are twofold. The first objective is to describe the UrbanSim model design at a level that facilitates an assessment of how it addresses emerging requirements for land use and transportation modeling, and how it compares to other existing modeling approaches. The second is to describe the application of the model to the Eugene-Springfield, Oregon metropolitan area, and assess its validity over an historical period. These objectives are addressed in turn in the following sections. The final section concludes with a description of further research and development priorities.

The Design of UrbanSim

Overview and Comparison to Other Operational Models

The design of UrbanSim differs significantly from several existing operational modeling approaches, including the spatial-interaction DRAM/EMPAL models developed by Putman (1983); the spatial input-output TRANUS and MEPLAN models, developed respectively by de la Barra (1989) and Echenique et al. (1990); the GIS-based California Urban Futures (CUF, CUF-2) Model (Landis, 1994, 1995; Landis and Zhang, 1998a, 1998b), the MUSSA model developed by Martínez (1992), and the CATLAS (and later METROSIM and NYMTC-LUM) model developed by Anas (1982). These models are discussed in detail in several recent reviews (Miller *et al*, 1998; Dowling *et al*, 2000; Parsons Brinckerhoff, 1998, U.S. EPA, 2000), which update a number of earlier reviews (Anas, 1987; Harris, 1985; Kain, 1985; Paulley and Webster, 1991; Southworth, 1995; Wegener, 1994, 1995). The pitfalls of large-scale urban models were convincingly articulated almost three decades ago (Lee, 1973; 1994), and remain significant concerns. The design of UrbanSim has been well informed by these criticisms of prior modeling efforts, as well as by advances in theory, computation, and econometric methods.

To clarify differences between UrbanSim and other operational urban models that have been reviewed elsewhere (Miller et al, 1998; Dowling et al, 2000; U.S. EPA, 2000), Table 1 compares the key features of four model approaches.

TABLE 1. Comparison of Operational Model Characteristics

Characteristic	DRAM/EMPAL	MEPLAN and TRANUS	CUF-2	UrbanSim
Model Structure	Spatial Interaction	Spatial Input-Output	Discrete Choice	Discrete Choice
Household Location Choice	Modeled	Modeled	Not Modeled	Modeled
Household Classification	Aggregate, 8 categories	Aggregate, User-Defined	Not Represented	Disaggregate, Income, Persons, Workers, Child
Employment Location Choice	Modeled	Modeled	Not Modeled	Modeled
Employment Classification	Aggregate, 8 categories	Aggregate, User-Defined	Not Modeled	Disaggregate, 10-20 Sectors
Real Estate Development	Not Modeled	Modeled	Modeled	Modeled
Real Estate Classification	4 Land uses	Aggregate, User-Defined	7 Land Uses	24 Development Types
Real Estate Measures	Acres	Acres Units Floorspace	Acres	Acres Units Floorspace
Real Estate Prices	Not Modeled	Modeled	Not Modeled	Modeled
Geographic Basis	Census Tracts or Aggregates	User-Defined Zones (2-300)	Grid Cells	Grid Cells
Temporal Basis	Quasi-dynamic, Equilibrium (5-10 year steps)	Cross-Sectional, Equilibrium	Annual, Dynamic	Annual, Dynamic
Interaction with Travel Models	Yes	Yes	No	Yes
Modular Model Structure	Partial	No	No	Yes
Software Access	Proprietary	Proprietary	NA	Open Source

A brief description of terms used in Table 1 is in order. The term *spatial interaction* refers to models that draw on the analogy of the physical relationship of gravity. The application to human geography, made popular by Wilson (1967) recognized the empirical pattern that trips between two locations increase as the activity (population and employment) in the origin and destination zones increases, and the travel cost decreases. A large class of location choice models draws on and extends this metaphor (Fotheringham and O'Kelly, 1989).

The term *discrete choice* refers to models that draw on discrete choice theory and the development of a class of econometric models known as *random utility maximization* (RUM). Daniel McFadden (1973, 2000) recently won the Nobel Prize in economics for his pioneering work in this area. The approach is suited to modeling choices between alternatives that are mutually exclusive. Many of the early applications of this class of techniques, including multinomial and nested logit models, were focused on the transportation mode choice problem.

The term *cross-sectional* refers to the use of one point in time for estimating a model, rather than using a *longitudinal* or *dynamic* approach that analyzes changes over multiple time periods. *Equilibrium* within the context of economic models describes a hypothetical long-term market steady-state condition in which supply, demand and prices are perfectly balanced so that no one can be made better-off without making someone else worse-off. It assumes that all buyers and sellers are operating within competitive markets and have full information about the current and future prices and benefits of consumption choices. *Disequilibrium* describes market conditions in which supply and demand are not perfectly balanced, reflecting various limitations in the responsiveness of economic agents. For example, short-term growth in demand for housing may outpace or lag the growth in housing supply, causing disequilibrium conditions we commonly refer to as the boom and bust cycle.

In short, the UrbanSim design departs from aggregate economic and spatial-interaction models that rely on cross-sectional equilibrium solutions using large geographic districts, and pursues an approach that is disaggregate and based on predicting changes over small time steps, as does the CUF-2 model (Landis and Zhang, 1998a, 1998b). Unlike the CUF-2 approach, however, the UrbanSim design explicitly represents the demand for real estate at each location, and the actors and choice processes that influence patterns of urban development and real estate prices. This design approach synthesizes and extends some of the best features of previous modeling efforts. It also uses an Open Source approach to provide free access to the underlying source code, and to make the model more open to scrutiny and to further extension and adaptation to emerging requirements for modeling.

A review of operational models for the Transit Cooperative Highway Research Program, project H-12, developed a specification for a proposed 'ideal' integrated land use and transportation model system, and assessed operational models compared to this framework, including DRAM/EMPAL, MEPLAN, TRANUS, NYMTC-LUM, MUSSA, and UrbanSim (Miller *et al*, 1998). The report concluded that UrbanSim came closest to their proposed 'ideal' specification. A more recent report by the National Cooperative Highway Research Project examined operational models and assessed their potential for use in evaluating the air quality impacts of highway capacity expansion (Dowling *et al*, 2000). The review of land use models included DRAM/EMPAL (ITLUP), MEPLAN, NYMTC-LUM, and UrbanSim, and singled out UrbanSim as a foundation for further development. A third recent review of models by the Environmental Protection Agency inventoried a large number of land use models and analytical tools, but did not undertake any assessment of them (U.S. EPA, 2000).

The preceding discussion provides a general description of the UrbanSim design, how it addresses emerging modeling requirements and compares to other modeling approaches. Other references provide empirical results from the original specification of the model (Waddell, 2000a), description of the data development process (Waddell, Moore and Edwards, 1998), detailed specifications of the current model implementation (Waddell *et al*, forthcoming), analysis of its relationship to land supply monitoring (Waddell, 2000b), and description of its theoretical foundations (Waddell, 2000c;

Waddell and Moore, forthcoming), its application as a decision support system (Waddell, 2001), and the underlying software infrastructure (Noth, Borning and Waddell, 2000). A more detailed description of the design of the UrbanSim database and model system is provided below, followed by a description of its application in the Eugene-Springfield, Oregon metropolitan area.

The Database

The data integration process for UrbanSim is depicted in Figure 1. The input data used to construct the model database, called the data store, include parcel files from tax assessor offices, business establishment files from the state unemployment insurance database or from commercial sources, census data, GIS overlays representing environmental, political and planning boundaries, and a location grid. A set of software tools, collectively referred to as the data integration tools, read these input files, diagnoses problems in them such as missing or miscoded data, and applies decision rules to synthesize missing or erroneous data and construct the model data store.

The data store represents each household in the metropolitan area as an individual object, with the primary characteristics relevant to modeling location and travel behavior: household income, size, age of head, presence of children, and number of workers. The household list is synthesized by integrating Census household-level data from the Public Use Microdata Sample with Summary Tape File 3A tabulations by census tract, and assigning synthesized households probabilistically to parcel data, using a variant of the procedure developed for the TRANSIMS model system (Beckman *et al*, 1995). Employment is represented in the data store as individual records for each job and its employment sector.

The data store represents locations using grid cells of 150 by 150 meters, which contain an area just over 5.5 acres (the cell size can be modified). This location grid allows explicit cross-referencing of other spatial features such as planning and political boundaries such as city, county, traffic zones, urban growth boundaries; and environmental features such as wetlands, floodways, stream buffers, steep slopes, or other environmentally sensitive areas.

Figure 2 shows one grid cell in a central Seattle neighborhood of Queen Anne, over a digital orthophoto and parcel boundaries. Parcel data are collapsed into the cells to generate composite representations of the mix and density of real estate at each location, labeled *development types*. These development types are somewhat analogous to the development typology developed by Calthorpe (1983), in that they represent at a local neighborhood scale the land use mix and density of development. Table 2 provides the rules for classifying grid cell development into types, based on the combination of housing units, nonresidential square footage, and the principal land use of the development. The grid cell shown in Figure 2 would be classified as a development type of R8, or high-density residential, on the basis of containing 98 housing units and no non-residential square footage.

The data store maintains an explicit accounting of real estate and occupants, linking individual households to individual housing units, and individual jobs to *job spaces* that can be either nonresidential square footage, or a residential housing unit, to account for home-based employment. When jobs or households are predicted to move, the space they occupy is flagged as becoming vacant, and when they are assigned to a particular housing unit or job space, that space is reclassified as occupied. By explicit assignment of housing units and nonresidential square footage to grid cells of fixed size, densities and mixtures of housing units and nonresidential square footage of industrial, commercial, or governmental types are inventoried. Land values and residential and nonresidential improvement values are also identified for each cell in the database. This integrated data store of households, jobs, land and real estate is what the model components update over time. Although this

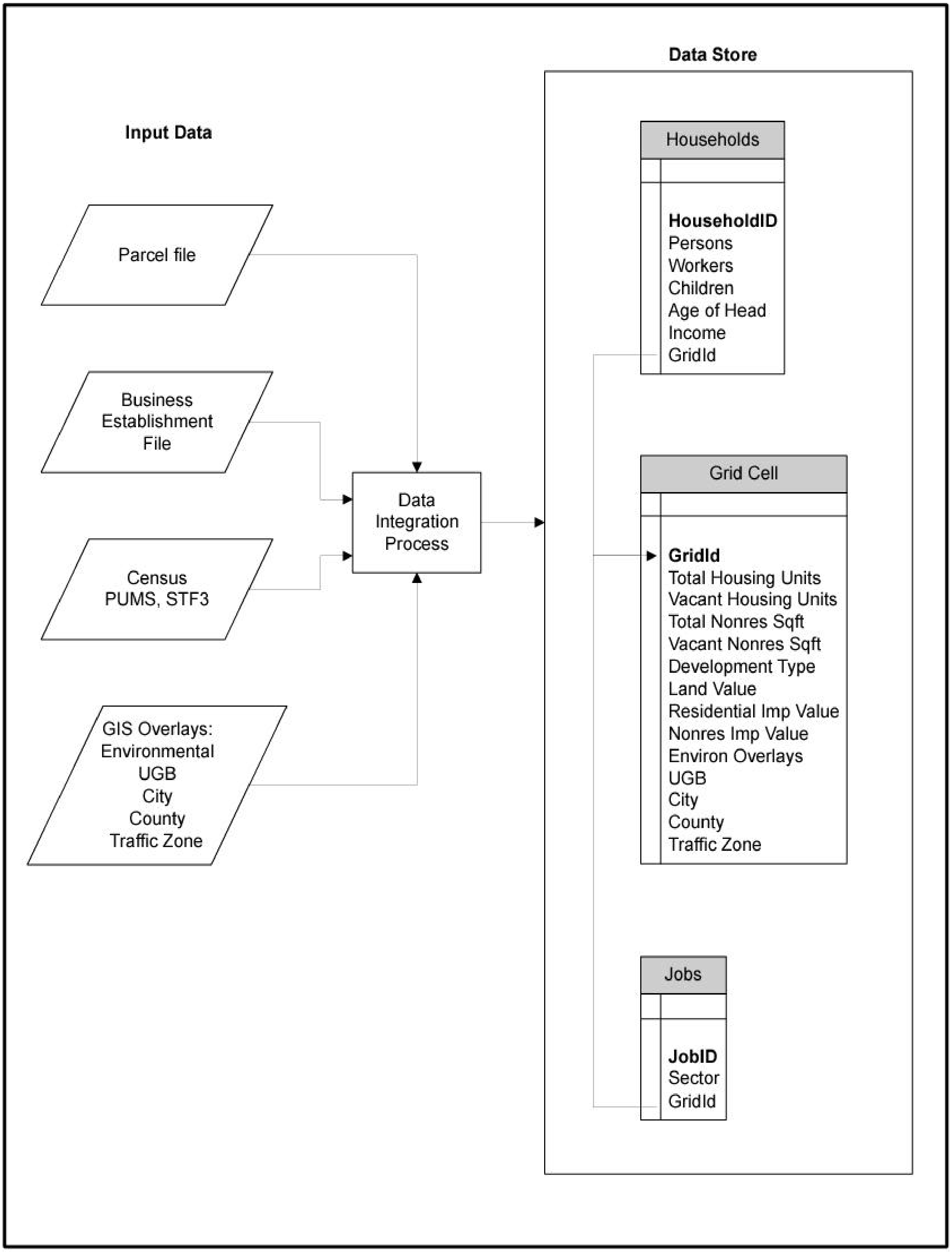

FIGURE 1. **UrbanSim Data Integration Process**

FIGURE 2. 150 Meter Grid Cells as Unit of Analysis for Location and Development

TABLE 2. Development Type Classification

DevType	Name	UnitsLow	UnitsHigh	SqftLow	SqftHigh	Primary_Use
1	R1	1	1	0	999	Residential
2	R2	2	4	0	999	Residential
3	R3	5	9	0	999	Residential
4	R4	10	14	0	2,499	Residential
5	R5	15	21	0	2,499	Residential
6	R6	22	30	0	2,499	Residential
7	R7	31	75	0	4,999	Residential
8	R8	76	65,000	0	4,999	Residential
9	M1	0	9	1,000	4,999	Mixed_R/C
10	M2	10	30	2,500	4,999	Mixed_R/C
11	M3	10	30	5,000	24,999	Mixed_R/C
12	M4	10	30	25,000	49,999	Mixed_R/C
13	M5	10	30	50,000	9,999,999	Mixed_R/C
14	M6	31	65,000	5,000	24,999	Mixed_R/C
15	M7	31	65,000	25,000	49,999	Mixed_R/C
16	M8	31	65,000	50,000	9,999,999	Mixed_R/C
17	C1	0	9	5,000	24,999	Commercial
18	C2	0	9	25,000	49,999	Commercial
19	C3	0	9	50,000	9,999,999	Commercial
20	I1	0	9	5,000	24,999	Industrial
21	I2	0	9	25,000	49,999	Industrial
22	I3	0	9	50,000	9,999,999	Industrial
23	GV	0	99,999	0	9,999,999	Government
24	VacantDevelopable	0	0	0	0	VacantDevelopabl
25	Undevelopable	0	0	0	0	Undevelopable

data store is derived from data about real households, businesses, and parcels, it is a synthetic database that represents only selected characteristics of people, jobs, real estate, and locations. Similarly, the models and their estimated parameters attempt to reflect the patterns of observed behavior of real agents, but are simplifications and abstractions of real behavior, as are all models.

Model Structure and Processing

UrbanSim includes model components reflecting the key choices of households, businesses, developers, and governments (as policy inputs) and their interactions in the real estate market. By focusing on the principal agents in urban markets and the choices they make about location and development, the model deals directly with behavior that planners, policy makers, and the public can readily understand and analyze. This behavioral approach provides a theoretical structure more transparent than 'black-box' models that do not clearly identify the agents and actions being modeled. The structure allows users to incorporate policies explicitly and to evaluate their effects.

UrbanSim is not a single model. It might be better described as an urban simulation system, consisting of a software architecture for implementing models and a family of models implemented and interacting within this environment. The models that are currently implemented employ a range of techniques and approaches. Some of the models, such as the economic and demographic transition models, are aggregate, non-spatial models that deal with the interface to external macro-economic changes. Other components such as location choice are discrete choice models of an agent (a household, for example) making choices about alternative locations, taking a top-down, or birds-eye view of the metropolitan area. The developer model, by contrast, takes a mostly bottom-up (worms-eye?) view, from the vantage point of a developer or land-owner at a single location (grid cell) making choices about whether to develop, and into what type of real estate. The bottom-up view in the developer model is tempered by market information that reflects the state of the market as a whole, such as vacancy rates.

The structure and processing sequence of UrbanSim are shown in Figure 3. Inputs to the model include the base year data store, control totals derived from external regional economic forecasts, travel access indicators derived from external transportation models, and scenario policy assumptions regarding development constraints arising from land use plans and environmental constraints. The individual model components predict the pattern of accessibility by auto ownership level (access model), the creation or loss of households and jobs by type (demographic and economic transition), the movement of households or jobs within the region (household and employment mobility models), the location choices of households and jobs from the available vacant real estate (household and employment location models); the location, type, and quantity of new construction and redevelopment by developers (developer model); and the price of land at each location (land price model). One special component, the model coordinator, manages the individual model components and handles the scheduling and implementation of events such as reads and writes to the data store. Taken together as a system, these components maintain the data store and simulate its evolution from one year to the next. For simplicity, the household and employment counterpart models for transition, mobility, and location are represented jointly in the diagram and described together in Table 3, since they are parallel and almost identical.

Simulated and User-Specified Events

The model system runs on events generated by the model components. A number of choices by households, businesses, and developers are simulated on an annual basis, and their outcomes are implemented as scheduled events. Large-scale development projects may be scheduled with multi-year timetables, defined using a template that describes the characteristics of different types of

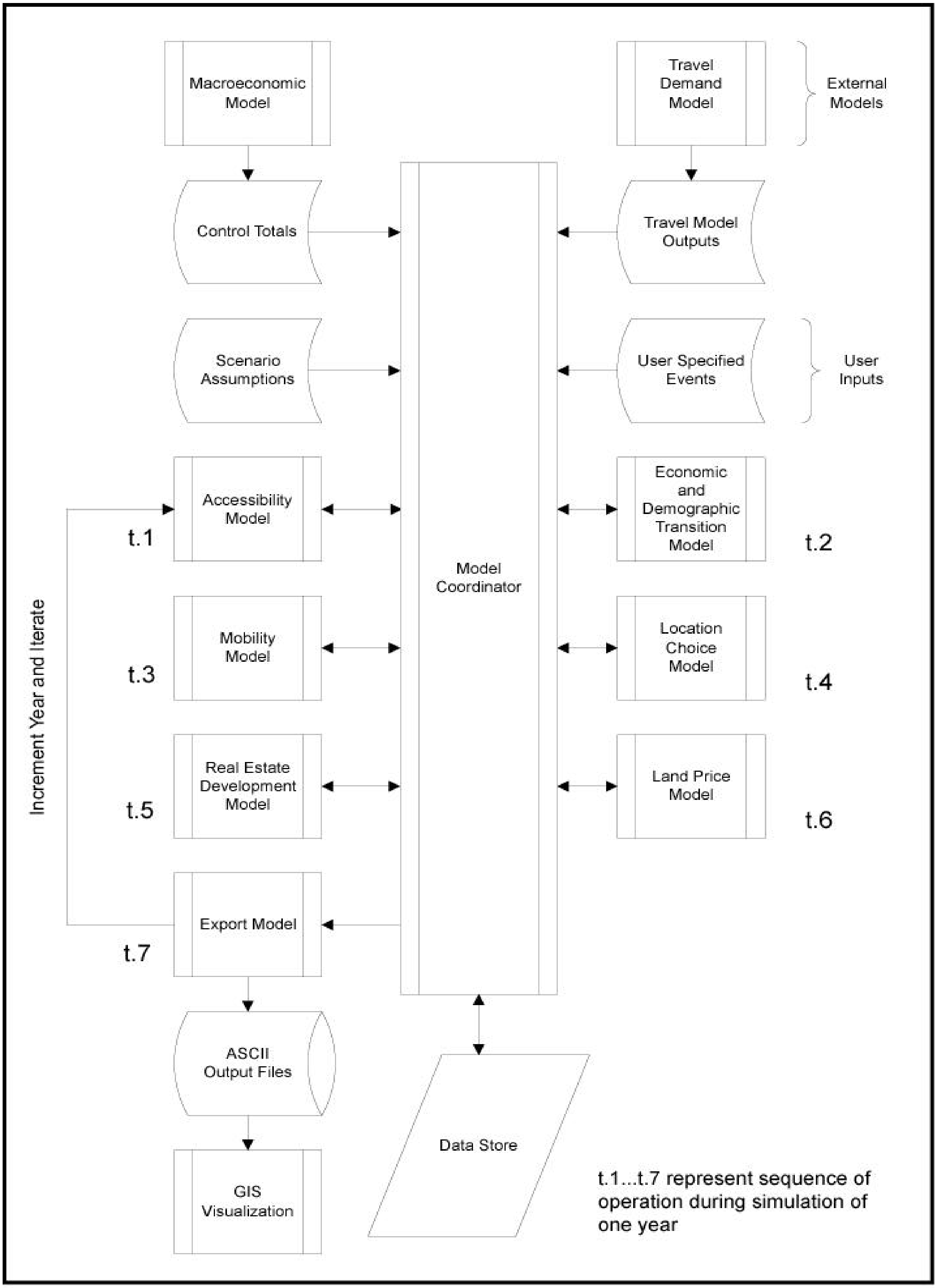

FIGURE 3. UrbanSim Model Structure and Processing

TABLE 3. Description of Core Models

Demographic and Economic Transition Models The Demographic Transition Model simulates births and deaths in the population of households. Externally imposed population control totals determine overall target population values, and can be specified in more detail by distribution of income groups, age, size, and presence or absence of children. This enables the modeling of a shifting population distribution over time. Iterative proportional fitting (Beckman *et al*, 1995) is used to determine how many households of each type are to be created or deleted. Newly created households are added to the household list but without an assignment to a specific housing unit (placed in limbo), to be placed in housing later by the Household Location Choice Model. Households to be deleted to meet the control totals are selected at random, drawn preferentially from households in limbo. The Economic Transition Model is responsible for modeling job creation and loss. Employment control totals are determine employment targets, and can be specified by distribution of business sector.
Household and Employment Mobility Models
The Household Mobility Model simulates households deciding whether to move. Movement probabilities are based on historical data. Once a household has chosen to move, it is placed in limbo to indicate it has no current location, and the space it formerly occupied is made available. The Employment Mobility Model determines which jobs will move from their current locations during a particular year using a similar approach to the Household Mobility Model.
Household and Employment Location Models
The Household Location Choice Model chooses a location for each household that has no current location. For each such household, a sample of locations with vacant housing units is randomly selected from the set of all vacant housing. Each alternative in the sample is evaluated for its desirability to the household, through a multinomial logit model calibrated to observed data. The household is assigned to its most desired location among those available. The Employment Location Choice Model is responsible for determining a location for each job that has no location. For each such job, a sample of locations with empty square feet, or space in housing units for home-based jobs, is randomly selected from the set of all possible alternatives. The variables used in the household location model include attributes of the housing in the grid cell (price, density, age), neighborhood characteristics (land use mix, density, average property values, local accessibility to retail), and regional accessibility to jobs. Variables in the employment location model include real estate characteristics in the grid cell (price, type of space, density, age), neighborhood characteristics (average land values, land use mix, employment in each other sector), and regional accessibility to population.
Real Estate Development Model
The Real Estate Development Model simulates developer choices about what kind of construction to undertake and where, including both new development and redevelopment of existing structures. Each year, the model iterates over all grid cells on which development is allowed and creates a list of possible transition alternatives (representing different development types), including the alternative of not developing. The probability for each alternative being chosen is calculated in a multinomial logit model. Variables included in the developer model include characteristics of the grid cell (current development, policy constraints, land and improvement value), characteristics of the site location (proximity to highways, arterials, existing development, and recent development), and regional accessibility to population. **Land Price Model** The Land Price Model simulates land prices of each grid cell as the characteristics of locations change over time. It is based on urban economic theory, which states that the value of location is capitalized into the price of land. The model is calibrated from historical data using a hedonic regression to include the effect of site, neighborhood, accessibility, and policy effects on land prices. It also allows incorporating the effects of short-term fluctuations in local and regional vacancy rates on overall land prices. Similar variables are used as in the Development Model.

development events. In addition to model-generated events, the system accommodates information that planners have about pending development, corporate relocations, or policy changes. We have developed a capacity to introduce user-specified events such as these into the model, both to allow planners to use available information about developments that are ‘in the pipeline’, and also to provide a capacity for testing the potential effects of a major project on further development and on traffic.

Scenario Assumptions

UrbanSim allows users to specify policy inputs and assumptions, generate and compare scenarios, compute evaluation measures, and query the database of results. The user interface of the model is focused on the interaction of the user with the inputs to each scenario. Scenarios consist of a combination of development policies, represented by appropriate input data such as comprehensive plans, infrastructure plans, urban growth boundaries, and development restrictions on environmentally sensitive lands. These policies are linked to locations at a grid cell, zonal, municipal, county, or metropolitan scale.

Broadly speaking, government agencies influence the land development process through a combination of land use regulations and infrastructure provision. These are frequently combined into packages that attempt to foster a development pattern in ways that promote planning objectives, for example by pursuing one or a combination of the following community visions:

- Containing development within an Urban Growth Boundary
- Focusing development along primary transportation corridors
- Focusing development within centers connected by multi-modal transportation
- Diverting development into new or existing satellite communities
- Encouraging development in parts of the region with underutilized infrastructure
- Promoting development of impoverished areas.

The use of the term ‘scenario’ differs in the UrbanSim context from its potential use to describe a particular ‘vision’ such as those listed above. An UrbanSim scenario is a collection of policy assumptions that can be input to the model to examine their potential consequences on outcomes such as urban form, land use mix, density, and travel patterns. In other words, the system allows interactive testing of how different policy strategies fare in achieving a particular vision or set of community objectives. It does not assume that a particular vision can be realized, but facilitates exploration of the trade-offs that may be involved in attempting to achieve it, given the range of policies available and their costs and consequences. The model does not attempt to ‘optimize’ policy inputs, but is intended to facilitate interactive use to support an iterative, participatory planning process.

The translation of these scenarios into inputs to UrbanSim involves interpreting policies and creating input files for the model that represent these policy interpretations. Interpreting the comprehensive land use plan is a key part of constructing a policy scenario in UrbanSim. Each land use plan designation (Planned Land Use, or PLU) may be described as a set of restrictions on development options. For example, the plan designation of ‘agricultural’ may not allow conversion to any developed urban category under restrictive interpretation of the land use plan, or may allow conversion to rural density single-family residential under a less restrictive interpretation. The adopted comprehensive plan guidelines for a local area should spell out the intended interpretation of these plan designations, but the user of the model may wish to assess the impact of altering these constraints as a matter of policy testing.

Development regulations may be coded for an entire metropolitan area, for individual counties, cities, or special overlays such as environmentally sensitive lands or urban growth boundaries. Overlays such as wetlands, floodways, steep slopes, or other environmental features may be used to specify environmental regulations that impose development constraints. The model interprets the cumulative impact of the policies by reflecting the most restrictive policies that apply to a given grid cell. For example, a general county plan might allow substantial development for a particular land use plan designation, but a more restrictive regulation that applies to wetlands would overrule this for any grid cell that was in a wetland. Figure 4 shows a portion of the UrbanSim user interface for specifying development constraints. PLU indicates 'Planned Land Use', which can represent either the land use plan designation or a zoning category.

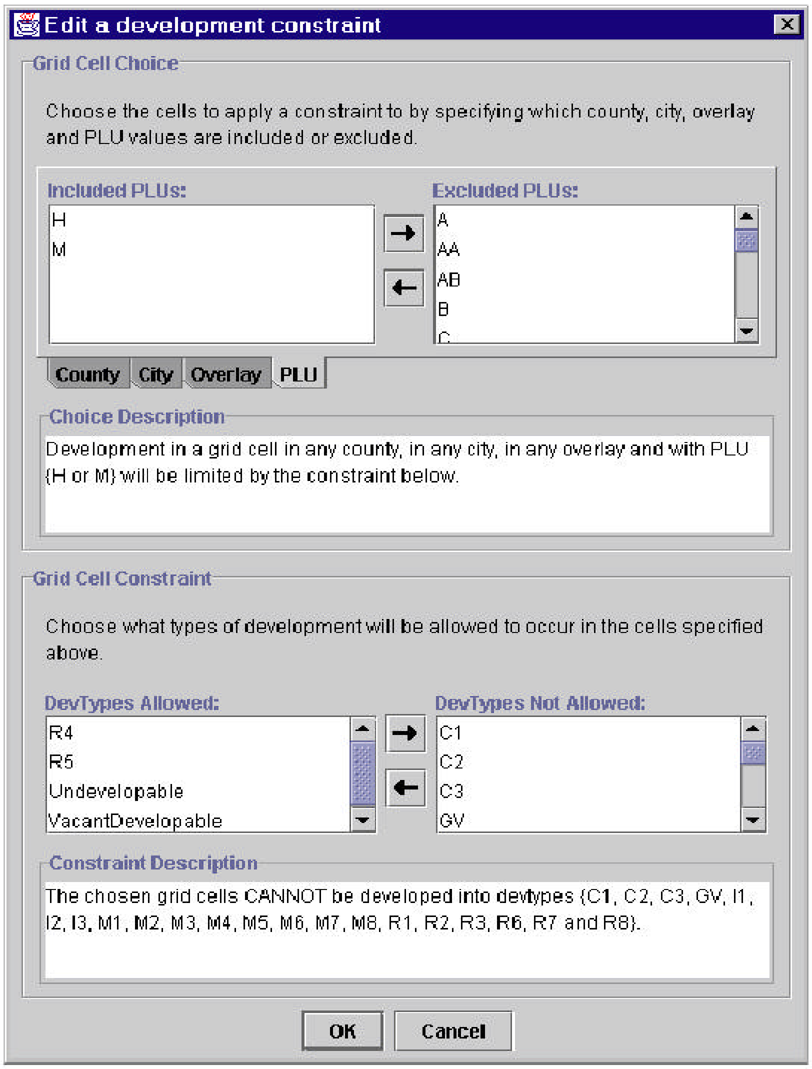

FIGURE 4. UrbanSim User Interface for Entering Development Constraints

In addition to development constraints, the scenario inputs include regional control totals from the external macroeconomic models, and assumptions about the space utilization rates (such as square feet per employee for different development types). Transportation policy assumptions are incorporated in the external transportation model, and are embedded in the travel time and utility outputs from the travel model that UrbanSim uses to calculate accessibility.

Local and Regional Accessibility

The Accessibility Model is responsible for maintaining accessibility values for occupants within each traffic analysis zone, including accessibility by residents and employees to shopping and other amenities, to employment, and to the central business district. The accessibility value for a zone to a specific type of activity is defined as the sum of the quantity of the activity (jobs, for example) at each possible destination, discounted by a weight between 0 and 1 reflecting the multimodal travel utility to the destination[1]. Handy (1993) and others have referred to this kind of measure as representing 'regional accessibility', in that it is regional in scope and uses the transportation network on a zone to zone basis to represent travel access. It is contrasted with 'local accessibility', which measures access to opportunities within a walkable neighborhood.

The link between land use and the travel model is two-way, since different accessibility values from the travel model will influence the decisions of developers, employers, and residents, giving rise to different travel demands, which then feed back into the travel model. The external travel model provides travel times and utilities to the Accessibility Model. The travel model is typically run only once every 5 simulated years or when there is a major change to the transportation system, since running it is relatively cumbersome and since its outputs generally change more slowly than other values in the simulation. However, UrbanSim is run annually, updating the accessibility values based on the evolving spatial pattern of activities.

UrbanSim also incorporates local accessibility measures, corresponding to the activities that can be reached by walking, over a distance of 600 meters (approximately 1/3 mile), using spatial queries of the grid cells in the data store. Achieving this scale of analysis makes UrbanSim the first operational urban model system to support analysis of location and travel behavior at a level that can effectively represent pedestrian and bicycle scales of travel. Given the ongoing debate over the potential influence of neo-traditional urban design on travel behavior, this innovation should provide a basis for making more systematic assessments of the effects of urban design-scale policies on both location and travel behavior. Traditional zone-based travel models are severely limited by poor performance on intra-zonal travel and insufficient representation of non-motorized travel modes. By creating a more detailed basis for the land use model, the main barrier to the improvement of transportation planning to address non-motorized modes and the integration of urban design policies has been effectively removed.

Data Export

The Data Export process is responsible for gathering, aggregating, and exporting data from the object store to a set of external files for subsequent analysis and graphical display. The user interface allows specification of desired output files and designation of specific simulation years for which to generate the outputs. Outputs are created at the grid cell level, and also summarized by traffic zone and for the region as a whole. The data are written in a standard format for ease of loading into ArcView, Excel, or other common desktop tools.

---

[1] The composite utility of travel from zone to zone is based on the logsum term in the mode choice model, which incorporates times and costs of all modes from origin to destination. It is scaled to a maximum value of 0, and exponentiated, to achieve a resulting weight between 0 and 1. The accessibility measure will therefore increase as modes are added, or as they are made faster or less costly.

Application to Eugene-Springfield, Oregon

The UrbanSim model system was first fully implemented in Eugene-Springfield, Oregon, in a project funded by the Oregon Department of Transportation TLUMIP project. The test site was chosen for a number of reasons, including the relatively small size of the metropolitan area and the availability of data needed for model application. The Lane Council of Governments (LCOG) is the Metropolitan Planning Organization for the area, and provided all the data used in the model development and application.

The Eugene-Springfield metropolitan area, as shown in Figure 5, lies in Lane County at the south end of the Willamette Valley in Oregon, a fertile agricultural region that contains most of the economic activity in the state, including the metropolitan areas of Portland and Salem. Eugene-Springfield is a relatively small metropolitan area, with a population of 322,959 in Lane County in the year 2000, of which 137,893 resided in the city of Eugene, and 52,864 in Springfield (U.S. Census, 2000).

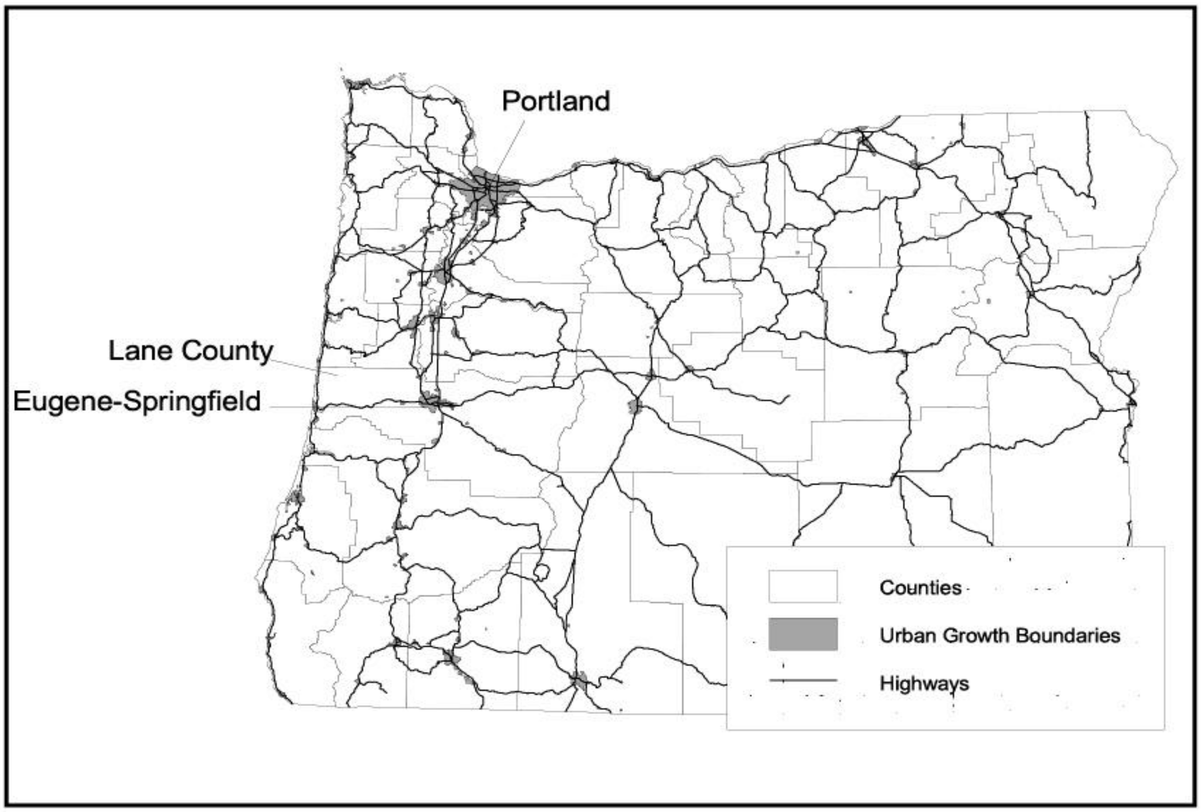

FIGURE 5. The Eugene-Springfield, Oregon Metropolitan Area

The base year for the Eugene-Springfield application is 1994, since consistent data were available for employment, household survey, and parcels. Population data were derived from a 1994 transportation home interview survey and from the 1990 census, and adjusted to the 1994 base year using the parcel distribution of the housing stock in that year. The scope of the study area is the extent of the 271 traffic analysis zone system represented in the Lane COG transportation model system. The study area was further subdivided into approximately 15,000 grid cells of 150 by 150 meters as the basic unit of spatial analysis in the residential location, employment location, real estate development, and land price models. Results are summarized by TAZ for input to the travel model system. The classification of households for use in the model is based on the stratification of five household characteristics, as shown in Table 4:

TABLE 4. Classification of Household Characteristics

Income	Age of Head	Persons	Workers	Children
Less than $5	15 to 24	1	0	0
$5,000 to $9,999	25 to 34	2	1	1 or more
$10,000 to $14,999	35 to 44	3	2 or more	
$15,000 to $24,999	45 to 54	4		
$25,000 to $34,999	55 to 64	5 or more		
$35,000 to $49,999	65 to 74			
$50,000 to $74,999	75 or over			
$75,000 to $99,999				
$100,000 or more				

Employment was classified using 2-digit standard industrial classification codes, grouped into sectors that are generally consistent with those used in the LCOG transportation models. These are shown below in Table 5.

TABLE 5. Employment Sector Classification

Standard Industrial Classification Codes	Sector Description
99 – 999	Agriculture
2400 – 2499	Lumber and Wood
2500 – 2599, 3200 – 3999	Other Durable
2000 – 2099	Food Products
2100 – 2399, 2600 – 3199	Other Nondurable
1500 – 1799	Construction
1000 – 1499	Mining
4000 – 4999	Transportation
5000 – 5199	Wholesale Trade
5200 – 5999	Retail Trade
6000 – 6999	Fire
7000 – 8199	Services
8200 – 8299	Education
9000 – 9999	Government

Results from this processing of parcel data for the Eugene-Springfield application of the model are shown in Figures 6 and 7. The housing distribution in the 1994 database is shown in Figure 5, with traffic analysis zones and the Urban Growth Boundary superimposed. Figure 6 depicts the distribution of nonresidential square footage in 1994, and conveys a pattern of decentralized employment centers beyond the central business district.

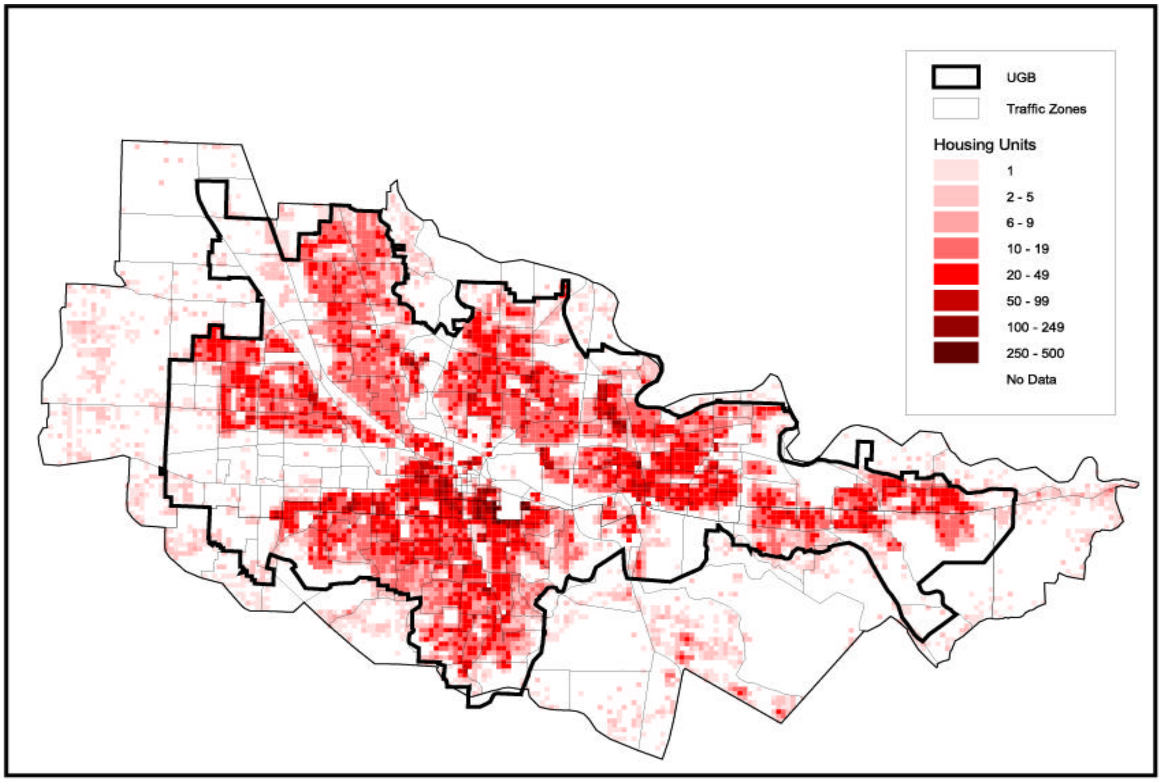

FIGURE 6. Housing Units by Grid Cell in 1994

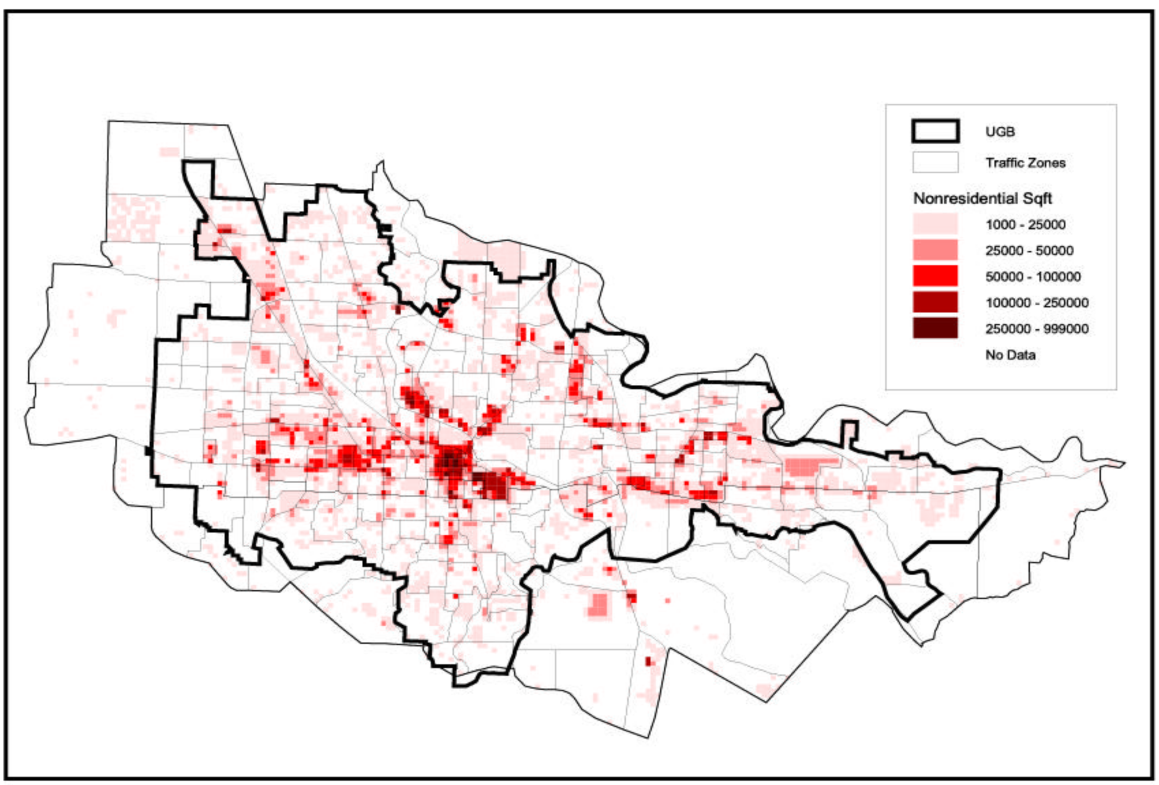

FIGURE 7. Nonresidential Square Footage by Grid Cell in 1994

After developing the database and estimating the model parameters using standard statistical software, the performance of the UrbanSim model was evaluated over a historical period in Eugene-Springfield, Oregon. A 1980 database was developed, and the 1994 database that was used to calibrate the model became the observed target for comparison of simulation results. The model was run in annual steps from 1980 to 1994, then compared to the 1994 observed data. Table 6 summarizes the correlation coefficients between the simulated to the observed 1994 data.

TABLE 6. Correlation of Simulated to Observed 1994 Values

	Cell	Zone	Average Over One-Cell Radius
Employment	0.805	0.865	0.917
Population	0.811	0.929	0.919
Nonresidential Sq ft	0.799	0.916	0.927
Housing Units	0.828	0.927	0.918
Land Value	0.830	0.925	0.908

The simulation results are compared to observed data at three units of geography. The unit used in the models is the 150 meter grid cell, although the model results are not generally intended for use at this level of detail. Nevertheless, the model simulation results correlate well after 15 years of simulation to the observed data. Aggregation of the results to traffic analysis zones used in transportation models produced higher correlations, with all but one category above a 0.9 correlation coefficient. Another spatial comparison was made on the grid cells averaged over the cells within one cell radius, and these produced correlations at least comparable to the zonal aggregations.

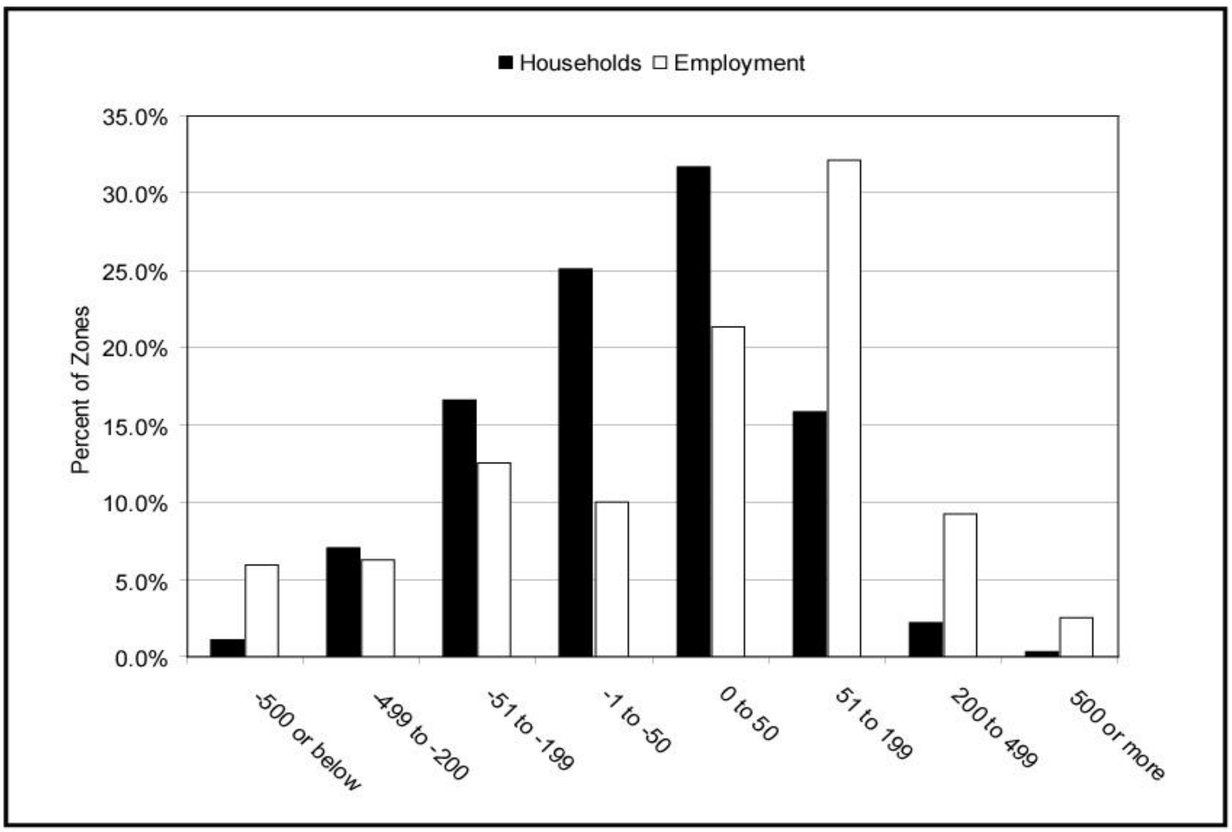

FIGURE 8. Difference Between Simulated and Observed 1980 to 1994 Change by Zone

A more stringent benchmark than the preceding comparison of simulated to observed 1994 values is the comparison of the observed changes from 1980 to 1994 to the simulated changes during this period. Figure 8 portrays this comparison for households and employment, using the 271 Traffic Analysis Zones as the basis for comparison. The graph shows the percentage of zones classified according to the size of the difference between the observed and simulated change in households and employment from 1980 to 1994. For example, the category of *–500 or below* indicates that the model under-predicted the change in households or jobs from 1980 to 1994 by at least 500. Figure 8 shows that the model predicted household change within a range of 50 households for approximately 57% of the zones, and employment change within a range of 50 jobs for 31% of the zones. Household and employment change was predicted within 200 households and jobs for 89% and 76% of the zones, respectively. These results, while mostly encouraging regarding the ability of the simulation to reproduce observed changes over time, nevertheless show that change in some zones was fairly significantly over- or under-predicted. Employment change was not predicted as well as household change, an outcome that is not entirely surprising given that employment tends to be more concentrated and volatile than housing. It should be noted that the model does not contain adjustment factors (sometimes known as K-factors) common in the cross-sectional validation of models to observed data, so these results are obtained purely from the underlying behavior of the model and not from adjustment to correct for errors.

Several factors condition the interpretation of these results. First, Eugene-Springfield is a fairly small metropolitan region, and the model simulated a period during which change was modest, with population in the study area growing from 185,000 in 1980 to just over 200,000 in 1994, while jobs grew from 75,000 to near 100,000. Second, policies such as the Urban Growth Boundary were essentially in place at the beginning of the period, and transportation system changes were relatively minor, so changes resulting from policy interventions were not significant. Third, the difficulties of assembling the base year data for model implementation, which were substantial even with current data, were compounded in the process of assembling a 1980 base year. Archival parcel maps and employment records were difficult to work with, and in the end, considerable error remained in the historical data, diminishing the capacity of the historical validation exercise to inform us about how well the model performed.

The model did not predict isolated events that occurred in the region during this period (nor was it designed to). One was a significant downsizing of a Weyerhouser plant in Springfield, and another was the opening of the Gateway Mall on Interstate 5. These kinds of large-scale events will not be accurately modeled by any model system, and are reminders of the limits of modeling.

Conclusions

The UrbanSim project has made significant progress toward developing models to support land use and transportation planning and growth management. While much has been accomplished, many challenges remain. Further development priorities include:

- Developing software tools for robust data preparation and integration to facilitate applying the system to other areas
- Developing a version of the system that would be suitable for classroom use in courses dealing with urban development, transportation and environmental planning
- Adding environmental components to simulate land cover change, water demand, and nutrient emissions (Alberti and Waddell, 2000; Waddell and Alberti, 2000)

- Adding an evaluation component that computes pre-defined indicators and allows users to construct new ones, and provides useful visualizations of them across multiple scenarios
- Adding an economic evaluation capacity that supports cost-benefit analysis and least-cost planning, incorporating social and environmental externalities and equity considerations
- Developing a web interface for distributed access to the system over the Internet, to support coordination of the model application across local governments within a region, and to provide public access to participate in scenario development and evaluation
- Adding microsimulation of demographic processes (household change)
- Adding a more behavioral real estate development model that represents the roles of landowners, lenders, investors, and specialized developers
- Integrating an activity-based travel model
- Integrating a macroeconomic model that reflects the potential macro-effects of local choices about major infrastructure and land policies
- Developing more robust methods for calibrating and validating the system, incorporating uncertainty about models and data more explicitly
- Leveraging Open Source development to enlist collaborators in the development of new tools and the application of the system in other areas.

Metropolitan land use, transportation and environmental planning must be more effectively integrated than has traditionally been the case. This integration requires robust analytical methods, and should be open to public scrutiny and deliberation in ways that have not been accomplished in the past. Simulation models can and should be part of this deliberative policy process, but they will have to come out of the 'black box' and become instruments that facilitate discussion between local governments and their constituents. The challenge of balancing multiple objectives and agendas within urban areas in the U.S. and abroad have grown increasingly intractable politically, and this work represents a small effort to contribute to more deliberative and informed metropolitan governance. It is, in closing, only one step forward. What lies ahead is a challenging agenda to refine the analytical tools for metropolitan and local planning, make them more accessible and robust, and to generate collaboration in the development and use of planning methods such as these to help communities that want to grow smarter.

Acknowledgements

The development of the current version of UrbanSim has been funded in part by the National Science Foundation, grant CMS-9818378, "Reusable Modeling Components for Simulating Land Use, Transportation and Land Cover," and by the University of Washington, Puget Sound Regional Synthesis Model project. We also wish to acknowledge the sponsors of earlier development of the system, including Oahu Metropolitan Planning Organization, Oregon Department of Transportation, National Cooperative Highway Research Program 8-32(3), "Integration of Land Use Planning and Multimodal Transportation Planning," and The Governor's Office of Planning and Budget, State of Utah. We particularly want to acknowledge the assistance of Bud Reiff of the Lane Council of Governments.

The development of the current version of UrbanSim was undertaken by an interdisciplinary team at the University of Washington led by Paul Waddell, Associate Professor of Public Affairs and of Urban Design and Planning, in coordination with Alan Borning, Professor of Computer Science and Engineering, Marina Alberti, Associate Professor of Urban Design and Planning, and Scott Rutherford, Professor of Civil and Environmental Engineering. Michael Noth developed the UrbanSim software architecture, and Nathan Freier and Michael Becke developed much of the current model implementation and the user interface. Gudmundur Ulfarsson provided substantial assistance with the calibration of the system.

References

Alberti, M. and P. Waddell. (2000). An Integrated Urban Development and Ecological Simulation Model, *Integrated Assessment* **1**(3) 215-227.

Anas, A. (1982) *Residential Location Markets and Urban Transportation: Economic Theory, Econometrics, and Policy Analysis with Discrete Choice Models*. Academic Press: New York.

Anas, A. (1987). *Modeling in Urban and Regional Economics*. New York, Harwood Academic Publishers.

Beckman, R. J., et al. (1995). *Creating Synthetic Baseline Populations*. Transportation Research Board Annual Meeting, Washington, D.C.

Ben-Akiva, M. and S. Lerman (1987) Discrete Choice Analysis: Theory and Application to Travel Demand. The MIT Press: Cambridge, MA.

Blizzard, M. (1996) *Creating Better Communities: The LUTRAQ Principles*. Sensible Transportation Options for People.

Calthorpe, P. (1983) *The Next American Metropolis: Ecology, Community and the American Dream*. New York: Princeton Architectural Press.

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

Dowling, R., R. Ireson, A. Skabardonis, D. Gillen, P. Stopher,.A. Horowitz, J. Bowman, E. Deakin, and R. Dulla. (2000) Predicting Short-Term and Long-Term Air Quality Effects of Traffic-Flow Improvement Projects. NCHRP 25-21. Transportation Research Board.

Downs, A. (1992). *Stuck in Traffic: Coping with Peak-hour Traffic Congestion*. Washington, D.C.: The Brookings Institution and Cambridge, MA: The Lincoln Institute of Land Policy.

Downs, A. (1994). *New Visions for Metropolitan America*. Washington, D.C.: The Brookings Institution.

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

Fotheringham, A.S. and M.E. O'Kelly (1989). *Spatial Interaction Models: Formulations and Applications*. Boston: Kluwer Academic Publishers.

Garrett, M. and M. Wachs (1996). *Transportation Planning on Trial: The Clean Air Act and Travel Forecasting*. Thousand Oaks, CA: Sage Publications.

Handy, S. (1993). Regional Versus Local Accessibility: Implications for Nonwork Travel. *Transportation Research Record*. 1400.

Harris, B. (1985). Urban Simulation Models in Regional Science. *Journal of Regional Science* 25, 4: 545-567.

Kain, J. F. (1985) Computer simulation models of urban location, in *Handbook of Regional and Urban Economics, Volume II*, Edited by E. S. Mills, Elsevier Science Publishers B.V.

Landis, J. (1994). The California Urban Futures Model: a new generation of metropolitan simulation models. *Environment and Planning B* 21: 399-420.

Landis, J. (1995). Imagining Land Use Futures: Applying the California Urban Futures Model. *Journal of the American Planning Association* 61, 4: 438-457.

Landis, J. and M. Zhang. (1998a). The second generation of the California urban futures model. Part 1: Model logic and theory. *Environment and Planning B: Planning and Design*, 25: 657 - 666.

Landis, J. and M. Zhang. (1998b). The second generation of the California urban futures model. Part 2: Specification and calibration results of the land-use change submodel 25: 795 - 824.

Lee, D. (1973). Requiem for Large Scale Urban Models. *Journal of the American Institute of Planners* 39, 3:163-178.

Lee, D. (1994). Retrospective on Large-scale Urban Models. *Journal of the American Planning Association* 60, 1: 35-40.

Martinez, F. (1992). The Bid-Choice Land Use Model: An Integrated Economic Framework. *Environment and Planning A* 24: 871-885.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*. P. Zarembka. New York, Academic Press.

Miller, E. J., D. S. Kriger and J. D. Hunt (1998). *Integrated Urban Models for Simulation of Transit and Land-Use Policies*, Transit Cooperative Research Project, National Academy of Sciences.

Noth, M, A. Borning, P. Waddell (2000). An Extensible, Modular Architecture for Simulating Urban Development, Transportation, and Environmental Impacts. Submitted for publication; draft available from http://www.urbansim.org.

Orfield, M. (1997). *Metropolitics: A Regional Agenda for Community and Stability*. Washington, D.C.: Brookings Institution Press and Cambridge, Mass: The Lincoln Institute of Land Policy.

Parsons Brinckerhoff Quade and Douglas, I. (1998). *Land Use Impacts of Transportation: A Guidebook*, Transportation Research Board, National Research Council.

Paulley, N. J. and F. V. Webster (1991). Overview of and international study to compare models and evaluate land use transport policies. *Transport Reviews* 11: 197-222.

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

Southworth, F. (1995). *A Technical Review of Urban Land Use--Transportation Models as Tools for Evaluating Vehicle Reduction Strategies*. Oak Ridge, TN, Oak Ridge National Laboratory.

U.S. EPA (2000) Projecting Land-Use Change: A Summary of Models for Assessing the Effects of Community Growth and Change on Land-Use Patterns. EPA/600/R-00/098. U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, OH. 260 pp.

U.S. Census (2000) Census 2000 Summary Tape File 1, U.S. Census Bureau. Available online at http://factfinder.census.gov.

Waddell, P. (2000a). A behavioral simulation model for metropolitan policy analysis and planning: residential location and housing market components of UrbanSim. *Environment and Planning B: Planning and Design* 27, 2: 247 – 263.

Waddell, P. (2000b). Monitoring and Simulating Land Capacity at the Parcel Level, in *Monitoring Land Supply with Geographic Information Systems: Theory, Practice and Parcel-Based Approaches*, Vernez-Moudon, A. and M. Hubner, eds., John Wiley & Sons, Inc.: New York.

Waddell, P. (2000c) Towards a Behavioral Integration of Land Use and Transportation Modeling, resource paper presented at the 9th International Association for Travel Behavior Research Conference, Queensland, Australia, July 2000.

Waddell, P. 2001. Between Politics and Planning: UrbanSim as a Decision Support System for Metropolitan Planning, in *Building Urban Planning Support Systems: Combining Information, Models, Visualization*, Klosterman, R. and R. Brail eds. Center for Urban Policy Research and Environmental Systems Research Institute.

Waddell, P. and M. Alberti (2000) Integrated Simulation of Real Estate Development and Land Cover Change, Paper prepared for the 4th International Conference on Integrating Geographic Information Systems (GIS) and Environmental Modeling, Banff, Canada, September, 2000.

Waddell, P., A. Borning, M. Noth, N. Freier, M. Becke, G. Ulfarsson. (forthcoming). UrbanSim: A Simulation System for Land Use and Transportation. *Networks and Spatial Economics*.

Waddell, P., T. Moore and S. Edwards (1998). Exploiting Parcel-Level GIS for Land Use Modeling, *Proceedings of the 1998 ASCE Conference on Transportation, Land Use and Air Quality: Making the Connection*, Portland, Oregon, May 1998.

Waddell, P. and T. Moore (forthcoming). Forecasting Demand for Urban Land, in *Land Market Monitoring for Smart Urban Growth*, G. Knaap ed., Lincoln Institute for Land Policy.

Weatherby, C. (1995) Summary of Workshop Observations and Recommendations, in G. Shunk, P. Bass, C. Weatherby and L. Engleke (eds.) *Travel Model Improvement Program Land Use Model Conference Proceedings*, Washington, D.C.: U.S. Department of Transportation.

Wegener, M. (1994). Operational Urban Models: State of the Art. *Journal of the American Planning Association* 60, 1: 17-30.

Wegener, M. (1995). *Current and Future Land Use Models*. Travel Model Improvement Program Land Use Model, Dallas, TX, U. S. Department of Transportation.

Wilson, A.G. (1967). Statistical Theory of Spatial Trip Distribution Models. *Transportation Research* 1: 253-269.