Design of a Statewide Land Use Transport Interaction Model for Oregon

JD Hunt
Department of Civil Engineering, University of Calgary
2500 University Drive NW
Calgary, Alberta, CANADA T2N 1N4   email: jdhunt@ucalgary.ca

R Donnelly
PbConsult
5801 Osuna Road NE, Suite 200
Albuquerque, New Mexico, USA 87109   email: rdonnelly@pbtfsc.com

JE Abraham
Department of Civil Engineering, University of Calgary
2500 University Drive NW
Calgary, Alberta, CANADA T2N 1N4   email: jabraham@ucalgary.ca

C Batten
EcoNorthwest
888 SW Fifth Avenue, Suite 1460
Portland OR 97204   email:batten@portland.econw.com

J Freedman
PbConsult
400 SW Sixth Ave, Suite 802
Portland Oregon,, USA 97204   email: freedman@pbworld.com

J Hicks
PbConsult
5801 Osuna Road NE, Suite 200
Albuquerque, New Mexico, USA 87109   email: jhicks@pbtfsc.com

PJ Costinett
PbConsult
8011 NE 121st St
Kirkland, Washington, USA 98034   email: pcostinett@worldnet.att.net

and

WJ Upton
Planning Section, Oregon Department of Transportation
555 13th Street, Suite 2
Salem, Oregon, USA 97301-4178   email: William.J.UPTON@odot.state.or.us

ABSTRACT

This paper describes the design of a land use transport interaction model of the State of Oregon in the United States. The model framework itself is a combination of 7 connected modules, some aggregate representations relying on equilibrium solutions and others fully dynamic dis-equilibrium agent-based micro-simulations. The system evolves through time in discrete year-by-year steps. Household demographic changes, residential location decisions, employment and associated training choices and daily activity patterns are represented using micro-simulation. Travel tours by household members and commodity movements are also micro-simulated. Each trip is loaded to the transport supply networks link-to-link with randomly assigned utility function sensitivities that allow a dispersion of travel choices with the use of minimum path assignment. A generalized form of spatially-disaggregated input-output model is used to identify the patterns of spatial locations and interactions among the different economic sectors, including the interfaces with households, in an aggregate treatment. Area-wide levels of production activity and population in-migration are established using an aggregate regional economic model. A Monte Carlo treatment of developer actions in small area grid cells is used to simulate changing development patterns.

KEYWORDS

Transportation Land Use Interaction; Spatial Economic Modeling; Agent-Based Micro-simulation; Equilibrium and Dis-Equilibrium Modeling; Micro-Assignment of Traffic

1.0 Introduction

For a number of years now the State of Oregon Department of Transportation has been pursuing an ambitious program of improving their transportation and related planning models. This program of improvement has included the development of 1st generation land use transport interaction models covering both the urban and the regional (statewide) scales. The 1st generation work has now come to a close and work has started on a 2nd generation model, which is intended to benefit from the lessons learned in the 1st generation work and from the very fast moving recent and anticipated advances in low cost computing capabilities.

This paper describes the design of the 2nd generation integrated land use transport interaction model, presenting elements of the design itself – some of them particularly novel – and outlining the various approaches used. The specified model requirements are listed in Section 2. The model design - in terms of its treatment of space, time and system behavior – is described in Section 3. Elements of the larger modeling system architecture are considered in Section 4, including the ‘data store’, ‘process controller’, ‘calibrator’ and the user interface components. Some issues concerning implementation of the model are covered in Section 5, and conclusions are offered in Section 6.

At the time of writing work on the development of the model is still underway, but some preliminary testing has been done that establishes the validity of many of the model elements.

2.0 Model Requirements

The intention is to provide a model that can adequately and practically support the analysis of a diversity of relevant issues, impacts and policies. This leads to the specification - based in part on the experience gained in the 1st generation work and in work done elsewhere - that the 2nd generation model:

- Should operate at a geographic scale roughly equivalent to traffic analysis zones in the metropolitan areas and census tracts outside of them - in order to provide a treatment of space that is sufficiently disaggregate for policy guidance.
- Must provide an integrated representation of the land use, transport and economic components - in order to provide an internally consistent treatment of assumptions and forecasts and provide a representation that is sufficiently comprehensive in its scope. Environmental components should also be integrated as well, although the extent and manner of doing so is not clear at this time.
- Should be dynamic – thereby representing processes that result in the evolution of the system through time.
- Should not rely unduly on an equilibrium formulation – and thus could use a dis-equilibrium treatment of land markets and activity interactions together with an equilibrium treatment of transportation and commodity markets – in order to provide a treatment of temporal dynamics that is sufficiently disaggregate for policy guidance.
- Should be activity-based – in order to be sufficiently behavioral in the representation of system responses to policy inputs.
- Must be affordable in terms of both time and money resource requirements.

3.0 Model Design

3.1. Treatment of Space

The model covers the entire State of Oregon and a ‘ring’ for about 50 miles just beyond the state boundaries to the north, east and south as indicated in Figure 1. This area is divided into a co-ordinated system of zones, link tributary areas and grid cells as depicted in Figure 2. These different divisions of space and related network conditions are used

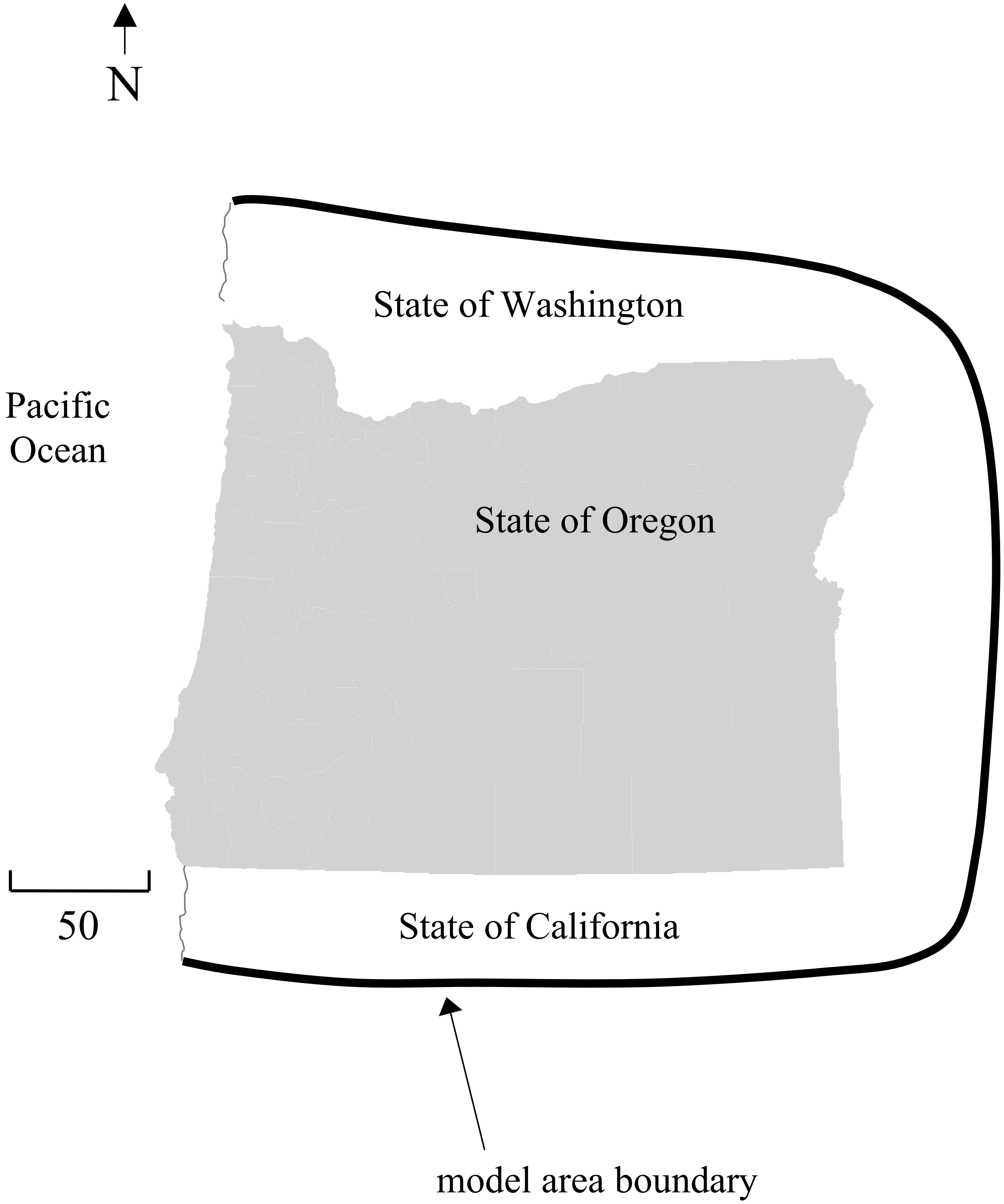


Figure 1: Area covered by model

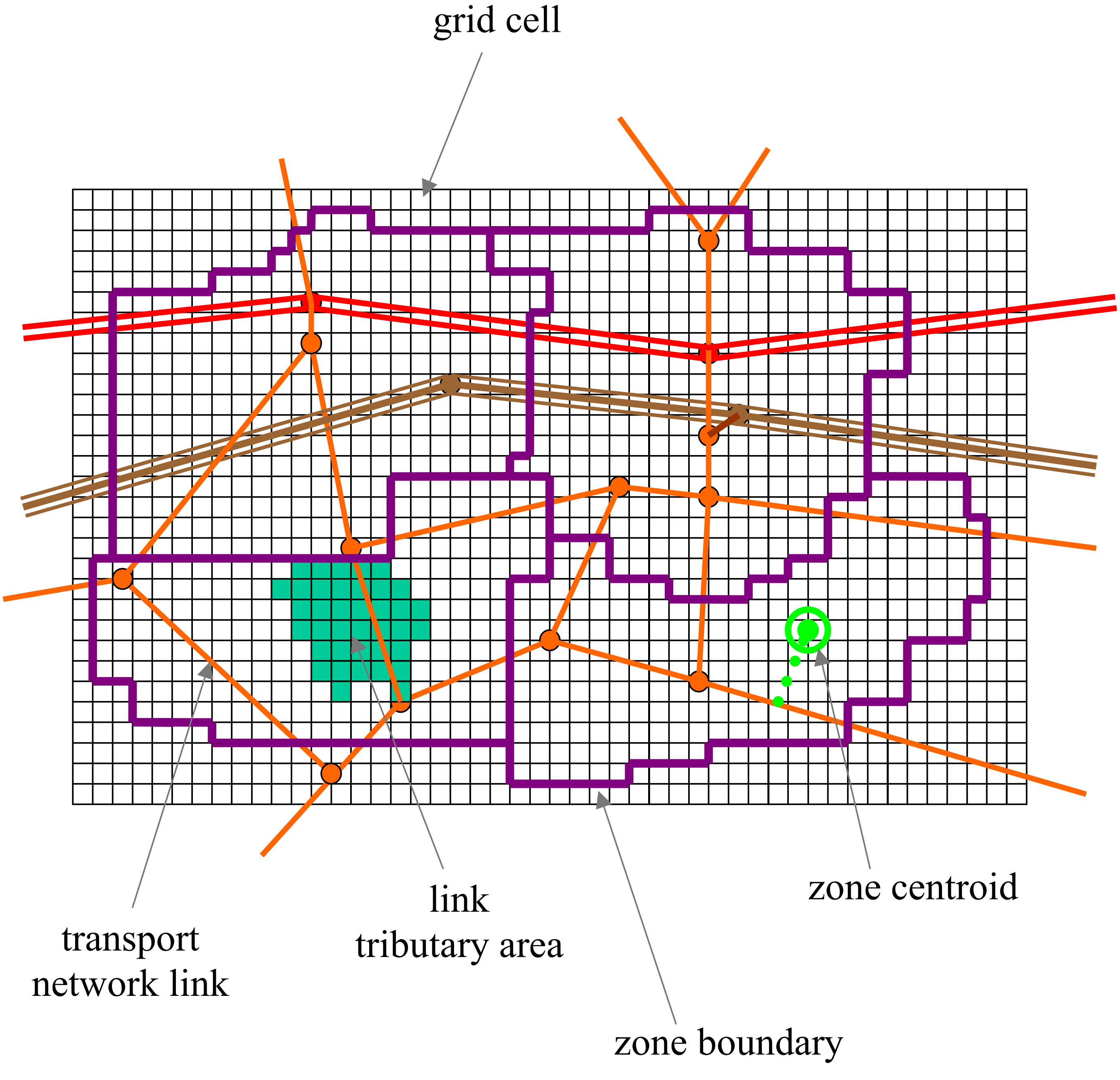


Figure 2: Illustration of co-ordinated system of zones, grid cells and link tributary areas used in the treatment of space in the model

for different parts of the model, reflecting the different perceptions and processes involved.

Each of these three forms of division covers the entire model area. The grid cells are the smallest, and they nest completely into both of the other two. That is, all zone boundaries and link tributary areas are consistent with grid cell boundaries and there is a whole number of grid cells in each zone and link tributary area.

The zones are roughly on the order of traffic analysis zones, and are connected to the transport supply networks using centroid connectors. There is a total of about 3,000 zones covering the entire model area.

The grid cells are squares of land sufficiently small enough that just one type of developed space (one category of building floorspace) can be attributed to a given cell without substantial loss of accuracy. The current thinking is that 30m x 30m cells within and near to built-up areas and 300m x 300m (or even larger) cells in more wide-open spaces are acceptable in this regard. The current estimate is that there is a total of about 14.5 million grid cells covering the entire model area.

The link tributary areas are the areas containing the origins and destinations of the trips feeding the corresponding links in the transport system, and there is a separate, mutually exclusive and collectively exhaustive set of tributary areas for the set of links constituting a network for each mode.

3.2. Treatment of Time

The model steps through time in a series of one-year steps that allow the entire system to evolve as indicated in Figure 3. The representation for year t+1 is influenced in part by the conditions determined for year t, which provides for explicit representation of various lagged effects and system inertia.

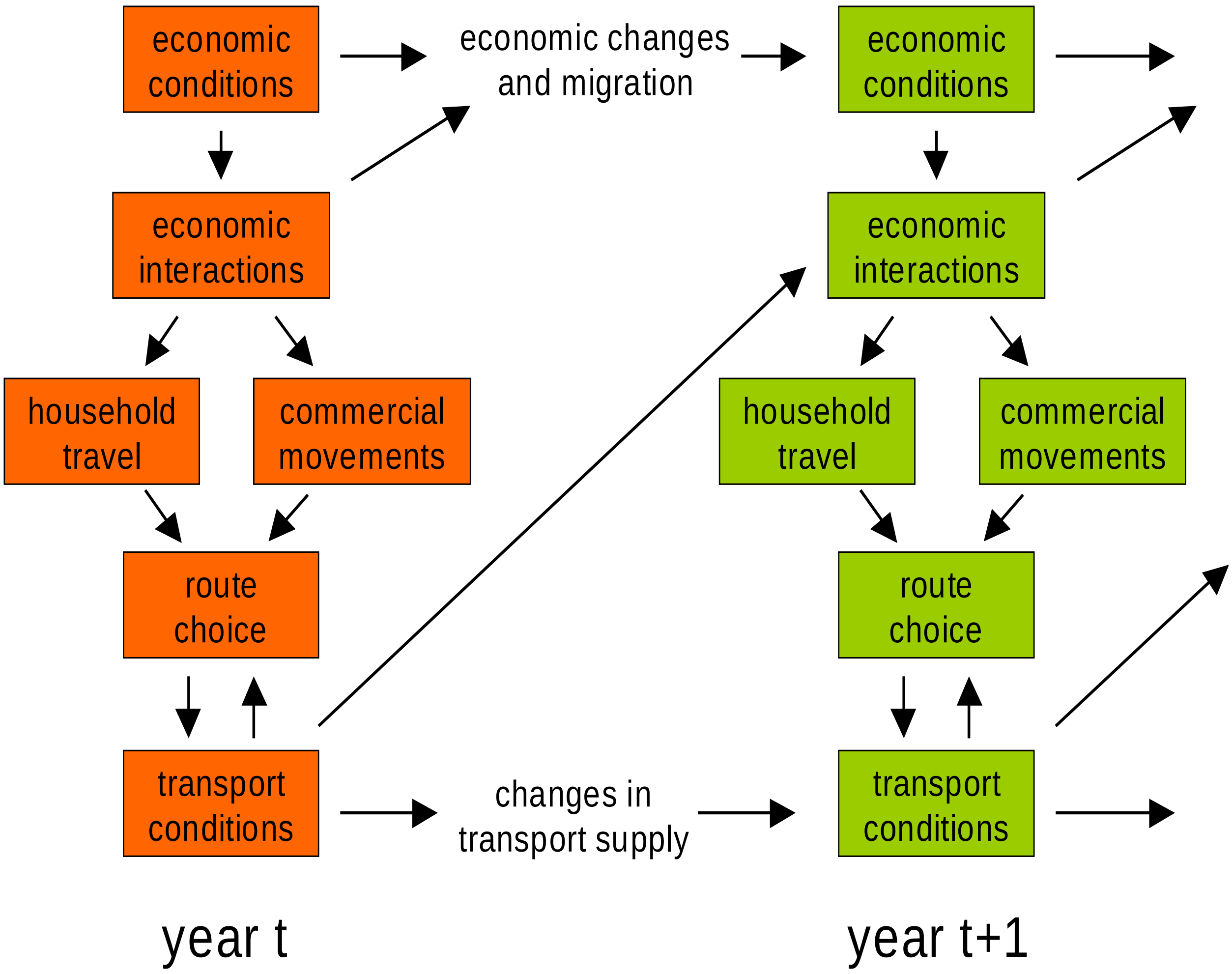


Figure 3: Process used to step the model through time

As explained below, an equilibrium is used for some aspects of the representation but not others. In addition, some of the simulation processes within the model involve time steps much shorter than one year, and therefore are performed multiple times during a given 1-year step.

3.3. Representation of System Behaviour

The behavioral representations in the model are based on utility values. In particular, the aggregate quantities considered in the aggregate components of the model are allocated among alternatives according to utility values and the individual actors considered in the disaggregate components of the model react to utility values.

Five different categories of utility value are defined for use in different components of the model. These differ in terms of how they are determined and how they are used.

With regard to how they are determined: a particular utility value is the dependent variable of a utility function that operates on sensitivities together with attribute values, as follows:

$$U_{i,a} = f\,(\,\alpha_i\,,\,X_a\,)$$

where:

i = index representing individuals
a = index representing alternatives
$U_{i,a}$ = utility determined for alternative a for individual i
α_i = vector of utility function coefficients indicating sensitivities of individual i to attributes of alternative a
X_a = vector of attribute values for alternative a.

The five categories of utility value are as follows:

- 'Rutility' values: These are used for allocations of aggregate quantities. The attribute values, X_a, are average, zonal or typical values. The sensitivity values, α_i, are typical values for the category of aggregate quantity being allocated. The 'R' placed before 'utility' stands for 'representative'.
- 'Zutility' values: These are used for agent-based micro-simulations of individual household and person decisions. The attribute values, X_a, are average, zonal or typical values, but the sensitivity values, α_i, are specific values assigned to the household or person. In this case, differences in the sensitivity values among a set of agents gives rise to dispersion in the behavior of the set. The 'Z' is for 'zonal'
- 'Iutility' values: These are used for network path selection for aggregate, zone-to-zone trip flow assignment. The attribute values, X_a, are specific link values, and the sensitivity values, α_i, are aggregate values assigned to the flow being assigned. The 'I' is for 'interchange'.
- 'Lutility' values: These are used for network path selection for individual trip assignment. The attribute values, X_a, are specific link values, and the sensitivity values, α_i, are specific values assigned to the trip-making agent. The 'L' is for 'link'.
- 'Cutility' values: These are used for micro-simulations of land development decisions. The attribute values, X_a, are specific grid cell values, and the sensitivity values, α_i, are typical values assigned to the developers as a single category. The 'C' is for 'cell'.

The model represents the behavior of the full system using a set of 7 separate but highly connected modules that cover different components of the full system. This set of modules and the flows of information among them are shown in Figure 4.

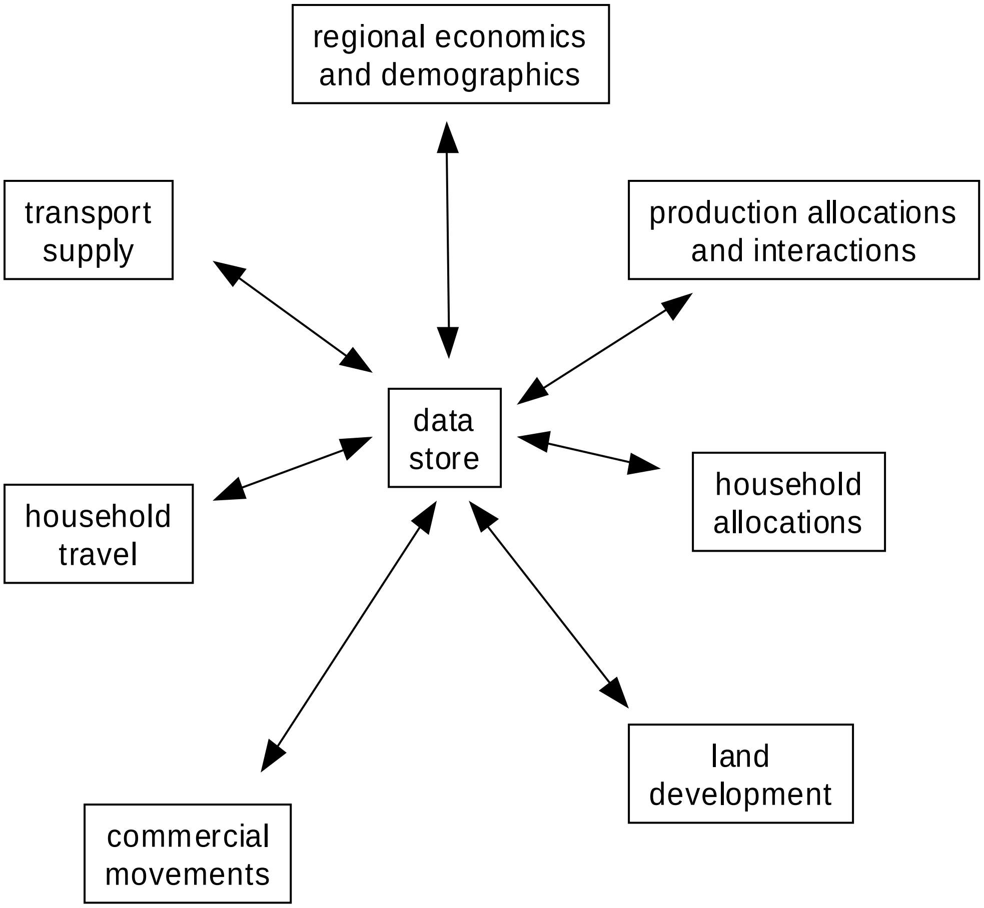


Figure 4: The modules in the model and the flows of information among them

This modular design allows a sort-of ‘plug&play’ upgrading and flexibility. It also allows different approaches to be used for different components, such as:

- fully dynamic vs quasi-dynamic;
- equilibrium vs dis-equilibrium;
- aggregate representation of the behavior of flows vs disaggregate representation of the behavior of individual agents; and
- statistical reproduction of patterns vs simulation of processes.

Each of the modules indicated in Figure 4 is described further below:

3.3.1. Regional Economics and Demographics

The regional economics and demographics module provides the rest of the model with regional control totals for production by economic sector, imports and exports by economic sector, employment by labor category, in-migration, and payroll by sector for each year. It requires exogenous forecasts of United States national production, employment, and population as inputs. It includes simple macroeconomic models of sub-national regions to the north, south, and east of the model area to assist with trade flow predictions.

A combined input-output and econometric model is used, based on the coupling strategy identified by Rey (1998) and various approaches and refinements developed for the Washington Project and Simulation Model by Conway (1990). A set of simultaneous linear equations representing the input-out put structure of the model area economy are used to predict the dollar value of production by sector. Additional linear equations for various components of final demand and for the labor market are solved simultaneously with the input-output equations.

A list of the production sectors included in the model is provided in Appendix A.

Final demand is divided into four categories:

- exports;
- consumption;
- investment; and
- state and local government.

The national and external-region economic forecasts serve as a proxy measure of export demand.

Consumption demand is forecast taking into account demand for motor vehicles, other durable goods, non-durable goods, and services. The motor vehicle demand uses the results for the previous year from the auto ownership model within the household allocation module.

Investment demand has three components:

- residential structures;
- non-residential structures; and
- equipment.

The equations for the residential and nonresidential structures include the results for the previous year from the land development module, which also forecasts redevelopment.

State and local government demand is developed using the results for the previous year concerning education demand and the resulting education operating expenditures as well as results concerning land development. Information about changes to the transportation network that indicate construction activity can also be used to guide an exogenous input to the quantity of state government demand.

Other equations in the regional model include:

- in-migration trends;
- employment by labor category; and

- labor-force participation.

Trade flows between the model region and other regions are forecast by allocating gross imports and exports by economic sector to the regions using the method developed for the production interactions module, which is described below.

3.3.2. Production Allocations and Interactions

The production allocations and interactions module determines for each year:

- The distribution of production activity among zones;
- The consumption of space by these production activities;
- The flows of goods and services and labor from the location (zone) of production to the location (zone) of consumption; and
- The exchange prices for goods and services, labor and space.

This is done using an aggregate treatment, where the quantities of production, the consumption of space, the flows of goods and services and labor (called 'commodities') and the exchange prices are determined for the entire economy, with the households located as determined in the household allocation module in the previous year.

The theoretical basis of the aggregate treatment of the economy is an extended form of the spatially-disaggregated input-output approach used in MEPLAN (Hunt and Simmonds, 1993) and TRANUS (de la Barra, 1998). Figure 5 provides a depiction of the approach. Production activity by economic sector is quantified in terms of dollar value of production, and production functions based on the technical coefficients in the input-output 'make' tables are used to determine quantities of input commodities required, including goods and services, labor and space. Commodities flow from production locations to 'exchange locations' to consumption locations. Exchange locations represent the places where exchange prices are determined and where commodities are transferred from the seller to the buyer, thereby allowing an allocation of transport costs. Flows of commodities are allocated from production locations to exchange locations, and from

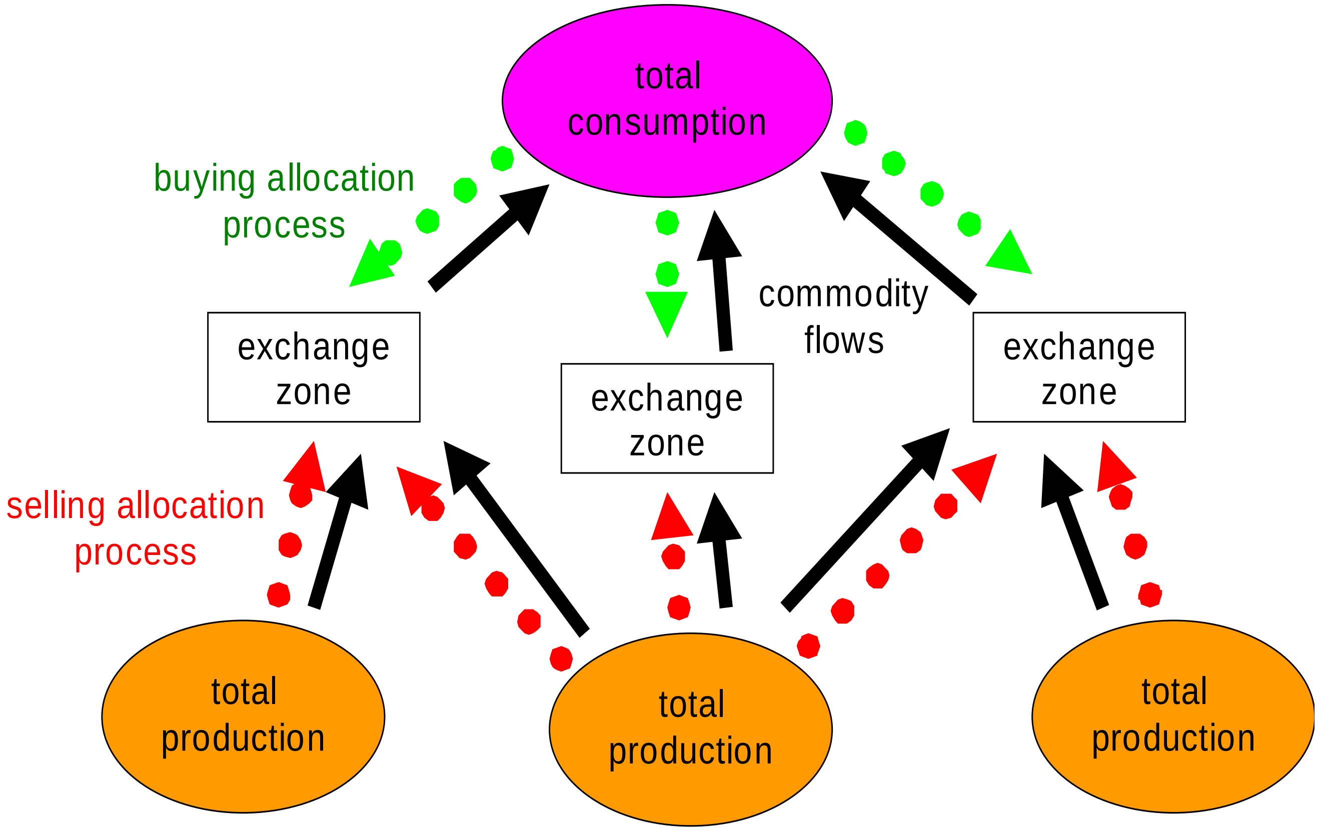


Figure 5: Depiction of the extended form of spatially-disaggregated input-output approach used in the production allocations and interactions module

exchange locations to consumption locations according to logit formulations that take into account the exchange prices and the relevant transport costs. This is in contrast to the treatment in the standard spatially-disaggregated input-output approach, where it is the demand for additional production that is allocated, rather than the flows of commodities, and where the prices are always determined at the production location. The introduction of the concept of exchange locations in this way permits a much more flexible range of possible representations, and allows a much more realistic treatment of labor markets in particular.

A list of the commodity categories included in the model is provided in Appendix B.

For all economic sectors other than households, the total production activity in the sector over the entire model area for each year is determined in the economic and demographic module. This total is allocated among the zones in the model area using a logit formulation with a utility function that includes terms representing the impacts of:

- Lags and inertia;
- The Rutility (including transport costs and exchange prices) associated with selling output commodities, based on the zone-to-zone attributes for a representative time slot and fixed utility function coefficients;
- The Rutility (including transport costs and exchange prices) associated with buying input commodities;
- The utility of business travel; and
- Location-specific taxes and subsidies.

This is a further generalization of the approach used in standard spatially-disaggregated input-output modeling, where production activity is allocated from the zone where the demand for this activity arises. This further generalization is important because it allows a consistent interpretation - based on random utility theory - of the (composite) utility of production activity in each zone and for the entire model area, which can be used in assessment of policy alternatives and in the forecasting of net in-migration of production activity to the entire model area.

During the allocation of production activities, households stay in the zones where they were located in the household allocation module in the previous year. But the flows of labor produced by these households are allocated to exchange locations as part of the allocation of production activities. Similarly, the flows of commodities consumed by households are allocated from the exchange locations as part of the allocation.

The consumption of goods and services and labor and the consumption of space in zones by production activities is elastic with respect to the buying disutility, which is influenced by the exchange prices that are available. The production of labor by households is also elastic with respect to the selling utility - with a function that constrains the ability of labor supply in each exchange zone to change beyond certain limits in a given year. The exchange price for a commodity at an exchange location is adjusted in response to the levels of aggregate demand versus the aggregate supply at that exchange location. An equilibrium solution is identified by adjusting exchange prices and reallocating production activities and flows of commodities in an iterative process to convergence. It is certain that the equilibrium solution that is identified will not be beyond the capacity of the economic system because the production totals have been established in the economic and demographic module.

3.3.3. Household Allocations

The household allocations module determines for each year:

- Changes in household composition, in the characteristics of household members and in the characteristics of households, including the effects of demographic and economic processes;
- The specific sensitivities (relevant utility function coefficients) of household members and households;
- Household actions regarding home locations; and
- Residential space use quantities and prices.
- The vector of inputs for each household;

- Household car ownership changes;
- The employment/SOC status and workplace location (zone), as relevant, for each household member; and
- The school status and school location (zone), as relevant, for each household member .

This is done using micro-simulation, where there is an explicit consideration of each individual household in each time period. In general, Monte Carlo simulation is used to assign specific conditions to each household, with the selection probabilities for alternative possible conditions based on either (a) exogenously specified conditional distributions or (b) choice probabilities provided by models of choice behaviour that include an explicit representation of influences on decision-making.

The theoretical basis for the representation of household behaviour in the determination of household composition and household conditions is reflected in the processes used to assign the specifics to each household and in the way these processes are combined. In some cases the specifics for each household are assigned by random sampling from aggregate distributions, which is in essence a statistical approach. In some cases the specifics for each household are assigned using probabilities determined by a logit formulation with utility functions, which means that the process can be viewed to have a basis in random utility theory. The particular processes used and aspects of their theoretical basis are described below.

A synthetic population of discrete households must be generated for the first time period (first year). An iterative proportional fitting process developed by Beckmann et al (1996) is used to generate a population of households that respects the aggregate marginal distributions for the study area in each source of data.

Each household in the resident population is then considered in turn.

First the household is taken through demographic transitions including the aging of the household members, births, deaths, the departure of members to form new households and possible dissolution of the household. The transition probabilities are functions of various location and socio-economic factors and are adjusted so as to respect the specified marginal distributions for the total population determined in the demographic component. As part of this process, the income for the household and the qualifications to work in the various SOC categories considered in the model for the household members are also updated to reflect the influences of labor prices and the attractiveness of working in different SOC categories consistent with the results obtained in the economic and the production allocation and activity interaction components of the model. New households are formed using members departing from existing households and the household and personal characteristics for these households are 'filled out' as required by randomly selecting values from known marginal distributions as appropriate.

Then the relevant sensitivities (specific utility function coefficients) concerning household location choices and activity pattern and travel-related choices considered throughout the model for each household and each member in it are updated (or assigned if not previously assigned). The sampling distributions for these sensitivities are specified as part of the model parameter estimation and calibration, enabling an explicit treatment of the distributions of sensitivities within the population. The potential exists for these sensitivities to evolve over time in response to changing conditions and exposures, reflecting changes in attitudes and learning as the system evolves and the population gains further experience.

Then a vector of technical coefficients for inputs is assigned to the household and updates to employment/SOC status and a school status, along with workplace locations (zones) and school locations (zones) as appropriate are assigned to the household members. Changes in car ownership for the household are also assigned. These assignments to households are done using Monte Carlo selection where the selection probabilities are determined using logit formulations with utility values that are functions of the age and gender of the person, the household composition and income and the zonal level attributes of travel and the prices of labor determined in the production allocations and interactions module.

Each household is then taken through a series of decisions concerning residential location, as follows:

- Whether or not to move the location of the primary residence;
- The zone for the new location of the primary residence when there is to be a move;
- Whether or not to have a secondary residence;
- Whether or not to move the location of the secondary residence when there is one; and
- The zone for the location of the secondary residence when there is one and there is to be a move.

The results of these decisions are determined using Monte Carlo selection with the selection probabilities assigned to the available alternatives using logit formulations with utility functions that include indications of the Zutility values associated with alternative zones as residential locations.

Each household undergoing a move of the primary residence is assigned to be moving to a location within the model region or to a location outside the model region. This is done using Monte Carlo selection with the selection probabilities influenced by the utility values associated with alternative residential locations and consistent with the distributions of characteristics determined for the out-migrating population in the demographic component.

A population of discrete in-migrating households is synthesized consistent with the expected marginal distributions for the characteristics of this group using the process developed by Beckmann et al (1996), as was done in the development of the base year population. This population is then added to the set of households moving to primary residence locations within the model region.

As moving households are allocated to the zones containing the selected new residential locations, the quantities of available residential space and the resulting prices for this space are updated and the results of these updates are allowed to impact subsequent location decisions. In particular, a decrease in the supply of space in a zone (decreasing the vacancy rate in the zone) brings about an increase in price in the zone. Subsequent households respond to this price increase in two ways: (1) the location utility for the zone decreases as a result of the increased price and therefore the probability of selecting that zone decreases and (2) the quantity of space used decreases because of a negative elasticity of space demand with respect to price. As the vacancy rate approaches 0 in a zone, the price increases are sufficiently rapid to force all demand to go elsewhere. If the point is reached where there is not sufficient space in the entire model area to accommodate all the households seeking new residential locations, then those households than cannot be accommodated are added to the pool of households seeking new locations in the next year. The set of moving households is considered in this process in random order - in an effort to take into account the very large and complex effects that the temporal ordering of individual decisions has on this process in reality.

3.3.4. Land Development

The land development module determines the changes in space from one year to the next.

The supply of (developed) space in a particular year is fixed, and the other modules operating for the year operate given this fixed supply. These other components determine a price for each category of space in each zone, and the primary task of the land development module is to adjust the quantity of space over time in response to changes in price. This is done in a highly disaggregate manner, one grid cell at a time.

Three decisions are simulated for each grid cell. The first is whether the cell should be redeveloped or not. The second is the choice of the category of space that the cell should be redeveloped into. The third is the quantity of new development in the cell. These three decisions are represented using logit models. The utility functions for each of the alternatives in each choice set are based on rents and vacancy rates established in the equilibrium model.

The land development model is behavioral, simulating the decisions of land owners regarding how to improve their properties. Land owners (developers) make their decisions based on current prices and vacancy rates, leading to behaviour consistent with the notion that developers react to changes in conditions because they expect future conditions to be similar to the most current conditions.

The changes to grid cells are assigned using Monte Carlo selection where the selection probabilities are determined using logit formulations with Cutility values that are functions of the current space type, the number of years since the more recent change, the zonal prices for space of difference categories and both zonal and model-area-wide vacancy rates. This leads to a patchwork of different space types in zones and in link tributary areas that helps simulate the diversity of trip origins and destinations. In addition, the potential also exists to further influence this process and to adjust the results using information about both (a) proximity to specific transportation infrastructure and (b) the content of adjacent cells (possibly applying forms of cellular automata).

A list of the types of developed space included in the model is provided in Appendix C.

For certain mixed use developments, such as 'office with retail' and 'multi-family residential with retail', the cell is split into two cells with the same attributes except that one of the cells contains the office space and associated zoning limits and the other cell contains the corresponding retail space and associated zoning limits; then both cells are considered as with all the other cells

3.3.5. Commercial Movements

The commercial movements module determines the truck movements arising during a particular representative workday for each year.

A fully disaggregate list of truck movements is synthesized, providing the following for each truck movement:

- Vehicle type (light single-unit, heavy single-unit, articulated);
- Starting link;
- Ending link;
- Starting time;
- Commodity carried; and
- Trans-shipment organization.

Commodity flows between aggregations of zone pairs are considered in turn.

For a given commodity flow and aggregation of zone pairs, a shipment size of that commodity is randomly selected from the Commodity Flow Survey dataset and the flow is reduced by this amount. The vehicle type, starting time, and trans-shipment information for the shipment is synthesized taking into account similar shipments and vehicle movements being made nearby and this information is added to the list being generated by the model.

The origin and destination of the movement are attributed to particular grid cells in each zone by randomly selecting from the set of cells with space types consistent with the commodity being shipped. In this way the link for the start of the trip and the link for the end of the trip are identified for each trip, as required by the transport supply module.

This continues until there is no further commodity flow for the aggregation of zone pairs and consideration then proceeds to the next interaction and eventually to the next commodity.

3.3.6. Household Travel

The household travel module establishes a list of the specific individual trips made by members of households during a particular representative workday for each year, providing the following for each trip:

- Starting link;
- Ending link;
- Starting time;
- Tour mode;
- Vehicle occupancy if automobile mode;
- Utility attribute coefficients defining personal preferences of traveler(s) for this trip; and
- Non-network related utility components.

The process starts by assigning each household member an activity pattern for the day. The activity pattern is a listing of the sequence of activities undertaken by the household member as a series of tours made out from the home (and from the workplace as appropriate). For example, one such activity pattern is ‘Home-Shop-Home-Work-Lunch-Work-Recreation-Home’, which includes two ‘home-based’ tours and one ‘work-based’ tour. Each household member is assigned an activity pattern using a Monte Carlo process where the selection probabilities are based on a logit function using Zutility values as inputs. These Zutility values are influenced by the age and gender of the household member, household income, work and school status, expenditure level and transport accessibilities at both the home location and the workplace location as appropriate. The data required to develop these models were obtained in large-scale, special-purpose surveys of the activity patterns and travel behavior of household members over two consecutive days for households drawn throughout Oregon.

The activities in a given pattern are then assigned durations using a Monte Carlo process, with the selection probabilities based on hazard models that are functions of the general nature of the rest of the activity pattern assigned the person, the age and gender of the person, household income, work and school status, expenditure level and transport accessibilities at both the home location and the workplace location as appropriate.

Following this, each home-based tour and work-based tour is considered separately. The tour is assigned a ‘primary’ destination location (zone) using a Monte Carlo process with the selection probabilities determined using logit formulations with Zutility values that are functions of the type of activity, the zonal level attributes of travel and related accessibilities. After the ‘primary’ destination is assigned, then the tour mode is assigned, again using a Monte Carlo process with the selection probabilities determined using logit formulations with utility values that are functions of the travel attributes for the round trip from the zone containing the home (or workplace) to the zone containing the ‘primary’ destination.

The following tour mode alternatives are considered:

- ‘driver’ - drives for all trips on tour;
- ‘cyclist’ - cycles for all trips on tour;
- ‘passenger-AA’ - auto passenger for all trips on tour;
- ‘passenger-AT’ - auto passenger before primary destination and transit passenger after;
- ‘passenger-TA’ - transit passenger before primary destination and auto passenger after;
- ‘passenger-TT’ - transit passenger for all trips ‘passenger-PP’ - transit passenger with auto access to and from base of tour; and
- ‘walker’ - walks for all trips.

The networks of links available for the individual trips in the tour, considered in the transport supply module, are dictated by the tour mode that is assigned. For example, if the tour mode is auto drive alone for both the outbound and return components of the tour, then just the road network links are available for each of the trips in the tour.

The number and locations (zones) of any intermediate stops are then assigned to each tour, up to a maximum of one for each of the outbound and return portions of the tour, again using a Monte Carlo process with the selection probabilities determined using logit functions with Zutility values that reflect the nature of the tour, the tour mode and the accessibility at the home and workplace as appropriate. For each auto trip, including both driver and passenger tour modes, a further specification of the number of people in the vehicle is performed using a Monte Carlo process with the selection probabilities determined using logit formulations with Zutility values that are functions of the relevant travel attributes. This is in order that the trip assignment process can make the right set of link types (with or without HOV) available for the specific trips in each case.

The start time of each trip is then established according to the durations previously assigned the corresponding activities. The origin and destination of each trip are attributed to particular grid cells in each zone by randomly selecting from the set of cells with space types consistent with the activity at the stop. In this way the link for the start of the trip and the link for the end of the trip are identified for each trip, as required by the transport supply module.

3.3.7. Transport Supply

The transport supply module determines transportation network loadings and transportation network attributes given transport demands from the household travel and commercial movements modules. This is done as a ‘micro-assignment’, at the level of individual vehicles and individual travelers (using public transportation, using their own personal vehicle, or grouped with other travelers sharing a vehicle).

The primary mode of the trip dictates which network links are available for the trip. Road-based vehicle trips are loaded to the road network and person-based transit trips are loaded to the transit service network and related walking networks separately and then the results are combined by adding the road-based transit vehicles to the road network. Walk-based and cycle-based trips are loaded to walk and cycle networks derived from the road network using link distances, with no impact on congestion levels.

For the loading to the road network, the vehicle trips (including both those established in the household allocations module and those established in the commodity movements module) are first organized into an aggregate zone-to-zone trip table and this used to establish an aggregate solution that becomes the starting point for the micro-assignment. For the aggregate solution, the Frank-Wolfe assignment process (Ortúzar and Willumsen, 1990) is used (with aggregate flow assignment sensitivities used to develop Iutility values for the links) for a sufficient number of iterations to ensure that an acceptable level of equilibrium convergence has been reached. Initial testing has indicated that this takes about 15 iterations. Based on this equilibrium set of link flows and travel times, the micro-assignment procedure is implemented. Micro-assignment involves loading the individual trips (with individual route choice sensitivities represented in the Lutility values calculated for the links) from specific origin point locations to destination point locations one at a time in order to obtain a dispersion of trips consistent with the range of sensitivities in the population and in order to avoid artificial overloadings of trips at the points where the centroid connectors join the road network. To ensure conservation of the total number of trips and the corresponding network loadings, it is necessary to remove the portions of a single trip from paths built during the Frank-Wolfe solution process when loading a new trip in micro-assignment. The result is a detailed all-or-nothing assignment of trips based on equilibrium prevailing network conditions.

The specific procedure has the following steps:

1. Define traffic analysis zone system with zonal centroid connector links for network;

2. Aggregate trip list to O/D matrix consistent with traffic analysis zones defined;
3. Solve a conventional Frank-Wolfe solution procedure for user-optimal network equilibrium link flows based on generalized link disutilities to an acceptable level of solution convergence;
4. Begin with first trip in trip list;
5. Select trip origin and destination locations and utility coefficients from the trip list;
6. Compute disutility (negative Lutility) for each network link;
7. Compute shortest disutility path from origin link to destination link and save path,
8. For each link in shortest path, increment link volume by one;
9. For each link in each shortest path used in the Frank-Wolfe solution (step 3) between the TAZs in which the origin and destination links are located, remove the portion of O/D flow allocated to each respective shortest path;
10. If trips remain in trip list, repeat steps 5 - 9, otherwise stop.

Road network assignment is typically the most computationally burdensome component of running an integrated land use transport model, and this model is no exception. The model considers as many as 12 million trips in its list each year, which requires several hours in the current setup. This leads to rather long run times overall, and efforts are underway to identify ways of reducing these run times, including parallel processing, reorganization of the computer code, and the possibility of running the transport assignment less frequently than every year.

The loading to the transit network is done using a form of the optimal strategies procedure as implemented in the EMME/2 modeling software (INRO, 1996).

With this procedure all ‘reasonable’ paths through the transit network are identified, and the flow of travelers from origin zone to destination zone is allocated among these paths according to their relative attractiveness. It is assumed that travelers arrive at a transit station uniformly during the headway period and board the first reasonable transit route that becomes available. Thus, if competing routes are available, wait time is determined by a composite headway of the competing lines that provide a path through to the final destination.

In contrast to the process with the road network, an initial aggregate solution is not first determined with the transit network. This is because the impacts of transit user loads on transit services are not considered, which means that conditions do not change with different loadings. The process starts with the micro-assignment, where each trip is considered in turn. Lutility values for the transit network are determined for the trip (using the specific sensitivities assigned the corresponding traveler) and the optimal strategies process is used to identify the alternative reasonable paths along with the probabilities of use of each of these paths. These probabilities of use are then used to select a specific path for the trip using a Monte Carlo process. The loads on all the transit links in the selected path are increased by 1 and then the next trip in the list is considered.

The walk and cycle trips are then loaded one at a time using a path selection process link that used for the road-based trips. The results of the road network assignment are used and the walk and cycle trips are accumulated on links as the process continues.

At the time of writing, work on the implementation of these network assignment processes is still underway.

4.0 System Architecture

The principle components of the software system for running the model, apart from the model modules as described above, are:

- data store;
- process controller;
- user interface; and
- calibrator.

4.1. Data Store

The data store is the database where all the information input and output form the modules is stored. All information flowing between modules is ‘passed through’ (written to and read from) the data store.

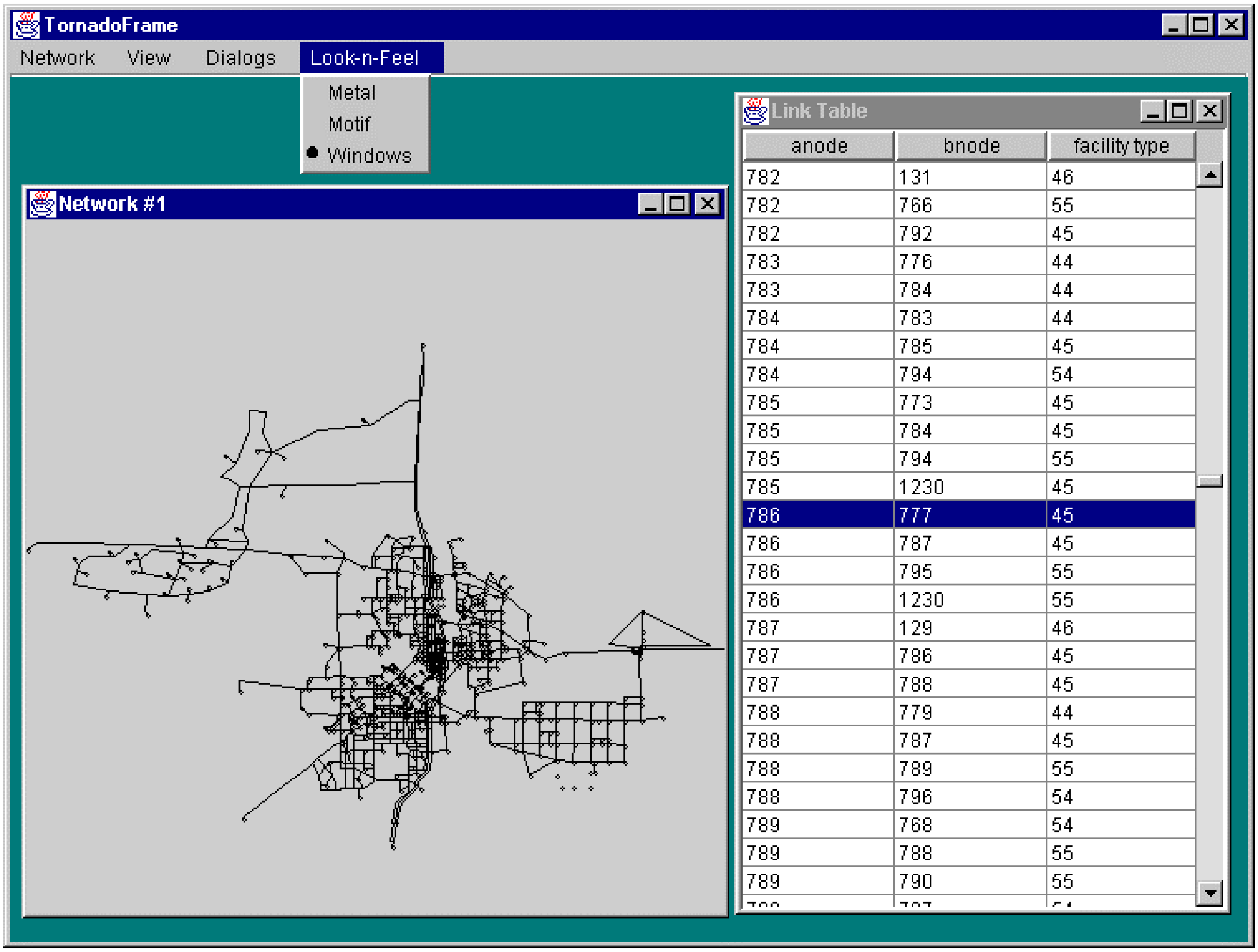
4.2. Process Controller

The process controller dictates the sequence of operation of each of the model modules in turn in order to facilitate a ‘run’ of the model. In a given year, the economic and demographic module is run first, followed by the production allocations and interactions module, and so on following a clockwise circuit ‘around’ the data store as shown in Figure 4. The process controller also informs the modules if it is the first year of the run – so that the modules can perform any additional set-up routines as required – and it continues year-by-year until the specified last year is reached. It would also be the process controller that would control whether or not the transport supply module is to be skipped in a given year as part of any strategy to reduce the total model run time.

4.3. User Interface

The user interface includes a graphic interface for facilitating both input and output, a run archiver and a process visualizer for monitoring the progress of a given model run in real time.

With the graphic interface, inputs are written to the data store and specified outputs from the data store are presented, in graphical or map format as appropriate. An example of the form of the graphical presentation of a network is included as Figure 6.



anode	bnode	facility type
782	131	46
782	766	55
782	792	45
783	776	44
783	784	44
784	783	44
784	785	45
784	794	54
785	773	45
785	784	45
785	794	55
785	1230	45
786	777	45
786	787	45
786	795	55
786	1230	55
787	129	46
787	786	45
787	788	45
788	779	44
788	787	45
788	789	55
788	796	54
789	768	54
789	788	55
789	790	55

Figure 6: Example of the form of presentation for the graphical interface system

The run archiver merely stores a record of the full set of inputs to the model together with specified outputs – and is intended to provide the user with a record of a given run and the results that it provided.

The real-time process visualizer provides the user with an indication of the current state of the process controller, including the module being run and the degree to which convergence has been achieved or the relevant list has been processed as appropriate.

4.4. Calibrator

The calibrator facilitates the estimation of specified model parameters given various observations of system behavior, considering the fit of the model across modules.

Overall, a Bayesian sequential procedure is used in estimation. Sub-models and individual relationships within the various modules are calibrated first, separate from the overall modeling system, and then the entire model is calibrated. The calibrator facilitates the calibration of the entire model by running the model and comparing its outputs with a selection of weighted observed data to provide a goodness-of-fit measure. Certain parameters are adjusted, and the model re-run to determine the effect of changes in various parameters on the goodness-of-fit. This provides information to a search procedure within the calibrator that searches for the set of highest level parameters that gives the best possible model fit.

Experience has shown that problems in a model design often show up as an inability to achieve goodness-of-fit in some values without compromising goodness-of-fit in other values (Abraham, 2000). The sensitivity of the estimated parameter values to changes in the weights in the goodness-of-fit measure can help to uncover these issues. The calibrator allows this information to be explored interactively after the estimation has converged.

5.0 Model Development and Implementation

The development of the model is very much a team effort. One of the benefits of the modular approach has been the ability to assign different modules to different team members, thereby spreading the effort, allowing for faster progress and the opportunity to draw on the differing strengths and knowledge-bases of the different team members.

A modeler and a computer programmer are assigned to each module. These assignments are listed in Table 1.

Java computer code has been used, following an objected-oriented approach in design and allowing for different modules to be developed in different locations and later brought together. The strategy that has been followed is to develop initial versions of the modules and conduct experiments with them, both separately and in combination, in order to test the design concepts and learn from what is tried in order to guide further development. Re-organization of the software code and the use of other languages in the interest of reducing run times is to be done as the development work continues and more is learned about the suitability of the design.

6.0 Conclusions

The work described here is certainly ‘pushing the envelope’ in modeling, with very ambitious goals and very exciting possibilities.

The use of micro-simulation in some of the modules is an essential element of the work. There are many ideas about the behavioral processes involved in land use and transport systems and the approaches to modeling these process that simply do not fit into an aggregate modeling framework. The use of micro-simulation in this case is allowing some of these ideas to be explored.

Module or Component	Modeler	Programmer
Regional Economics and Demographics	Carl Batten	Abe Dunn
Household Allocations	JD Hunt	John Abraham
Land Development	John Abraham & JD Hunt	John Abraham
Production Allocations and Interactions	JD Hunt	John Abraham
Household Travel	Joel Freedman & JD Hunt	Joel Freedman & John Abraham
Commercial Movements	Rick Donnelly	Rick Donnelly
Transport Supply	Jim Hicks & JD Hunt	Jim Hicks
Support and General Utilities	Tim Heier & Rick Donnelly	Tim Heier

Table 1: Assignments to modules

The use of specific utility function sensitivities that vary across agents – which produces a distribution in behavioral response consistent with behavioral theory – is a particularly successful component of the work. The micro-assignment processes that have been developed for the road network and the transit network considerations are particularly successful as well because they provide a practical method for avoiding the concentrated loading problems that arise with the use of centroid connectors and they give rise to a dispersion in behavior.

The calibrator is an essential component because it facilitates an appropriate estimation process within an acceptable level of effort.

The use of an aggregate treatment in some modules and a disaggregate treatment in others introduced some challenges concerning the interface between the two. This is perhaps most apparent with regard to the treatment of labor, where the aggregate treatment in the production allocations and interactions module has to be reconciled with the disaggregate treatment in the household allocations module.

The use of the object-oriented approach in design has facilitated a ‘fresh’ examination of some of the issues involved, and has in itself been useful.

The development of initial test version has been very useful, allowing for new ideas and approaches to be tested, guiding re-design and providing a ‘level of comfort’ as the work proceeds towards some of the longer-term goals.

Finally, at the time of writing the work described here is still in progress, and the design is still evolving. Perhaps the greatest factors leading to the reconsideration of elements of the design at this point are limitations in the available data for various parts of the model calibration.

Acknowledgements

As indicated above, the work described here is very much a team effort, and has benefited from the substantial efforts and ideas contributed by all those listed in Table 1. The work has also benefited from the comments and suggestions provided by Frank Koppelman, Michael Wegener, David Simmonds, Julie Dunbar, Keith Lawton and Gordon Shunk. The prime contractor, and team leader, for the work is Parsons Brinckerhoff.

The descriptions included here have in some cases drawn heavily from working documents produced by different members of the team.

The work itself is sponsored by the Oregon Department of Transportation, which is to be lauded for its appreciation of the far-reaching and long-term benefits expected from the work along with the more short-term ones already being realized.

Appendix A: Economic Sectors Included in Model

- Agriculture in Office Space: production in the Agricultural Industrial Sector that is located in Office Development Space and consumes Managerial, Professional and Clerical Labor
- Agriculture in Agricultural Space: production in the Agricultural Industrial Sector that is located in Agricultural Development Space and consumes Agricultural Labor
- Forest in Office Space: production in the Agricultural Industrial Sector that is located in Office Development Space and consumes Managerial, Professional and Clerical Labor
- Forest in Forest Space: production in the Agricultural Industrial Sector that is located in Forest Development Space and consumes Agricultural, Unskilled and Other Labor
- Light Industry in Office Space: production in the Light Industry Industrial Sector that is located in Office Development Space and consumes Managerial, Professional and Clerical Labor
- Light Industry in Light Industrial Space: production in the Light Industry Industrial Sector that is located in Light Industrial Development Space and consumes Assembly and Fabrication, Semi-Skilled Manual and Other Labor
- Heavy Industry in Office Space: production in the Heavy Industry Industrial Sector that is located in Office Development Space and consumes Managerial, Professional and Clerical Labor
- Heavy Industry in Heavy Industrial Space: production in the Heavy Industry Industrial Sector that is located in Heavy Industrial Development Space and consumes Assembly and Fabrication, Semi-Skilled Manual and Other Labor
- Wholesale in Office Space: production in the Wholesale Industrial Sector that is located in Office Development Space and consumes Managerial, Professional and Clerical Labor
- Wholesale Industry in Warehouse Space: production in the Warehouse Industrial Sector that is located in Warehouse Development Space and consumes Semi-Skilled Manual, Unskilled Manual and Other Labor
- Retail in Office Space: production in the Retail Industrial Sector that is located in Office Development Space and consumes Managerial, Professional and Clerical Labor
- Retail in Retail Space: production in the Retail Industrial Sector that is located in Retail Development Space and consumes Retail and Other Labor
- Hotel and Accommodation: all production in the Hotel and Accommodation Sector that is located in Hotel Development Space and consumes all categories of Labor
- Construction: all production in the Construction Sector that is located at construction sites and consumes all categories of Labor
- Health Care in Office Space: production in the Health Care Industrial Sector that is located in Office Development Space and consumes Managerial, Professional, Clerical and Health Care Labor
- Health Care in Hospital Space: production in the Health Care Industrial Sector that is located in Hospital Development Space and consumes all categories of Labor

- Health Care in Institutional Space: production in the Health Care Industrial Sector that is located in Institutional Development Space and consumes all categories of Labor
- Transportation Handling in Office Space: production in the Transportation Handling Industrial Sector that is located in Office Development Space and consumes Managerial, Professional and Clerical Labor
- Transportation Handling in Depot Space: production in the Transportation Handling Industrial Sector that is located in Depot Development Space and consumes Semi-Skilled Manual, Unskilled Manual and Other Labor
- Other Services in Office Space: production in the Agricultural Industrial Sector that is located in Office Development Space and consumes Managerial, Professional and Clerical Labor
- Other Services in Light Industrial Space: production in the Other Services Industrial Sector that is located in Light Industrial Development Space and consumes Assembly and Fabrication, Semi-Skilled Manual, Unskilled Manual Labor and Other Labor
- Other Services in Retail Space: production in the Other Services Industrial Sector that is located in Retail Development Space and consumes Retail Labor
- Grade-School Education in Office Space: production in the Grade-School Education Industrial Sector that is located in Office Development Space and consumes Managerial, Professional and Clerical Labor
- Grade-School Education in Grade-school Space: production in the Grade-School Education Industrial Sector that is located in Grade-School Development Space and consumes Grade-School Teaching Labor
- Post-Secondary Education: production in the Post-Secondary Education Sector that is located in Institutional Development Space and consumes all categories of Labor
- Government in Office Space: production in the Government Industrial Sector that is located in Office Development Space and consumes Managerial, Professional and Clerical Labor
- Government in Government Support Space: production in the Government Industrial Sector that is located in Government Support Development Space and consumes all categories of Labor
- Government in Institutional Space: production in the Government Industrial Sector that is located in Institutional Development Space and consumes all categories of Labor

Appendix B: Commodity Categories Included in Model

- Farm products
- Forest products
- Fresh fish or marine products
- Metallic ores
- Coal
- Crude petroleum, natural gas or gasoline
- Nonmetallic minerals
- Ordnance or accessories
- Food or kindred products
- Tobacco products, excluding insecticides
- Textile mill products
- Apparel or other finished textile products or knit apparel
- Lumber or wood products, excluding furniture
- Furniture or fixtures
- Pulp, paper or allied products
- Printed matter
- Chemical or allied products
- Petroleum or coal products
- Rubber or miscellaneous plastic products
- Leather or leather products
- Clay, concrete, glass or stone products
- Primary metal products, including galvanized
- Fabricated metal products
- Machinery, excluding electrical
- Electrical machinery, equipment or supplies
- Transportation equipment
- Instruments, photographic goods, optical goods, watches or clocks
- Miscellaneous products or manufacturing
- Waste or scrap materials not identified by producing industry
- Other (miscellaneous) freight shipments
- Containers, carriers or devices, shipping, returned empty
- Waste hazardous materials or waste hazardous substances
- Construction services
- Pipeline transportation services
- Transportation and storage services
- Radio and television broadcasting services
- Postal services
- Utilities services
- Wholesale margins
- Retail margins
- Other finance, insurance and real estate services

- Business services
- Education services
- Health services
- Amusement and recreation services
- Accommodation services
- Food services
- Other personal and miscellaneous services
- Managerial labor
- Professional labor
- Grade-school teaching labor
- Clerical labor
- Assembly and fabrication labor
- Agricultural labor
- Semi-skilled manual labor
- Unskilled manual labor
- Retail labor
- Health care labor
- Post-secondary teaching labor
- Other labor

Appendix C: Developed Space Types Included in Model

- Single-family residential
- Multi-family residential
- Light industrial
- Heavy industrial
- Office
- Retail
- Warehouse
- Hotel
- Depot
- Hospital
- Institutional
- Grade school
- Post-secondary school
- Government support
- Agricultural
- Forest
- Protected resource
- Interim parking
- Vacant

References

Abraham, J. (2000) Parameter Estimation in Urban Models: Theory and Application to the MEPLAN Model of Sacramento. Unpublished PhD Dissertation, Department of Civil Engineering, University of Calgary.

Beckmann, R., Baggerly, K. and McKay, M. (1996) "Creating synthetic baseline populations." *Transportation Research* 30A(6):415-435.

Conway Jr., R.S. (1990) "The Washington Projection and Simulation Model: A regional interindustry econometric model." *International Regional Science Review* (13:1)

de la Barra, T. (1998) The mathematical and algorithmic structure of TRANUS. Unpublished paper available at http://www.modelistica.com

Hunt, J.D. and Simmonds, D.C. (1993) "Theory and application of an integrated land-use and transport modelling framework." *Environment and Planning* 20B:221-244.

INRO (1996) *EMME/2 User's Manual; Release 8.0; April 1996*. INRO Consultants, Montreal QC, Canada.

Ortúzar JdeD and Willumsen LG (1990) *Modelling Transport*. Wiley, Chichester, UK.

Rey, S.J. (1998) "The performance of alternative integration strategies for combining regional econometric and input-output models." *International Regional Science Review* (21)