Integrating a traffic router and micro-simulator into a land use and transportation model

By:
Austin Troy
Dale Azaria (corresponding author)
Brian Voigt
Spatial Analysis Lab
Rubenstein School of Environment and Natural Resources
University of Vermont
81 Carrigan Drive
Burlington, VT 05405
802-656-8336
Austin.Troy@uvm.edu, Dale.Azaria@uvm.edu, Brian.Voigt@uvm.edu

Adel Sadek
University at Buffalo, The State University of New York
233 Ketter Hall
Buffalo, NY 14260
Phone: (716) 645-4367
asadek@buffalo.edu
Abstract

Because accessibility is a critical factor in determining land use, land use models have long been integrated with travel demand models. As travel modeling moves from the traditional four-step approach toward advanced modeling techniques, including microsimulation, integrated land use-transportation models are expected to evolve as well. This paper describes a first-of-its-kind attempt at integrating a dynamic, second-by-second traffic router/micro-simulator with a highly disaggregated and dynamic land use model. The traffic microsimulator captures the emergence of congestion and its impact on traffic flow, rather than looking solely at the ratio of traffic volume to capacity as the traditional four-step models do. In comparing the land use outputs of a model system that includes the micro-simulator with the results from the system without micro-simulation for our study area of Chittenden County, Vermont, statistical tests found only slight differences in the land use predictions between the two model integrations for a 40-year simulation. Although these differences were slight, their spatial patterns shed light on how transportation models influence the outcome of land use models. In particular, differences in land use predictions appear to relate to the traffic micro-simulator’s predictions of emergent traffic bottlenecks along routes that serve peripheral areas where there is poor redundancy in route choice. These results suggest that land use models are at least somewhat sensitive to the type of transportation model that is used to generate accessibility measures. Our study site is a small metropolitan area with only modest population pressures and limited traffic congestion. Differences in predictions between model integrations grow as population forecasts are artificially increased, suggesting that integration of traffic micro-simulation may be of greater use in more congested areas.

Introduction

The linkages between land use and transportation and the need to incorporate those linkages in planning are well established (1-4). Under the Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 and the Transportation Equity Act for the Twenty First Century (TEA-21) of 1997 (to a lesser extent), in order to receive certain types of federal transportation funds, state or regional transportation agencies are required to model the effect of transportation infrastructure development on land use patterns and to consider whether transportation plans and programs are consistent with land use plans. Metropolitan Planning Organizations (MPOs) are increasingly integrating dynamic land-use modeling into those efforts to evaluate transportation infrastructure performance, investment alternatives, and air quality impacts. Dynamic-coupled models differ from stand-alone models in that they simulate the dynamic interactions between transportation and human activities. Because accessibility is an important factor in determining land use, dynamic land use models have long been integrated with four-step travel demand models (5).
However, as dynamic components are added, model integrations become increasingly complex and difficult to implement. Little guidance exists about what levels of complexity or disaggregation is needed or appropriate for modeling land use and transportation and how that changes for different planning applications. The correct balance likely depends on the particular application of the model. Many new approaches to comprehensive model-integration are being unveiled in the research community. However, as noted by Hunt et al. (6), few of these models have been conclusively shown to increase the accuracy of the model output.

This paper presents one of the first known attempts to integrate a traffic router/micro-simulator operations model with a highly disaggregated and dynamic land use model. Ongoing work by others demonstrates the possibilities of incorporating activity-based modeling into land use (Waddell 2010) and of incorporating activity-based modeling with traffic micro-simulation (Lin 2010).

Three components are used in this modeling effort: UrbanSim for land use (7-9), TransCAD (Caliper, Inc) for travel demand modeling and traffic routing and assignment, and TRANSIMS for traffic routing through micro-simulation (10-11). We compare the more commonly-used integration of the land use model with the static traffic assignment (TransCAD) to the novel integration of the land use model with the dynamic router/micro-simulator (TRANSIMS). The latter integration also requires use of TransCAD for trip generation, so we refer to the simpler integration as the “two-way model” and the more complex one as the “three-way model.”

UrbanSim is a land-use allocation model that simulates urban growth for a region based on externally derived estimates of population and employment growth (control totals). Expected growth is spatially allocated across the landscape to simulate the pattern of future development and land use. Agents in UrbanSim include households, employers, and real estate developers. The landscape is divided into grid cells of a user-defined size (geographic units like parcels can also be used). Each simulated development event is assigned to one of those cells based on factors like accessibility, site constraints, zoning, and land value. Model features include the ability to simulate the mobility and location choices of households and businesses; developer choices for quantity, location and type of development; fluxes and short-term imbalances in supply and demand at explicit locations; and housing price adjustments as a function of those imbalances.

While almost all other urban growth models rely on aggregate cross-sectional equilibrium predictive approaches, UrbanSim is an agent-based behavioral simulation model that operates under dynamic disequilibrium, which allows for more realistic modeling of economic behavior; supply-demand imbalances are addressed incrementally in each time period but are never fully satisfied (Iacono 2008). Because of its dynamic nature, UrbanSim endogenizes factors that older models took as exogenous, such as location of employment and the price of land and buildings.
Because accessibility can play an important role in land use decisions, UrbanSim is generally integrated with some type of transportation model. The assumption is that accessibility changes over time, so transportation should be dynamically linked with land use to improve model results. The degree to which accessibility affects land use in a given implementation of the model system depends on the way that the various statistical models in UrbanSim are parameterized and the extent to which the data reveals a relationship. In our version of UrbanSim, the residential development choice location model and the commercial development choice location model both include coefficients for accessibility.

TransCAD is a traditional four-step travel demand model, including trip generation, trip distribution, mode split and traffic assignment. The trip generation step quantifies the number of incoming and outgoing trips for each zone based on land use and employment patterns, and classifies these trips according to their purpose (e.g., home to work, home to shopping). Trip distribution assigns the incoming and outgoing travel from the trip generation step to specific zones. The mode split step estimates the number of trips by mode of transport. Finally, the traffic assignment identifies the route for each trip. Traffic assignment is based on an equilibrium model which employs an iterative procedure to reach convergence.

TRANSIMS is a detailed, data-intensive operations model that is designed to simulate traffic behavior with great spatial and temporal disaggregation. It consists of four modules: (1) synthetic population generator; (2) activity generator; (3) router; and (4) micro-simulator. In standalone implementations, TRANSIMS starts by creating a synthetic population based on census and land use data, among other data sets. The Activity Generator then creates an activity list for each synthetic traveler. The router then computes combined route and mode trip plans to accomplish the desired activities. Finally, the micro-simulator simulates the resulting traffic dynamics based on a cellular automata model, yielding detailed, second-by-second trajectories of every traveler in the system over a 24-hour period. The micro-simulator allows for a highly detailed characterization of traffic flows and is able to take into account factors like cueing, car-following, and lane changing behavior. As an operations model, it is designed to help optimize microscopic factors such as signal timing and actuation.

While TRANSIMS allows for an activity-based approach to transportation demand modeling (using the population synthesizer and activity generator), the model’s router and micro-simulator modules can be applied using standard Origin-Destination (O-D) matrices. Implementing only TRANSIMS’s router and micro-simulator is typically referred to as a “Track 1” TRANSIMS implementation. “Track 1” TRANSIMS implementation has been the focus of the current work so far. While some have suggested that using only the traffic supply modules of a microsimulator and not the traffic demand modules fails to exploit the overall purpose of microsimulation (12), in our view the Track 1 implementation provides a cost-effective approach for regional planning.
organizations, which can take advantage of the increased resolution of the
TRANSIMS micro-simulator, while continuing to depend upon familiar O-D
matrices. It also sets the stage for a more complete implementation in the future.

Since we have not yet incorporated TRANSIMS’ activity-based approach to
transportation demand in this model system, the primary difference between the
2-way model and the 3-way model is the way each one characterizes traffic and
resulting accessibilities (which are an input into UrbanSim). TransCAD uses a
volume-delay function, where the congested travel time on the link is equal to
the ratio of the number of vehicles on the link divided by the total capacity of the
link. It assumes that inflow equals outflow for all individual links in the network.
TRANSIMS, on the other hand, calculates congested travel times based on a
simulated interaction of vehicles on the roadway that takes into account factors
like weaving, merging, queuing, traffic signals, and intersection spill-back.
TRANSIMS is designed to replicate the real-world phenomenon that lead to
increased travel time and congestion that cannot be explained by just a simple
volume-to-capacity ratio. This means that failure can occur at some intersections
where inflow no longer equals outflow. As a result, TRANSIMS is likely to
predict more localized bottlenecks.

Objectives

The first purpose of this paper is to introduce and describe the integration of the
TRANSIMS router/micro-simulator with the UrbanSim land use model. The
second purpose is to determine whether the two model integrations lead to
different land use predictions. To the extent the land use predictions differ, we
analyze the pattern of outputs to better understand how the two approaches to
calculating accessibilities in each transportation model contributes to these
differences. By characterizing and analyzing these differences we hope to shed
light on the role that transportation and accessibility modeling play in long-term
land use predictions and the tradeoffs to added complexity in such modeling
efforts.

Methods

Modeling Site

Our models are run for Chittenden County, VT (Figure 1), the most populous
county in the state and the home to its largest city, Burlington. Chittenden
County is among the smallest metropolitan areas where UrbanSim has been
implemented, with an estimated 2009 population of 152,000. It is an excellent
location for modeling for two reasons: first, its small size makes highly
disaggregate and data-intensive modeling tractable; second, its isolation from
other cities (the nearest metropolitan area is Montreal, more than 90 miles away),
means it approximates “closed city” modeling conditions (although we do use 17
external TAZs to account for inter-county traffic, this is a small component of the county’s overall transportation). Despite its small size, Chittenden County has its own Metropolitan Planning Organization, which conducts extensive modeling.

![Map of Chittenden County](image)

**FIGURE 1** Map of Chittenden County.

**Description of the Models**

This analysis was conducted by integrating previously developed implementations of three models. We used an implementation of UrbanSim developed for Chittenden County, Vermont, by Austin Troy and Brian Voigt (6, 14, 15). We used the Chittenden County Metropolitan Planning Organization’s (CCMPO) implementation of TransCAD, which was developed for the MPO by Resource Systems Group, Inc. The model includes 335 internal traffic analysis zones (TAZs) to simulate traffic flow, and includes an additional 17 external zones to represent traffic entering (or passing through) the County from outside its borders (14). The travel model is based on household travel diaries collected for the CCMPO. Customized scripts were developed that automated the
integrated models. We used the implementation of TRANSIMS developed by Resource Systems Group and Adel Sadek (12, 13).

The 2-way configuration consists of UrbanSim, which generates the socio-economic land use data like total number of households and employment in each traffic analysis zone, and TransCAD, which derives accessibilities using travel times from the static vehicle assignment. These travel times are then sent as input to UrbanSim. After every five years of model time TransCAD is rerun using updated land use data from UrbanSim, and in turn updating UrbanSim’s accessibilities (6, 14, 15).

The 3-way configuration adds a third component: the TRANSIMS router/microsimulator. In this configuration (Figure 2), TransCAD performs trip generation, trip distribution, and mode choice, and exports a PM peak vehicle trip matrix to TRANSIMS. TransCAD’s static vehicle assignment is replaced by TRANSIMS’ regional vehicle micro-simulation. The amount and distribution of regional auto travel demand is identical in the two models, but in the 3-way model accessibilities are derived using the simulation-based auto travel times and sent as input to UrbanSim.

![FIGURE 2 Three-way model configuration.]

**Integration of the Traffic Micro-simulation Model**

Because the 3-way model still uses TransCAD for trip generation and because TransCAD operates at an aggregate level, a significant task in integrating the 3-way model was to convert the PM peak hour vehicle trip matrices produced by TransCAD to daily vehicle trips that could be used by the microsimulator. Using diurnal distribution data collected during the development of the daily CCMPO TRANSIMS model, and daily peak PM hour traffic volume (defined as 5:00 pm to 6:00 pm in the TransCAD model), we derived a PM peak hour to daily adjustment factor for each of five trip types. The diurnal distribution data is
The second significant difference in the 3-way model is the calculation of auto travel times, which are the most significant component of accessibility in the model system. In the 2-way model, TransCAD generates a file that contains auto, walk/bike, and transit utilities as well as the logsum (composite measure of accessibility across modes) for each zone-to-zone pair. This file is fed back to UrbanSim for the next iteration. By incorporating TRANSIMS into the model chain in the 3-way model, we replace the auto utilities in this file with auto utilities based on zone-to-zone travel times calculated by the TRANSIMS microsimulator instead of the TransCAD model assignment module.

TRANSIMS-based auto utilities are calculated using the following regression equation:

\[
\text{Utility (Auto)} = -1.09438 - 0.020795 \times \text{TRANSIMS Time}
\]

Logsum value for each zone-to-zone pair are calculated based on the new auto utilities.

\[
\text{Logsum} = \ln(\exp[\text{Utility(Walk-Bike)}] + \exp[\text{Utility(Transit)}] + \exp[\text{Utility(Auto)}])
\]
A python script reads the existing logsum file generated by the TransCAD model as well as a TRANSIMS zone-to-zone travel time skim file. The program updates the logsum file by calculating a new auto utility and then recalculating the logsum for each zone pair using the equations presented above. The revised logsum file can then be used as input to UrbanSim to complete the feedback process.

A new module was added to the CCMPO TRANSIMS model that writes out a zone-to-zone travel time skim matrix. The skim file output contains the zone-to-zone congested travel time for the 5:00pm to 6:00pm hour calculated by the microsimulator.

Model Runs and Analysis

We ran forty year simulations of both the 2-way and 3-way model integrations using the same data sets, starting in 1990 and ending in 2030. In both cases, UrbanSim iterated every year while the transportation model ran every five years. A fixed seed was used in choice-set delineation for UrbanSim to minimize stochasticity and maximize comparability between the model integrations. Both model integrations use the same UrbanSim model coefficients.

Two versions of each model were run, one using population and employment forecasts obtained from the MPO as control totals, (the “baseline scenario”), and another using controls totals artificially increased by 50% (the “increased control total scenario”). This was done to help determine whether differences in the model outputs relate to population development pressures.

Finally, we analyzed the outputs. While a large number of indicators are produced by these model integrations, we focus this analysis on three: residential units (at the town and TAZ level), commercial square footage (at the town and TAZ level) and accessibilities, characterized as logsum values (at the TAZ level only). Because our model base year is 1990, we were able to conduct a preliminary validation of both model integrations against observed data from later years (2006 for household development and 2009 for commercial development). We found no statistically significant differences in prediction accuracies for the two model integrations. For that reason, we do not present the results here. Nevertheless, we ran statistical analyses to look for differences in the 2030 outputs of the two models and analyzed geographic patterns in those differences.

Results

Statistical differences in models

Variance ratio tests for the whole population of TAZs revealed no significant difference in variance across the whole population of TAZs between models for both sets of indicators for 2030. Using paired t-tests, slight significant differences
were found in predicted commercial square footage for 2030 at the TAZ level when grouped by town. For the baseline population scenario, significant differences in commercial square footage were found at the 95% confidence level for the town of Williston (t=2.654, p=0.011), which has the third largest number of TAZs in the county. Westford had significant differences at the 90% confidence level (t=-2.366, p=0.099). With the increased control total scenario, differences were fewer: there were no significant differences in commercial square footage at the 95% confidence level, although Burlington (t=-1.825, p=0.072) and Shelburne (t=-1.867, p=0.092) were different at the 90% level. A significant difference in residential units was found for Milton at the 95% confidence level in the baseline scenario (t=-2.487, p=.03). In the increased control total scenario, significant differences at the 95% level were found in residential units for Jericho (t=-3.61, p=.037) and at the 90% level for Milton (t=-2.12, p=.058). A spatial statistical analysis was also conducted using Moran’s I (Moran (1950) to see if measures of spatial autocorrelation differed between the outputs of the two models, but no difference was found.

Preliminary Comparison of Travel Times

Figure 4 shows the difference in predicted logsum accessibilities between the 2-way model and the 3-way model for the year 2030 under a scenario with baseline population forecast control totals. Because accessibility is one of the factors in the land use development choices, the fact that there are clear differences in the spatial pattern of accessibility served as an indication that differences in land use outputs were a distinct possibility, and that further analysis was warranted.
Long-term trends

We looked at graphs of key indicators to see when large discrepancies emerge between the models, if at all. Figure 5 shows the percent difference in predicted housing units between the two models for a sample of six towns from 1990 to 2030. It indicates a continuously growing difference for the outlying towns of Milton and Underhill. Milton has higher predictions for the 2-way model, while Underhill has the opposite. Other towns, like Bolton, show divergence between the models in early years and then return to smaller differences later. Several towns start to show patterns of divergence between models and then return to small differences in later years, such as South Burlington, Richmond and Colchester. Others are in close agreement throughout all forty years of model time, such as Charlotte and Burlington. Commercial square footage prediction graphs (not shown here) show a somewhat similar pattern with Milton also having increasingly positive 2-way prediction differences over time, several
outlying towns with the opposite pattern and a number of towns in the middle, with relatively little difference.

FIGURE 5 Percent difference in predicted residential units between models (2-way minus 3-way divided by total units) for a sample of 6 towns.

Side by side maps in Figure 6 and 7 show percentage differences in predicted residential units (a) and commercial square footage (b) for 2030 at the town level and the TAZ level, respectively, under the increased control total scenario. Baseline control total maps are not shown in the interests of space and because the patterns are similar but much weaker.

FIGURE 6 Town-level comparison under increased control totals: (a) Percent difference in residential development forecasts from the 2-way and 3-way models for 2030 using baseline control totals. (b) Percent difference in commercial development forecasts from the 2-way and 3-way models for 2030 using baseline control totals. Blue
Discussion

This project was the first of its kind to integrate a traffic router/micro-simulator with a highly disaggregated and dynamic land use model. This project shows that such an integration is feasible, although it is also difficult, time consuming, and expensive. With hundreds of gigabytes of outputs, far more analysis of the results of these models remains to be done before any final assessment of the value of this project can be made. However, this analysis represents a preliminary attempt to address it, at least in the context of a region with little competitive pressure on land use and low rates of congestion.

The fact that accessibilities are far more spatially heterogeneous in the 3-way model (Figure 4), would lead us to believe that, theoretically, there could be systematic differences in the land use outputs. Our UrbanSim implementation
consists of ten statistical models that drive activities like household and employment moves, land price, and development events. While many include spatial parameters such as location within the “urban core,” or the amount of commercial or residential development within walking distance, only the residential and commercial development models include parameters on accessibility from the travel model. Because TRANSIMS predicted more localized areas of reduced accessibility within the interior of the county, we expected to find that some more centrally located areas might develop slightly less in the 3-way model than in the 2-way model.

While the results of our two models are different, it is not clear that these differences are important enough to matter for the purposes of land use change prediction. Our validation results (not presented here) show minimal differences between the two in predicting intermediate-year data. Statistical pairwise comparisons of TAZ-level results grouped by town suggest that differences in predicted indicators for 2030 are present for only a few towns. Tests of the whole population of TAZs found no significant difference in variance for both land use indicators.

Nonetheless, our maps of 2030 prediction differences in commercial development under increased control totals (which was use because it emphasizes the differences between models more) show some interesting patterns that suggest potential systematic spatial differences in predictions. As Figure 6 shows, all the peripheral towns along the northern and eastern boundaries of the county have more commercial development under the 2-way model than under the 3-way. The same pattern is evident at the TAZ level, although heterogeneity is slightly greater along the periphery. This result is intuitive given what we know of the models. As population grows, TRANSIMS predicts more congestion and delay and hence lesser accessibility in the outer TAZs than TransCAD. This pattern is particularly evident for TAZs that do not adjoin the Interstate (where the Interstate runs through, there are fewer red TAZs). Redundancy of routes is very poor the further out one travels in the county, so just a few high-delay links can make a big impact on accessibility in areas that require a long drive on non-Interstate routes. Our preliminary analysis of TRANSIMS’s link level outputs (not presented here) shows a number of predicted traffic bottlenecks along such key arterials that connect outer suburbs to the urban core that TransCAD does not capture. Not all of these “red TAZs” are on the outer periphery. Some are more central, but require significant driving on bottleneck-prone arterials.

Interestingly, as is reflected in Figure 8, most of the TAZs containing an Interstate exit appear to have higher employment predictions in the 2-way model, which is consistent with this explanation.
FIGURE 8 Blow-up of Figure 7(b) showing Interstate Exits.

No clear spatial pattern is evident for differences in residential predictions. Figure 6 suggests that only one of the towns included in the graph experience steadily increasing differences over time between models. Otherwise, differences oscillated within a small range over time. This difference between residential and commercial indicators is likely due to the model coefficients that relate to output from the transportation models. The residential developer model includes a parameter for home accessibility to employment while the commercial developer model includes a parameter for work accessibility to employment. Further, the commercial development coefficient is almost twice the magnitude of the residential coefficient.

Conclusion

TRANSIMS is designed as an operations model for assessing and optimizing microscopic factors in the traffic network. Some believe that models like this are inappropriate for coupling with long-term land use change models. Our land use results from the 2030 simulation look generally reasonable, but our preliminary
analysis of link level data from TRANSIMS indicates that after forty years of simulation, a number of unrealistic bottlenecks and congestion points develop. This is probably because, as an operations model, TRANSIMS runs with an assumption that factors like signal timing and lane rules are to be changed over time. When they remain static over long periods like forty years, this may lead to unrealistic characterizations of accessibility. Nonetheless, these bottlenecks only had a very minor impact on development predictions. This may be because of our model coefficients, which were estimated in an area where traffic congestion is relatively minimal. Had we estimated these coefficients in a larger urban area with extensive congestions, it is possible that the impacts of these accessibility differences on development would have been greater. Hence, the impact of transportation model type on land use results is extremely sensitive to model coefficient specification. It is also possible that had we run the TRANSIMS Track 2 implementation which includes the activity model with disaggregated activity locations, differences would have been more pronounced.

Given our current results, there appears to be little justification for expending the large amount of time and money required to implement TRANSIMS for the purposes of long-term land use modeling in a context like Chittenden County. However, this approach might be more valuable in large metropolitan areas where population pressures and traffic delays are much greater. In such cases, we would expect to find delay-related (as opposed to distance-related) accessibility having a greater impact on land use. It is possible that in such cases a land use model integrated with TRANSIMS would yield a more accurate characterization of accessibility, leading to better land use predictions. However, such a model should probably only be run for short-term predictions in highly congested areas, as long-term simulations could result in unrealistic localized stoppages of traffic flow which, in real life, could be addressed through minor interventions, like re-timing signals. Further research is warranted to determine the usefulness of including a micro-simulator in land use modeling for more populous and congested regions and to determine the appropriate time frame of modeling in this context.

The integration of TRANSIMS with a land use model may also be valuable in assessing how hypothetical changes to the transportation network might influence the spatial pattern of development, potentially even in smaller metropolitan areas. We are currently in the process of running the 2-way and 3-way models on an alternative scenario involving the construction of a large number of new roads to determine if the 3-way model’s land use predictions are more spatially sensitive to the new infrastructure. This and other future research will help us better understand the usefulness and cost effectiveness of complex integrated modeling tools for the planning process.

Acknowledgements
This work was funded by grants from the USDOT administered through the University of Vermont Transportation Research Center and through FHWA. The authors also wish to thank the Chittenden County Metropolitan Planning Organization and Regional Planning Commission for their collaboration, as well as Stephen Lawe, Brian Grady, and John Lobb of Resource Systems Group, Inc., and James Sullivan of the University of Vermont Transportation Research Center.
References


