New England University Transportation Center
77 Massachusetts Avenue, E40-279
Cambridge, MA 02139
utc.mit.edu

Year 25 Final Report

Grant Number: DTRT13-G-UTC31

Project Title:

Scenario Discovery for Resilient Urban Systems (or, The Future *is* "Big Data")

Project Number:	Project End Date:	Submission Date:
MITR25-17	31 Dec, 2016	25 April, 2017

Principal Investigator:	Chris Zegras	**Co-Principal Investigator:**	Francisco Pereira
Title:	Assoc. Prof.	**Title:**	Professor
University:	MIT	**University:**	DTU
Email:	czegras@mit.edu	**Email:**	camara@dtu.dk
Phone:	617 452 2433	**Phone:**	(former Research Scientist, MIT)

INTRODUCTION

Our original primary aim was to formalize techniques for representing and propagating uncertainty in integrated land-use and transportation (LUT) models and demonstrate how this more rigorous characterization of uncertainty can lead to more robust and flexible systems, and improve decision-making. The project successfully completed the development, calibration and partial validation of an integrated land use-transport (LUT) model for the Boston metropolitan area and ran the model under alternative scenarios up to 2030. The LUT model was implemented in commercial software, using Citilabs' Cube Voyager (four-step transport model) and Cube Land (land use model).

Our original intention was to implement the models in a cloud-based computing environment, but the costs were too prohibitive. In the end, we used a machine with 20 dual core I7 chips (40 threads) which allowed parallelization of model processes. For the operational LUT model, we compiled a multi-source database on travel demand, transport supply, demographics, the housing market, employment, and land uses from 1970 to 2010 for the model area (Boston Metropolitan Area).

Researchers involved in this project included:

- PI: Chris Zegras, Associate Professor, DUSP
- Postdoctoral Associate: Victor Rocco, DUSP
- Research Associate: Mikel Murga, CEE
- PhD Student: Yafei Han, CEE/Transportation
- Masters Students: Michael Dowd (MST/MCP), Shenhao Wang (MST/MCP), Menghan Li (MST/MCP), Pablo Posada (MCP)

This work leveraged research resources from a Joint Research Project between the MI/MIT Cooperative Program at MIT and Masdar Institute (model development), the Center for Advanced Urbanism (sea level rise analysis), and the Future Urban Mobility Interdisciplinary Research Group under the Singapore MIT Alliance for Research and Technology (scenario frameworks).

TRANSPORT MODEL

On the transport side of the LUT model, we developed an enhanced four-step transport model for Boston Metropolitan Area. This included individual model estimation, four-step model implementation and validation for the years 1990 and 2010. Compared to conventional practices, we introduced more complexity and flexibility to the model, which provides a richer representation of reality and the ability to explore the impact of alternative models (model uncertainty) on forecast results.

For model estimation, work included:

- Logit model estimation of vehicle ownership choice for 1990 and 2010;
- Logit model estimation of mode choice for 1990 and 2010

- Gravity model calibration for trip distribution by worker' earning category for 2010.

For the supply side, we updated the transport network and its evolution by collecting all available historical data on the transit system and road network to reflect system changes from 1990 to 2010. The final four-step model is implemented in a combination of Python and Cube Voyager (Citilabs' commercial software). The first three steps of the model are implemented in Python and then integrated via Python scripts into Cube Voyager for network assignment. This approach allows us to introduce more complex model specifications, which cannot be easily implemented in the commercial software and also gave us the flexibility to run uncertainty simulations inside a four-step model.

Important/new features of the transport model include:

- A rich characterization of households (224 household types);
- Work-at-home likelihood (varying by earning groups) considered for home-based work trip generation rates;
- Trip distribution varying by income category;
- Road traffic assignment disaggregated by household income.

The transport model was calibrated and validated by comparing model outputs with observed data, including travel surveys, census data, traffic counts, and transit ridership data.

Uncertainty Analysis: Transport Model

We conducted a series of analysis of exogenous uncertainty, model uncertainty and behavioral uncertainty. Han (2015) includes a literature review of the sources of uncertainty, uncertainty analysis methods (including uncertainty propagation, and model transferability assessment), and empirical precedents and findings. We also depict the major demographic trends from 1980 to 2010 and travel behavior changes from 1990 to 2010 in the Greater Boston Area. Since demographics are a main source of input uncertainty, different aspects of demographics can have different degrees of uncertainty: some are highly predictable, while others are more surprising. Such empirical evidence is valuable to studies on input uncertainty. Han (2015) presents the empirics for the study area.

To examine a timely and, for Boston, relevant future source of exogenous uncertainty, we examined the transport network impacts of potential Sea Level Rise (SLR) due to climate change. Dowd (2015) demonstrates a method (Inundation Impact Assessment) for quantifying transport network impacts under six different inundation levels, one-foot to six-feet. The results indicate that inundation has widespread impacts on the ability of persons to complete trips and the performance of both the auto and transit networks. Dowd (2015) demonstrates how four-step transport models can be used to plan for SLR by modeling two different demographic scenarios for the year 2030 with two different public transport infrastructure alignments. A visualization of the SLR scenario and impacts can be found at: http://mdgis.github.io/#.

To examine other sources of uncertainty (behavioral, data, and model), Han (2015) analyzed the temporal transferability of vehicle ownership models and trip generation models for Boston metropolitan area from 1990 to 2010. The statistical tests show significantly changed preferences in household vehicle ownership choice and trip production. The prediction tests suggest that failing to consider preference changes cause significant bias in population demand forecasts. Han and Zegras (2016) further analyze model uncertainty and behavior uncertainty in vehicle ownership modeling.

Uncertainty Propagation

Within the full four-step model system we analyzed the propagation of uncertainty arising from two sources: model uncertainty and behavioral uncertainty (Table 1).

Table 1. Summary of Four-Step Model Uncertainty Propagation Analysis Approach and Findings

Four-Step Model Sub-models	**Model uncertainty**		**Behavior uncertainty**	**Findings**
	Model specification	*Model parameter*	*Model parameter*	
Vehicle ownership (VO)	Yes	Yes	Yes	Behavioral uncertainty largest impact on VO forecast difference - Largest propagation to MC o Non-work trips; bus trips (8%) - Little impact on VMT, VHT
Trip generation (TG)		Yes	Yes	Parameter uncertainty - ±10% difference in total trips - 6% difference in VMT - -5% to 10% difference in transit ridership Behavior uncertainty - 14% difference in total trips - 15% difference in VMT - 15% difference in transit ridership
Mode choice (MC)		Yes	Yes	Parameter uncertainty: little impact Behavior uncertainty - -22% difference in VMT - -14% difference in transit ridership

The overall finding is that behavior uncertainty has substantial impacts on model forecast results, while sampling uncertainty of parameters leads to a smaller range of variations relative to the point estimates.

LAND USE MODEL

The land use model is implemented in Cube Land, the commercial version of MUSSA, an operational land use model based on auction theory. The model distributes agents – households and firms – by real estate type to specific locations using a market equilibrium model, which has three sub-models: a Demand Model, a Supply Model, and a Rent Model. In the demand model, a consumer decides the bid for each type of property in each zone. The bid, or willingness to pay, comes from the utility function, which depends on the consumer characteristics, real estate attributes and location. Given fixed supply, the highest bidder wins. In the supply model, supply agents decide the amount of each type of real estate to offer based on profit maximization. Rent connects supply and demand. The auction process adjusts rents and bids until all agents are located without incentive to move. In the end, we did not include the firm rent module into the final land use model because of the lack of data.

Our LU model version allocates 12 types of households (defined by age, income, size), and 11 types of firms (defined by industry) to 12 kinds of real estate units (e.g., single family, large apartment) in each traffic analysis zone (TAZ). We estimated household and firm location choice models using the 2010-2011 Massachusetts Travel Survey (2010MTS), InfoGroup data, MAPC/MassGIS parcel data, and Census data. The household location choice model has a rich representation of zonal attributes, such as accessibilities to jobs and shopping, race, SAT scores, income, crime, FAR, and taxes, and household characteristics including size, age, income, student, children etc. The firm location choice model includes variables such as distance to highway entries, accessibility to population and employment, density, job density by sectors, FAR etc. We calibrated the land model for 2010.

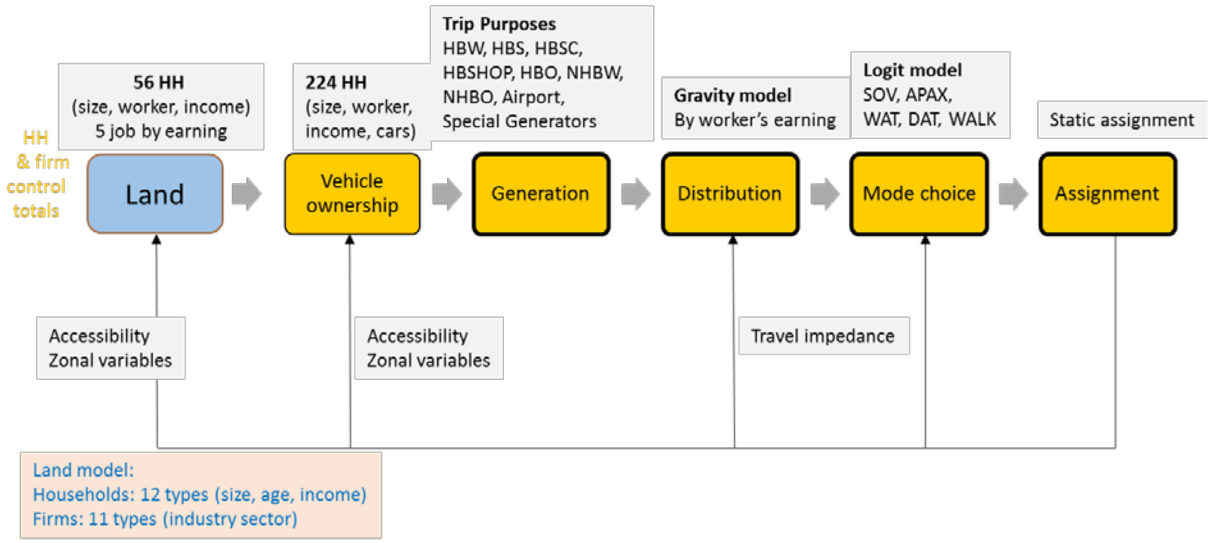
Uncertainty Analysis: Land Use Model
In the land use model Posada (2015) examined data-related uncertainty (how model estimation changes with different data sources) as well as temporal transferability (how do preferences change over time). The data-related uncertainty was analyzed in the housing model (for 2010) and the analysis indicates that the models are sensitive to the specific dataset used in the estimation. That is, two data sets that represent the same population in the same period of time can result in two different model estimations. The accuracy of the different models' estimations seems to be correlated to the category size (number of observations of the individual agent categories) in the sample data that was used for the estimation. Preference change uncertainty was examined in the firm location choice model (for 2000 and 2010). The location choice models suggest that firms' willingness to pay for clustering has changed over time. In 2000, the firms that valued proximity to jobs in the same industry were those in the government and other service sector, while in 2010 it was firms in the professional services, government, and education and health sectors

We tested the LU model accuracy by back-casting land use in 1990 using the model calibrated for 2010. The model inputs were observed zonal variables in 1990. The projection for households performs well for the residential market but not so well for the firm market. Several reasons may cause the relative inaccuracy in firm projections: more behavioral uncertainty may exist in firm location choice; the assumption of every firm agent (i.e. industry) having the same location strategy regardless of the firm size fails to capture intra-industry differences; the lack of firm rent price and firm numbers introduce input uncertainty. Furthermore, many variables available in 2010 (base year for model) were not available in 1990, which adds input uncertainty to the backcast.

UNCERTAINTY ANALYSIS: LUT MODEL

The land and transport models are linked by accessibility measures, household and firm locations, and other zonal variables derived from number of agents and travel skims, updated during each model iteration (Figure 1). The models are integrated via python scripting which automates the full model system runs, including the necessary translation between the different zonal structures and agent types across the land use and transport models and updating of the endogenous variables (e.g. accessibility) based on each model's outputs. In the available computational environment, each complete LUT loop required four hours and we ran 5 loops per scenario, meaning each model run required 20 hours (the majority of computational consumption came from the network assignment stage).

Figure 1. LUT Model Platform



Given the computational constraints, we used the LUT model system to only demonstrate one source of exogenous uncertainty: sea level rise. We used a simplified evaluation framework, utilizing relatively traditional measures of vehicle hours of travel. Additional performance indicators, such as accessibility, were not included due to time constraints. In the end, we were only able to demonstrate how a land use-transport (LUT) model can be used to forecast the short and longer term impacts of potential 4-foot sea level rise in Greater Boston by 2030 (Han et al., 2017). The short-term scenario represents the immediate transport system response to inundation, which provides a measure of resiliency in the case of an extreme event, such as a storm surge. In the short run, the results reveal that transit captive users will suffer more. Transit, in general, displays less resiliency, at least in part due to the center city's vulnerability and Boston's radial-focused transit system. Trip distances would modestly decrease, and average travel speeds would go down by over 50%. Rail transit ridership would be decimated and overall transit usage would go down by 66%. The longer term scenario aims to predict how households and firms would prefer to relocate in the "new equilibrium" where over ten square miles of land disappear and the transport network inundations become permanent. Assuming no supply constraints, new residential growth centers would emerge on the peripheries of the inundated zones, primarily in the inner-core suburbs. Some regional urban centers and traditional industrial towns would boom. Firms would be hit harder, due to their heavy concentration in the inner-core; firm relocation would largely follow households. Transit usage would again be decimated, but walking trips would increase.

Limitations

The project faced numerous technical challenges, including: data acquisition and model development for the land use model, ongoing challenges in integrating the land use-transport models, model run times (and de-bugging) for both models, and calibration and validation over-time (for model backcasting and forecasting). As mentioned, the prohibitive costs of running the software in a cloud-based server system, limited us to a single, relatively high-powered machine with 20 dual core I7 chips (40 threads) which allowed parallelization of model processes. Even then, the 2030 sea level rise scenarios required 20 hours per model run. This computing constraint limited the types of scenarios we could reasonably explore. In short, we were unable to reasonably explore the original idea of operationalizing the concept of how the "future is big data."

Publications/Presentations/Websites

Theses

Chheda, R (2014) Negotiating the Neighborhood: Modeling the Relationship Between the Built Environment and Transit Choice. Thesis, Master of Science in Transportation, MIT, September. http://dspace.mit.edu/handle/1721.1/95572

Dowd, M. (2015) Modeling Inundation Impacts on Transportation Network Performance: A GIS and Four-Step Transportation Modeling Analysis. Thesis, MST& MCP, Massachusetts Institute of Technology, Department of Urban Studies and Planning, June. http://dspace.mit.edu/handle/1721.1/99567

Han, Y. (2015) Temporal transferability assessments of vehicle ownership models and trip generation models for Boston Metropolitan Area. Thesis, MST& MCP, Massachusetts Institute of Technology, Department of Urban Studies and Planning, June. http://hdl.handle.net/1721.1/99566

Posada, P (2015) Location, location, location choice models. Thesis, M.C.P., Massachusetts Institute of Technology, Department of Urban Studies and Planning, June. http://dspace.mit.edu/handle/1721.1/99091

Peer Reviewed Journal Articles

Han, Y., Zegras, C. (2016) Exploring Model and Behavior Uncertainty: Temporal Transferability Assessment of Vehicle Ownership Models for Boston, Massachusetts, Metropolitan Area. *Transportation Research Record 2563,* pp. 122-133.

Han, Y., C. Zegras, V. Rocco, M. Dowd, M. Murga. When the Tides Come, Where Will We Go? Modeling the Impacts of Sea Level Rise on the Greater Boston, Massachusetts, Transport and Land Use System. Transportation Research Record: Journal of the Transportation Research Board, No. 2653, 2017, pp. 54–64. http://dx.doi.org/10.3141/2653-07

Presentations

Zegras, C. (2015) Starting to understand what we don't know: Uncertainty in urban mobility systems analysis. The Boston Case. Invited presentation at Singapore University of Technology and Design, Singapore, June.

Han, Y., Zegras, C (2015) Temporal transferability of vehicle ownership and trip generation

models for Boston metropolitan area. Presented at CUPUM 2015 (Computers in Urban Planning and Urban Management), Cambridge, July.

Han, Y, Zegras, C. (2016) Exploring Model and Behavior Uncertainty: Temporal Transferability Assessment of Vehicle Ownership Models for Boston, Massachusetts, Metropolitan Area. Presented at the Annual Meeting of the Transportation Research Board, Washington, DC, January.

Han, Y. Zegras, C., Rocco, V., Dowd, M., Murga, M. (2017) When the Tides Come, Where Will We Go? Modeling the Impacts of Sea-level Rise on Greater Boston's Transport and Land Use System. Presented at the Annual Meeting of the Transportation Research Board, Washington, DC, January.

Websites

- Visualizations of sea-level rise impacts can be seen here: http://mdgis.github.io/.

- General project overview website: http://mfc.mit.edu/strategically-adaptive-sustainable-mobility-systems.

- The Model system, including specification, estimation, validation, and forecasting is described here: http://web.mit.edu/czegras/www/Strategic_model.html.