UNCERTAINTY ANALYSIS AND ITS IMPACTS ON DECISION-MAKING IN AN INTEGRATED TRANSPORTATION AND GRAVITY-BASED LAND USE MODEL

Jennifer Duthie (corresponding author)
Division of Statistics and Scientific Computation, University of Texas at Austin, 2.104A WCH Building, 1 University Station, G2500, Austin, TX 78712
Email: jen.duthie@ssc.utexas.edu
Fax: 512-232-1045

Avinash Voruganti
Department of Civil, Architectural, and Environmental Engineering, University of Texas at Austin, 6.202 ECJ Hall, 1 University Station, C1761, Austin, TX 78712
Email: avinash.voruganti@gmail.com
Fax: 512-475-8744

Kara Kockelman
Department of Civil, Architectural, and Environmental Engineering, University of Texas at Austin, 6.904 ECJ Hall, 1 University Station, C1761, Austin, TX 78712
Email: kkockelm@mail.utexas.edu
Fax: 512-475-8744

S. Travis Waller
Department of Civil, Architectural, and Environmental Engineering, University of Texas at Austin, 6.204 ECJ Hall, 1 University Station, C1761, Austin, TX 78712
Email: stw@mail.utexas.edu
Fax: 512-475-8744

The following paper is a pre-print and the final publication can be found in *Journal of Urban Planning and Development*, Vol. 136 (4):294-302, 2010.

ABSTRACT

It is well established that uncertainty is present in all aspects of transportation and land use planning. Sources of uncertainty range from technological advances to changing economic conditions to enacted policy to errors in data and model structure. While much research has been devoted to analyzing the variation in model outputs due to uncertainty, little has been done to quantitatively answer the more important question of how decision making will change based on recognition of this uncertainty. This paper aims to begin to fill this gap by evaluating how roadway investment decisions will differ depending on whether or not uncertainty is recognized. Population and employment control totals, as well as trip generation parameters, are found via antithetic sampling and a full feedback integrated gravity-based land use and four-step travel model is used. It is found that the ranking of improvement projects may indeed be different if uncertainty is considered relative to treating all parameters and data as deterministic. Results for the experimental analysis conducted in this paper found this percent difference to be between 4% and 25% depending on the performance metric used: total system travel time, vehicle miles traveled, total delay, average network speed, and standard deviation of network speed were all examined.

INTRODUCTION

Major transportation improvements are costly investments, often made with the help of mathematical models to anticipate future transportation conditions. Given budget constraints and the wide-reaching and long-term nature of major project impacts, selection of system improvements should be done with care. This paper focuses on how variations in model parameters and inputs affect land use and transportation predictions, and how the formal recognition of uncertainty in integrated transport-land use modeling exercises may change roadway capacity expansion decisions, thereby resulting in decision-making that is more robust to future conditions. The model used in this paper is a very basic one with a relatively short running time that allows for a large number of realizations and improvement scenarios to be analyzed, without ignoring the feedback between land development and travel costs that provides a more realistic version of the future.

Integrated modeling of land use and transport is complex, and relatively few studies have evaluated the effects of input and parameter uncertainty in this context. Those that have done so focus primarily on the variation in model outputs. For example, Smith and Shahidullah (*1*) evaluated the performance of 10-year population forecasts (a key land use model input) by age group for three very different census tracts, and found mean absolute percent errors to range from 15% to 20%. Zhao and Kockelman (*2*) measured the propagation of uncertainty through a four-step travel demand model and concluded that trip assignment (to the network) serves to reduce uncertainty levels in outputs to initial levels of input uncertainty (rather than widening, for example), and that employment estimates, trip generation rates, and mode choice parameters are critical determinants of output variations (e.g., link flows and total VMT) – and that such final outputs can be closely predicted via relatively few input values (using a regression of their 100 simulated outputs). Pradhan and Kockelman (*3*) extended Zhao and Kockelman's study by examining propagation through an integrated transportation and land use model (ITLUM) using UrbanSim and a four-step travel model for Eugene, Oregon; they concluded that in the long run, only those inputs that have a cumulative effect (such as population and employment growth rates) are likely to have a significant effect on model outputs.

Krishnamurthy and Kockelman (*4*) conducted another study of uncertainty in ITLUMs, this time using a Disaggregated Residential Allocation Model (DRAM®) and Employment Allocation Model (EMPAL®) based on the DRAM®-EMPAL® Lowry-type models developed by Putman (*5*), along with a four-step travel model in TransCAD. This research was conducted on the three-county Austin metropolitan region and found output variations to be most sensitive to the exponent (of the volume-to-capacity ratio) in standard link performance functions, the split of trips between peak and off-peak periods, and trip generation and attraction rates.

More recently, Sevcikova et al. (*6*) illustrated the benefits of Bayesian Melding over simple random sampling in their quest to calibrate traffic analysis zone-level household counts using UrbanSim for Eugene, Oregon. Random numbers were used within the economic and demographic transition models, employment and household mobility models, employment and household location choice models, and the land development model for a total of 843 uncertain parameters. Two sets of 100 simulations were run, using a different random number seed for each set. The Bayesian Melding procedure led to wider confidence intervals for various outputs that are more likely to contain the true result than if a simple random sampling procedure had been used. Related to this, Gregor's (*7*) new land use model (LUSDR) makes uncertainty explicit through multiple model runs, in order to find the best extension of Jackson County, Oregon's urban growth boundary.

While this study may be one of the first to consider how investment decisions change given uncertainty in an ITLUM context, others have considered this topic in the travel model context, particularly with respect to traffic assignment. Lam and Tam (*8*) used Monte Carlo simulation methods to study the impact of uncertainty in traffic and revenue forecasts for road investment projects. They assumed normal distributions for each of several uncertain parameters, including population and demand elasticity. Moreover, Waller et al. (*9*) assigned independent distributions for each origin-destination (OD) pair's future year demand in three test networks (ranging from two to 100 OD pairs) and demonstrated how assignment models relying on expected values of all inputs will tend to underestimate future congestion and may (in 14% of cases studied) lead to selecting projects with higher average (future) travel costs (i.e., lower net benefits) than ideal, and higher variance in such costs (which implies more risk). Duthie et al. (*10*) extended this earlier research to allow for correlations in trip-making between OD pairs and showed how neglecting correlations when they exist will lead to errors in future travel cost predictions and suboptimal project selections (2% to 50% of the time, depending on the correlation structure and the objective function used). Also, Rodier and Johnston (*11*) found that errors in county population forecasts can affect whether or not the Sacramento region meets air quality conformity. The authors' focus, however, was sensitivity analysis (as opposed to uncertainty analysis), so uncertain parameters (i.e., population growth, income, and fuel prices) were only varied by a few percentage points, and counties with outlying population growth rates were eliminated. Harvey and Deakin (*12*) considered uncertainty in population growth, fuel prices, and household income levels in their Short-Range Transportation Evaluation Program model and found that plausible ranges of the input variables resulted in VMT values that differed by -25% to 15% relative to the original prediction.

This work builds off the existing literature and seeks to begin to fill the gap that exists in regards to improved land use and transport decision making. The great need for incorporating uncertainty into the decision making process has been noted by many researchers and practitioners (*13-15*) and the types of uncertainties presented in ITLUMs have been categorized many different ways (*16-18*). The research presented in this paper focuses on the uncertainty inherent in regional population and employment forecasts, as well as parametric uncertainty, and uses a Lowry-type land use model and a four-step travel model to appreciate how decision making may differ when uncertainty is recognized.

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

THE MODEL

The integrated transportation and gravity-based land use model (ITGLUM) used in this paper is described in this section. The transportation module of ITGLUM (contained within Figure 1's bottom gray box) is a very basic four-step model of trip generation, trip distribution, mode choice, and traffic equilibrium. (Readers may wish to review Martin [*19*] for a description of these steps.) The land use module of ITGLUM (contained within Figure 1's top gray box) is based on Putman's documentation of ITLUP equations (*5*), which represent a disaggregated residential allocation Model (DRAM®) and employment allocation model (EMPAL®). The land allocation module, LUDENSITY, was developed by the authors (however, its placement in the model series is based on Putman's LANCON model [*5*]). (An earlier alternate version of the land use model that was helpful in the development of ITGLUM can be found here [*20*]). Calibration determines the value of parameters that generate the best forecast for the base year based on the lagged year data. Each sub-module is described below.

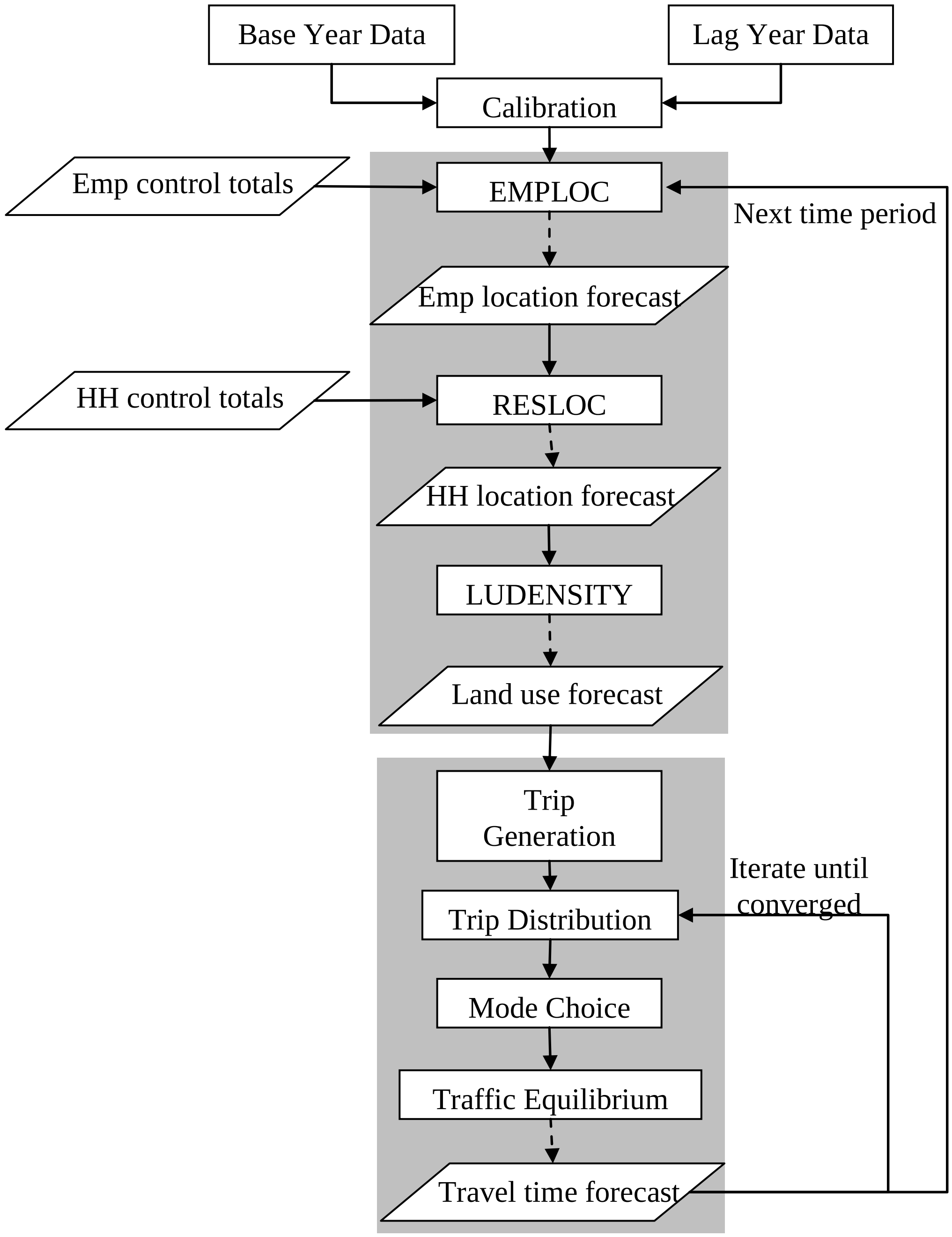

FIGURE 1 Diagram of ITGLUM

Gravity-based Land Use Model Description

The residential land use allocation module, RESLOC, uses Equation 1 and the employment land use allocation module, EMPLOC, uses Equation 2. The following indices are used: *n* – household type (low income, medium-low income, medium-high income, high income), *i* and *j* – travel zone, *k* – employment sector (basic, retail, service), and *t* – time period. Note that for each household type, the following equation holds:

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

where $a_{k,n}$ is the regional average ratio of type *n* households per type *k* employees, *u* is the unemployment rate, *E* is employment, and *Q* is a placeholder that is used in Equation 7. Other notation needed is *c* – travel cost between zones, *x* – the percentage of developable land already developed, L^{res} – residential land, and *N* – number of households. Parameters determined during calibration are η (constrained to vary between zero and one), ρ, *r*, *q*, *b*, and φ. Employment by sector and households by type are constrained by regional control totals and are scaled accordingly if a control total is not met.

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

$$E_{j,t}^{k} = \eta^{k} \sum_{i} \frac{\left(E_{j,t-1}^{k}\right)^{\varphi^k} \exp\left(\rho^{k} c_{i,j,t-1}\right) \sum_{n'} N_{i,t-1}^{n'}}{\sum_{j'} \left(E_{j',t-1}^{k}\right)^{\varphi^k} \exp\left(\rho^{k} c_{i,j,t-1}\right)} + \left(1-\eta^{k}\right) E_{j,t-1}^{k} \quad (2)$$

The land allocation module, LUDENSITY uses Equation 3, which allows for an exponential decrease in the amount of land allocated to each use as the number of uses in a zone increases. The index l denotes the use (basic employment, commercial employment [retail and service], or residential [all income levels]), ψ denotes the amount of use, β are parameters determined through calibration, and δ is a constant determined based on L_0 - an initial condition for the amount of land devoted to a use if the zone is empty. If the amount of allocated land exceeds zone size, then it is assumed that multi-story development is taking place.

$$L_{j,t}^{l} = \frac{\psi_{j,t}^{l} - \psi_{j,t-1}^{l}}{\beta_1 \left(\left(\sum_{l'} \psi_{j,t}^{l'} \right) + \delta^{l} \right)^{\beta_2}} + L_{j,t-1}^{l} \quad (3)$$

Travel Demand Model Description

The travel demand model consists of four steps (trip generation, trip distribution, mode choice, and traffic assignment) and is based on guidance given in National Cooperative Highway Research Program Report 365 "Travel Estimation Techniques for Urban Planning (*19*). In each time period, the three final steps (trip distribution, mode choice, and traffic assignment) are iterated until equilibrium is achieved between the travel demands and costs.

Each step is described below.

Trip Generation

The trip generation step determines the number of trips produced from and attracted to each zone. A look-up table is used (Table 1) based on Martin (*19*) to obtain trip production rates by purpose (home-based work, home-based other, and non-home-based) and by income level.

TABLE 1 Daily Person Trip Production Rates by Purpose and Income Levels

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

Trip attractions by trip type (Equation 4) are determined using modified versions of the equations provided on page 28 of Martin (*19*). A_j is the number of trips attracted to zone *j*, which can be further categorized by trip purpose: home-based work (*HBW*), home-based non-work (*HBNW*), and non-home-based (*NHB*). A_j is a function of the following zone *j* characteristics: total employment, retail employment, basic employment, service employment, and number of households.

$$\left.\begin{aligned} A_{j,HBW} &= 1.45*TOTAL\_EMPLOYMENT \\ A_{j,HBNW} &= 9.00*RETAIL+0.5*BASIC+1.7*SERVICE+0.9*TOTAL\_HH \\ A_{j,NHB} &= 4.10*RETAIL+0.5*BASIC+1.2*SERVICE+0.5*TOTAL\_HH \end{aligned}\right\} \quad (4)$$

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

Trip Distribution

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

$$T_{i,j} = \frac{P_i A_j e^{-\theta c_{i,j}}}{\sum_{j'} A_{j'} e^{-\theta c_{i,j'}}} \quad (5)$$

where $T_{i,j}$ is the flow of trips from zone *i* to zone *j* and P_i is the number of trips produced from zone *i*. The mean value of the parameter θ is, based on guidance from Martin (*19*), 0.125 for home-based work trips, and 0.1 for non-home-based work trips and for non-work trips. Intrazonal travel times are assumed to be a function of zone size, as follows (*19*):

$$c_{i,i} = \frac{0.5 * 3600\left[\frac{\sec}{hr}\right] * \sqrt{ZoneArea\left[mi^2\right]}}{15\left[mph\right]} \quad (6)$$

Since the travel demand model is iterated until convergence (when network and demand equilibria are achieved), the number of trips between each origin-destination pair reflects the travel costs of all travel options.

Mode Choice

A multinomial logit choice model is used to determine the share of trips that use transit, shared ride, and drive alone modes (denoted by indices *TR*, *SR*, and *DA* respectively). The utility equation for each mode is given by Equation 7 and the coefficients are based on recommendations from Martin (*19*). All travel times including in-vehicle travel time (*IVTT*) and out-of-vehicle travel time (*OVTT*) are assumed to be in minutes and *COST* is in units of cents.

$$\left.\begin{aligned} U_{TR} &= -0.025 IVTT_{TR} - 0.050 OVTT_{TR} - \beta_{COST} COST_{TR} \\ U_{SR} &= -0.025 IVTT_{SR} \\ U_{DA} &= -0.025 IVTT_{DA} \end{aligned}\right\} \quad (7)$$

Here, β_{COST} is a function of the origin zone's income distribution and an assumed value of time as percentage of income. To simplify the model only fare costs were considered. The table on page 66 of Martin (*19*) was used to determine appropriate values of this coefficient. $IVTT_{TR}$ was assumed to be twice $IVTT_{DA}$ plus twelve minutes (to account for out of vehicle travel time) (*21*), $IVTT_{SR}$ is assumed to be 150% of $IVTT_{DA}$, $OVTT_{TR}$ is assumed to be ten minutes, $COST_{TR}$ is assumed to $0.50. The share of trips by each mode, *m*, is calculated using Equation 8, where U_m denotes the utility for mode m.

$$\text{Share of mode } m = \frac{e^{U_m}}{\sum_{m'} e^{U_{m'}}} \quad (8)$$

Traffic Equilibrium

Person flows are converted to vehicle flows using the following equivalents: 7.1 for transit, 2 for shared ride, and 1 for drive alone. Additionally, bus flows are converted to passenger car flows by multiplying by a factor of 3 because of the additional space buses consume on roadways and their increased impact on capacity. The total passenger car units are assigned to the road network assuming user equilibrium (UE) behavior; no user can unilaterally switch routes without incurring a higher cost (*22*). A Bureau of Public Roads type link cost function is used as shown in Equation 9, where *tt* is a vector of link travel times, *v* is a vector of link volumes (*tt(0)* is a vector of link travel times when volume is zero), μ is a vector of link capacity, and α_1 and α_2 are parameters.

$$tt(v) = tt(0)\left(1 + \alpha_1 \frac{v}{\mu}\right)^{\alpha_2} \quad (9)$$

Link capacity is a function of original capacity, μ^0, lane additions, y, and single lane capacity, $\mu(1)$. The solution method for UE (originally formulated by Beckmann et al. [*23*]) is as follows, and more details can be found in Sheffi (*24*). First, find the shortest path between each OD pair using Dijkstra's algorithm (*25*), then average the current path flows with previous path flows using the golden section method. Calculate the new path costs based on these new flows, and check for convergence. If not converged, find shortest paths as before and repeat these steps.

METHODOLOGY

This research focuses on the decision-making impacts of uncertainty in future regional control totals of households (by income group) and employment (by sector), and in the trip generation parameters. The type of decision evaluated is the selection of roadway segments for improvement within an allotted budget. Twenty-four pairs of roadway improvements, where each improvement is approximately two miles in length, are considered. For each pair of improvements, the superior one is selected based on one of the following metrics: total system travel time (TSTT), vehicle miles traveled (VMT), total delay, average network speed, and standard deviation of network speed. The decisions made when the uncertainty is neglected are then compared to the decisions made when uncertainty is considered. This section presents how uncertain data and parameters are sampled, how candidate roadway improvements are selected, and how comparisons are made between decisions made in the deterministic and uncertain analyses.

Sampling Uncertain Data and Parameters

Each uncertain parameter is assumed to follow a lognormal distribution with coefficient of variation equal to 0.3 (as tested in Zhao and Kockelman [*2*]). (Future work may be helpful in ascertaining the impact of correlations between the parameters and inputs – for example, across regional population and employment control totals, in order to approximate economic booms and busts). Sampling is done via the antithetic technique, which is a popular and well studied variance reduction method. Antithetic sampling is expected to work best under a monotonic system response and a symmetric random variable (*26-28*), which implies it may perform well for the system response of consideration in this paper (since increased population and employment should lead to increased travel and travel cost). Antithetic has also been shown to work well for the traffic assignment problem (*10*).

In antithetic sampling, pairs of negatively correlated realizations of the uncertain parameters, denoted by $\tilde{\xi}^{\varphi}$ and $\tilde{\xi}^{\varphi'}$, are used to obtain an estimate of the expected value of the function $F(\tilde{\xi})$, which in this example could represent performance measures such as future roadway speeds. Negative correlation implies that when one realization of the random parameter is low, the other in the pair should be high, leading to a variance reduction effect. The average of

$F(\tilde{\xi}^{\hat{\omega}})$ and $F(\tilde{\xi}^{\hat{\omega}'})$, i.e., $F\left(\tilde{\xi}^{\omega}\right)=\dfrac{F(\tilde{\xi}^{\hat{\omega}})+F(\tilde{\xi}^{\hat{\omega}'})}{2}$, is evaluated for each negatively correlated pair of uncertain parameters. The estimate of the expected value of the function $F(\tilde{\xi})$ over points sampled from Ω $(\omega \in \Omega)$ is the sample average $\overline{F}$. The unbiased estimator of $V\left[F\left(\tilde{\xi}\right)\right]$ is the sample variance, s^2.

Note that the negative correlation between each pair of realizations for the uncertainty set is different from assuming correlations exist among the members of the uncertainty set. Antithetic sampling is relatively easy to implement depending on the complexity of drawing a sample from the population distribution and confidence intervals can be determined using standard statistical procedures as each realization of $F(\tilde{\xi})$ can be assumed to be independent and identically distributed.

Selecting Candidate Roadway Improvements

Candidate roadway improvements (lane additions) are selected from the subset of links with a volume that is at least 85% of μ^0 25 years into the future. Let Γ be the subset of links that meet this criteria and let each link in this set be indexed from $\gamma = 0\ldots\Gamma$ (ordering by v/c is not necessary). Let G_γ be the length of link γ, G_{max} be the maximum length of any expansion project, and y be the vector of lane additions. Also, let *count* be a counter that constrains the maximum number of attempts ($count_{max}$) to find a link to improve that meets the budget requirement.

1) Solve for UE on the network to get link flows, $v_\gamma \forall \gamma \in \Gamma$.
2) Assign two parameters, *LB* and *UB* to each link $\gamma \in \Gamma$ where
$$LB_\gamma = \sum_{\gamma'=1}^{|\gamma|-1}\left(\frac{v_{\gamma'}}{\mu^0_{\gamma'}}\right),$$
and
$$UB_\gamma = \sum_{\gamma'=1}^{|\gamma|}\left(\frac{v_{\gamma'}}{\mu^0_{\gamma'}}\right).$$
3) Set $y_\gamma = 0 \ \forall \gamma \in \Gamma$ and *count* = 0.
4) Choose a random number, $R \in (0, V)$ where
$$V = \sum_{\gamma\in\Gamma}\left(\frac{v_\gamma}{\mu^0_{\gamma'}}\right).$$
5) Select link γ corresponding to
$LB_\gamma < R \le UB_\gamma$.
6) If $G_\gamma < G_{max}$, $y_\gamma = y_\gamma + 1$, *count* = 0, and $G_{max} = G_{max} - G_\gamma$. Else *count* = *count* + 1.
7) If $K > 0$ and $count < count_{max}$, go to Step 4. Else $\mu = \mu^0 + y\mu(1)$.

In Step 1 ITGLUM is run using mean values for all inputs and parameters to determine the links that are in set Γ. (Restricting the set of links to improve was done to achieve a higher degree of realism in terms of projects that may be selected.) Step 2 evaluates two parameters for each link: one representing the sum of the volume to capacity ratio of all links assigned a lesser cardinal index, $|\gamma|$, and a second to represent the sum of the volume to capacity ratio of all links assigned a lesser than or equal cardinal index. Step 3 initializes y and *count* to zero. A random number from zero to the sum of the volume to capacity ratios for all links in Γ is generated in Step 4, and then in Step 5 this random number is translated to a specific link based on the parameter values calculated in Step 2. If money remains in the budget (where the budget is in terms of lane length added) as given by the constraint in Step 6, then this specific link is improved and the budget is subsequently decreased by an amount proportional to the length of the improved link. If budget remains for further improvements (Step 7) and $count_{max}$ has not been reached, another link is chosen according to the same procedure. When these conditions are no longer true, the vector of link capacities is updated to include the selected improvements.

Comparing Results of Deterministic and Uncertain Analyses

As mentioned earlier, five performance metrics are considered (as expected values for the uncertainty analysis): TSTT, VMT, total delay, average network speed and standard deviation of network speed. The last two are combined into a single metric by applying a weight of (1-w) (where $w \in (0,1)$) to average network speed and a weight of w to the negative of standard deviation of network speed and summing the two terms. The weight, w, is then varied across its range in increments of 0.1. Each of these performance metrics were evaluated to mimic the variety of such measures used in practice and because no one measure gives a complete portrayal of network conditions.

The weighted sum of the average and standard deviation of network speed is likely the best metric used because, if the weight is non-zero, it rates a network poorly that is very congested in some areas and nearly free flowing in others – a situation that may lead to a good rating if only the average speed was considered. TSTT and VMT are common measures, but can be problematic since mode and destinations are flexible in the ITGLUM. Total delay is a useful comparative measure since it is sensitive across different strategies, but cannot differentiate networks with heavy congestion in some parts and free flow on the others and networks with uniform congestion.

For the uncertainty analysis, let $Y^{\kappa}_{\phi-1,\phi}\left(\tilde{\xi}^{\omega}\right)$ be the difference in value of performance metric κ for uncertainty realization ω of network improvement pair $(\phi\text{-}1,\phi)$. For example, $Y^{TSTT}_{\phi-1,\phi}\left(\tilde{\xi}\right) = TSTT_{\phi-1}\left(\tilde{\xi}\right) - TSTT_{\phi}\left(\tilde{\xi}\right)$. If the number of samples is large (> 40 since the population standard deviation of Y is unknown [27]), the sample average $\overline{Y}^{k}_{\phi-1,\phi}$ can be assumed to behave according to a normal distribution, allowing the formulation of confidence intervals (CI's). If all points in the 95% CI around $\overline{Y}^{k}_{\phi-1,\phi}$ are greater than zero, then improvement ϕ *-1* is said to outperform improvement ϕ for metric κ in the uncertainty analysis if a higher value equates to a better performance metric, otherwise ϕ is said to outperform ϕ *-1*. Similarly, if all points in the

95% CI around $\overline{Y}^{k}_{\phi-1,\phi}$ are less than zero, then improvement ϕ is said to outperform improvement ϕ*-1* if a higher value equates to a better performance metric, otherwise ϕ*-1* is said to outperform ϕ. If the 95% CI around $\overline{Y}^{k}_{\phi-1,\phi}$ contains zero, then neither improvement is considered better. For the deterministic analysis, the winning improvement for each metric is trivial to calculate since there is only one realization of the parameters and data. The winning improvements for each metric are then compared between the uncertainty analysis and the deterministic analysis.

EXPERIMENTAL ANALYSIS

As shown in Figure 2, a highly idealized sample region (with zones denoted by gray lines) and roadway network (with nodes and links denoted by black circles and lines respectively) allow for ready analysis while helping clarify the impact and sources of modeling results. (Modeling an actual region, such as Austin, would likely obscure the sources of differences in deterministic and stochastic investment results.)

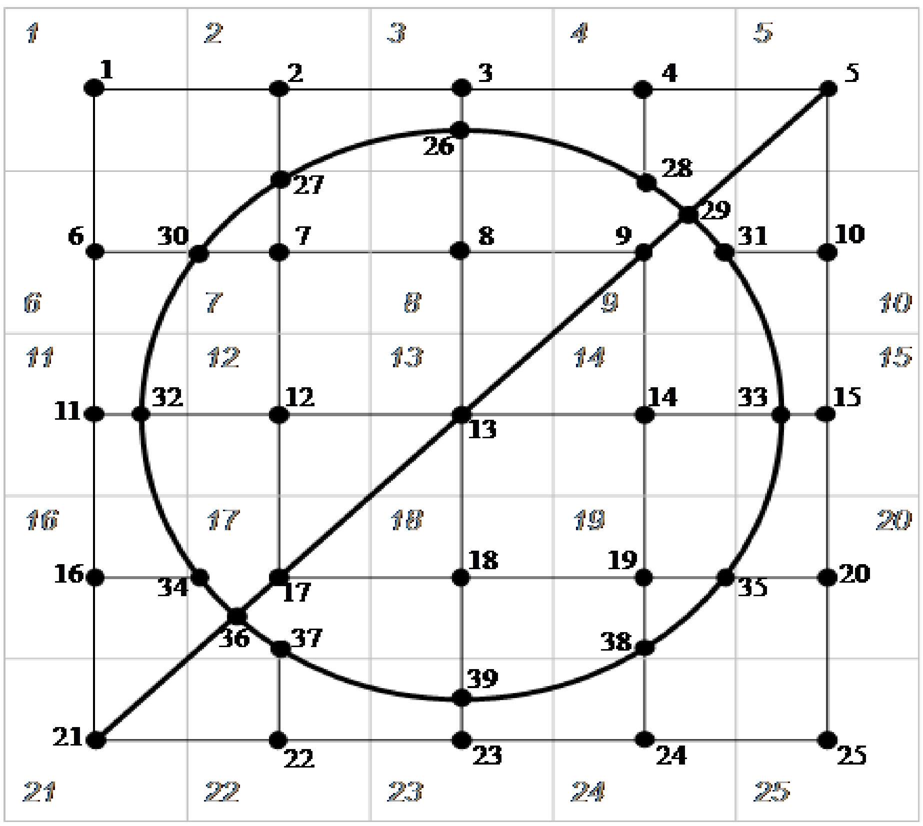

FIGURE 2 Sample Region and Network

Each zone is 1 square-mile (or 640 acres). Highways (circumferential and diagonal) are denoted by heavy black lines and local roads have lighter black lines. Capacity on local roads is 400 vehicles per hour per lane (vphpl) and capacity on highways is 1600 vphpl. All roads are bi-directional, having one lane in each direction. Free flow travel speed is 60 miles per hour (mph) on highways and 30 mph on local roads. Mean values of regional control totals for employment by sector and household by income, developed using rules of thumb for "reasonable" development, are given in Table 3. The base year data in Tables 6 and 7 were developed in a similar manner.

TABLE 3 Regional Control Totals

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

Table 4 gives the parameters for the EMPLOC and RESLOC sub-models obtained by calibrating Equations 1 and 2 using assumed base and lag year conditions, as presented in Tables 6 and 7 (which also were created using rules of thumb for “reasonable” development). Table 5 gives the parameters for the LUDENSITY sub-model obtained from calibrating Equation 3. Equation 3’s L_0 value was assumed to be 0.5 acres per basic job, 2.5 acres per commercial job, and 6 acres per household (based on approximations of the maximum amount of land typically allocated to these uses).

TABLE 4 Calibrated Parameter Values for EMPLOC and RESLOC Sub-models

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

TABLE 5 Calibrated Parameter Values for the LUDENSITY Sub-model

	Land Use		
Parameter	**Basic**	**Commercial**	**Residential**
β_1	0.0021	0.1347	0.0001
β_1	1	0.5498	1.1182

TABLE 6 Base Year Employment and Households

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

TABLE 7 Base Year Land Allocations

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

For each network improvement, 200 antithetic samples were run, using common random numbers across improvements. Due to averaging of the negatively correlated samples, the total number of uncertain realizations sampled from Ω $(\omega \in \Omega)$ is 100. All modeling was done in MATLAB on a Dell Optiplex GX620 with 4 CPUs at 3.4 GHz each and 2GB of RAM. Each ITGLUM run required approximately one minute.

RESULTS

Results for each performance metric are presented in this section for each network improvement pair. Figures 3 through 5 show the results for TSTT, VMT, and total delay, respectively. The y-axis value is 1 if the first improvement in a pair performed best, 2 if the second improvement in a pair performed best, and 0 if neither performed better than the other in a statistically significant way. For TSTT, four improvement pairs (16.67% of the total) were ranked differently between the uncertainty analysis (UA) and the deterministic analysis (DA), including one pair that did not show a statistically significant difference (at the 95% confidence level) in TSTT in the UA. A similar result was seen for VMT, but this time three of the four pairs that were ranked differently did not have a statistically significant difference in their UA performance. No pair was found to be different for both TSTT and VMT, emphasizing the importance of choosing performance measures carefully. The results for total delay indicate that 21 improvement pairs (or 87.5% of the total) were ranked differently in the UA and the DA; however, 19 of these pairs showed statistically insignificant differences (using a p-value of 0.05) in the UA. It is notable that the two pairs which ranked differently in the UA and the DA and differ in the UA in a statistically significant way, are a subset of the three pairs that meet similar criteria for the TSTT metric. The DA results for TSTT and delay are nearly identical (as expected, since both measure travel time); but, because so few of the UA delay results are statistically significant, their UA results differ significantly.

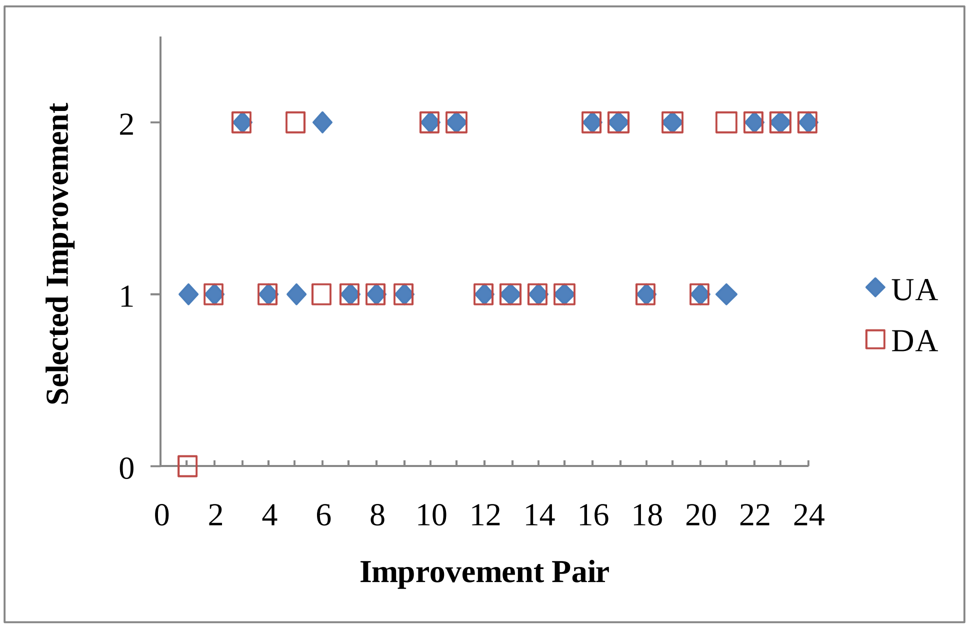

FIGURE 3 Minimizing Expected TSTT

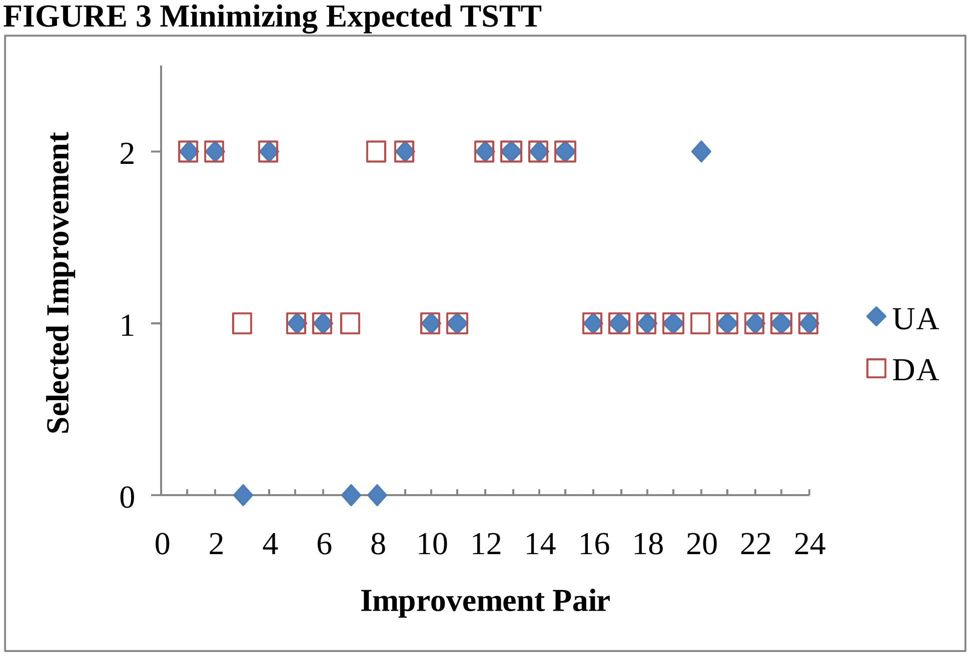

FIGURE 4 Minimizing Expected VMT

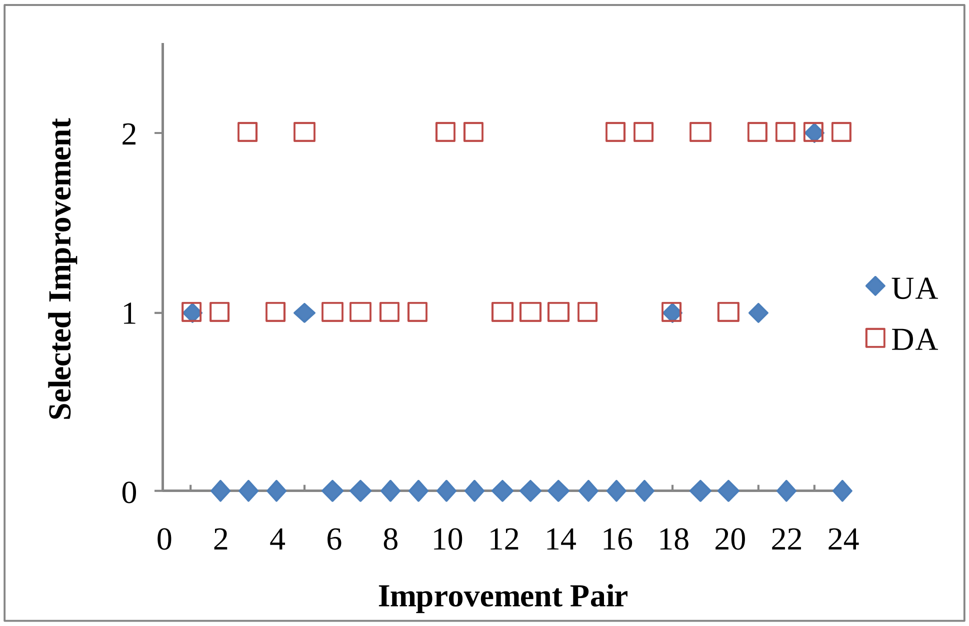

FIGURE 5 Minimizing Expected Delay

Since the network speed results are compared for 11 weights (*w* ranges from 0 to 1 in increments of 0.1), only the percentage of improvements that were ranked differently for each *w* are given in Figure 6. The percentage difference peaks when a 40 percent weight is applied to the standard deviation of network speed (and 60% to the average network speed). More tests are needed to determine if this result can be generalized across networks and coefficients of variances. Considering both the average and standard deviation is important because if only the average is considered, then it is possible that parts of the network are in near free-flow conditions while the other parts are very congested.

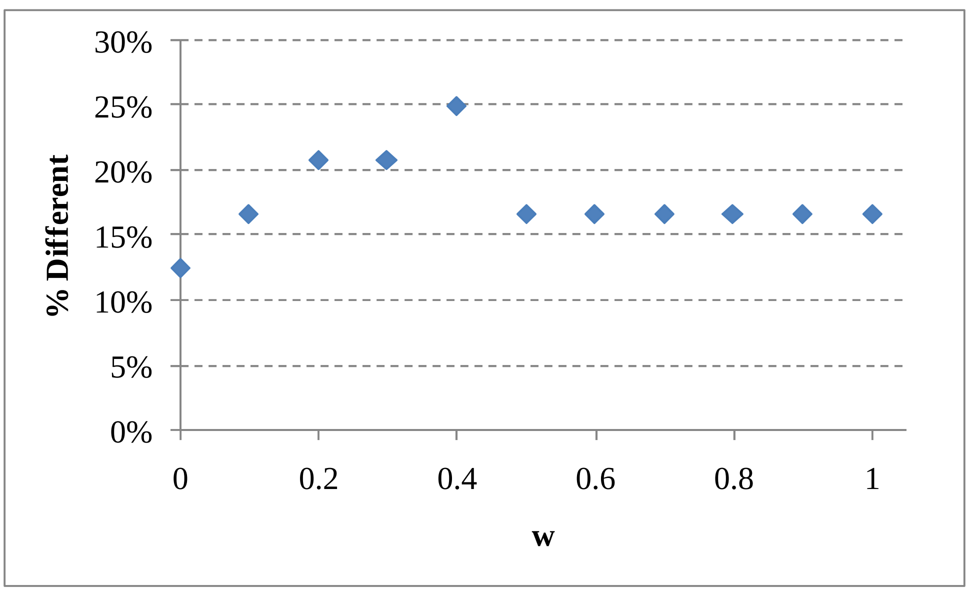

FIGURE 6 Percent of Trials where the Ranking of Network Improvement Projects

Differed when Minimizing Expected Weighted Sum of the Average Network Speed and Standard Deviation of Network Speed using UA and DA

CONCLUSIONS

In this paper, a very basic integrated gravity-based transportation and land use model (ITGLUM) was used to assess the impact of uncertainty on decision making, specifically the ranking of roadway improvements. The goal of using an ITGLUM with a simple structure was to gain insights into the problem of uncertainty and to evaluate emerging trends. Household and employment control totals, as well as parameters in the trip generation equations, were assumed to follow lognormal distributions with a 0.3 coefficient of variation. Antithetic sampling was used, and all parameters were assumed independent.

Results suggest that the ranking of network improvement projects often differs when recognizing uncertainty in model parameters and inputs. In evaluating pairs of links, UA and DA rankings differed significantly in just 4.17% of cases when using VMT as the criterion but 25% of cases when using a weighted combination of average network speed (60% weight) and the standard deviation of network speed (40% weight). Further research is needed to see what patterns emerge across different networks, land use patterns, and assumptions on uncertainty, but the results clearly indicate that neglecting uncertainty can lead to suboptimal network improvement decisions when certain common performance metrics are used.

ACKNOWLEDGEMENTS

The research contained in this paper was funded by Texas Department of Transportation project 0-5667, "Analysis and Guidelines for Establishing Unified Urban Land-Use and Transportation System Planning." The authors would also like to acknowledge Avinash Unnikrishnan for his assistance in developing the antithetic sampling code, and Brenda Zhou and Varun Valsaraj for their assistance in developing the land use model code.

REFERENCES

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

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

3. Pradhan, A. and K. Kockelman. Uncertainty Propagation in an Integrated Land Use-Transportation Modeling Framework: Output Variation via UrbanSim. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1805,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 128–235.

4. Krishnamurthy, S. and K. Kockelman. Propagation of Uncertainty in Transportation-Land Use Models: An Investigation of DRAM-EMPAL and UTPP Predictions in Austin, Texas. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1831,* Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 219–229.

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

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

7. Gregor, B. The Land Use Scenario Developer (LUSDR): A Practical Land Use Model Using a Stochastic Microsimulation Framework. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2003,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 93-102.

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

9. Waller, S.T., J.L. Schofer, and A.K. Ziliaskopoulos. Evaluation with Traffic Assignment under Demand Uncertainty. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1771,* Transportation Research Board of the National Academies, Washington, D.C., 2001, pp. 69-75.

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

11. Rodier, C.J. and R.A. Johnston. Uncertain Socioeconomic Projections used in Travel Demand and Emissions Models: could plausible errors result in air quality

nonconformity? *Transportation Research Part A*, 2002, Vol. 36(7), pp. 613-631.

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

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

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

15. Sussman, J., S.P. Sgouridis, and J.L. Ward. New Approach to Transportation Planning for the 21st Century: Regional Strategic Transportation Planning as a Complex Large-Scale Integrated Open System. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1931*, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 89-98.

16. Mahmassani, H. S. Uncertainty in Transportation Systems Evaluation: Issues and Approaches. *Transportation Planning and Technology*, Vol. 9, 1984, pp. 1-12.

17. Niles, J. S. and D. Nelson. Identifying Uncertainties in Forecasts of Travel Demand. Preprint for the Transportation Research Board's 80th Annual Meeting, Washington D.C., 2001.

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

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

20. Zhou, B., K.M. Kockelman, and J. Lemp. Transportation and Land Use Policy Analysis Using Integrated Transport and Gravity-Based Land Use Models. Proceedings of the 88th Annual Meeting of the Transportation Research Board and under review for publication in *Transportation Research Record*, 2009.

21. Lemp, J. Travel Demand Forecasting Models: Development, Application, and Comparison of Aggregate and Activity-Based Approaches for the Austin, Texas Region. M.S. Thesis, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, 2007.

22. Wardrop, J.G. Some Theoretical Aspects of Road Traffic Research. *Proceedings of the Institute of Civil Engineers II*, Vol. 1, 1952, pp. 325-378.

23. Beckmann, M., C. McGuire, and C. Winston. Studies in the Economics of

Transportation. Yale University Press, New Haven, Conn., 1956.

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

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

26. Cheng, R.C.H. The Use of Anti-thetic Variates in Computer Simulations. *Journal of Operational Research Society*, Vol. 33, 1984, pp. 229-237.

27. Law, A.M. and W.D. Kelton. *Simulation Modeling and Analysis*. McGraw-Hill Higher Education Press, 1992.

28. Devore, J.L. *Probability and Statistics for Engineering and the Sciences*. Brooks/Cole, 2004.