SENSITIVITY-BASED UNCERTAINTY ANALYSIS OF A COMBINED TRAVEL DEMAND MODEL

Chao Yang [a], Anthony Chen [a,b*], Xiangdong Xu [b], S.C. Wong [c]

[a] *Key Laboratory of Road and Traffic Engineering, Tongji University, Shanghai 201804, China*
[b] *Department of Civil and Environmental Engineering, Utah State University, Logan, UT 84322-4110, USA*
[c] *Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, China*

ABSTRACT

Travel demand forecasting is subject to great uncertainties. A systematic uncertainty analysis can provide insights into the level of confidence on the model outputs, and also identify critical sources of uncertainty for enhancing the robustness of the travel demand model. In this paper, we develop a systematic framework for quantitative uncertainty analysis of a combined travel demand model (CTDM) using the analytical sensitivity-based method. The CTDM overcomes limitations of the sequential four-step procedure since it is based on a single unifying rationale. The analytical sensitivity-based method requires less computational effort than the sampling-based method. Meanwhile, the uncertainties stemming from inputs and parameters can be treated separately so that the individual and collective effects of uncertainty on the outputs can be clearly assessed and quantified. Numerical examples are finally used to demonstrate the proposed sensitivity-based uncertainty analysis method for the CTDM.

Keywords: Uncertainty analysis; sensitivity analysis; nonlinear program; combined travel demand model

[*] Corresponding author. Tel.: +1-435-797-7109; fax: +1-435-797-1185.
E-mail: anthony.chen@usu.edu (A. Chen).

1. Introduction

Transportation planning and project evaluation are both based on travel demand forecasting, which is subject to different types of uncertainties (de Jong *et al*., 2007; Rasouli and Timmermans, 2012). These uncertainties stem from the predicted socioeconomic inputs, calibrated parameters, and the travel demand model itself (i.e., model structure and assumptions). Without considering uncertainty in the travel demand model, transportation planning, project evaluation, and investment decision are likely to take on unnecessary risk and any decisions based on these forecasts may be inaccurate and misleading (Zhao and Kockelman, 2002). Although transportation planners and decision makers may realize the existence of uncertainty, the vast majority does not employ any particular technique or methodology to systematically quantify it in the planning process. One of the reasons is that most of the existing procedures in the travel demand forecasting are deterministic, and there lacks a systematic methodology to conduct the uncertainty analysis of a travel demand model. Planners usually use *point estimates* of traffic forecasts in practice. However, they are not aware of the reliability and/or risk of the *point estimates*. Variance and confidence interval are typical ways used to quantify the reliability/risk of a point estimate. Also, covariance or correlation analysis can inform planners the relationship between outputs and inputs/parameters that may not be apparent just from the model. A systematic uncertainty analysis can provide insights into the level of confidence on the model outputs, and also identify critical sources of uncertainty for enhancing the robustness of the travel demand model.

In the literature, Waller *et al*. (2001) studied the impact of demand uncertainty on the results of the traffic assignment model. They found that the traffic assignment results tend to overestimate the network performance when ignoring demand uncertainty. Bowman *et al*. (2002) proposed a simplified method to estimate the probability distribution of a travel demand forecast. Given a forecast of any variable of interest (e.g., revenue, ridership), this method identified independent sources of uncertainty, estimated a probability distribution of each source, estimated the sensitivity of the variable to each source, and then combined the effects of multiple sources. Zhao and Kockelman (2002) addressed the uncertainty propagation issue of a sequential four-step procedure using Monte Carlo simulation, and concluded that the error of the model tends to amplify in the first three steps (i.e., trip generation, trip distribution and modal split) and reduces in the last step (i.e., traffic

assignment) of the sequential four-step procedure. Moreover, Pradhan and Kockelman (2002) and Krishnamurthy and Kockelman (2003) investigated the uncertainty propagation of an integrated land use-transportation model over time. Along a different line, Ševčíková et al. (2007) developed a Bayesian melding method for assessing uncertainty about variables of interest using urban stochastic simulation models. It combined all the available information about inputs and outputs (in terms of prior probability distributions and likelihoods) in a Bayesian approach to obtain the posterior distribution of variables as a function of inputs and/or outputs. Recently, Rasouli and Timmermans (2012) provided a comprehensive review on the uncertainty analysis in travel demand forecasting, including four-step models, discrete choice models, and activity-based travel demand models.

Typically, uncertainty analysis of a model consists of the following three steps: (1) characterization of input/parameter uncertainty, (2) uncertainty propagation, and (3) characterization of output uncertainty. The first step is to estimate the distribution characteristics (e.g., mean, variance, skewness, etc) of input/parameter uncertainty in the model. For the purpose of depicting their respective effect, we separate the input and parameter uncertainty in the analysis. Input uncertainty is a phenomenon that inherently exists in the real world. It can be measured, analyzed, and where appropriate explained. By contrast, parameter uncertainty is an aspect of knowledge which can (at least theoretically) be reduced by collecting more and better data (Brattin *et al*., 1996; Rai *et al*., 1996; Vose, 2000). However, the distinction between them is to a great extent a matter of convention since it may not be feasible to eliminate error in measurement (reducible uncertainty) beyond a certain level. In this paper, if not mentioned explicitly, the uncertainty refers to the combination of input and parameter uncertainty or the total uncertainty. The second step estimates the output uncertainty resulting from the input/parameter uncertainty. It is concerned with how the input/parameter uncertainty is converted or propagated by the model to the output uncertainty. The third step studies the characteristics of the output uncertainty, such as mean, variance, confidence level for a certain output, relationship between input and output (e.g., whether they are dependent or not, how strongly they are related if dependent) as well as between parameter and output.

The purpose of this study is to develop a systematic and computationally efficient network equilibrium approach for quantitative uncertainty analysis of a combined travel demand

model (CTDM) using the analytical sensitivity-based method. Our approach has the following three key features:

(1) Although the sequential four-step procedure has been widely employed by practitioners, it suffers from inconsistent consideration of travel times and congestion effects amongst various steps since it is not based on a single unifying rationale that would explain or legitimize all dimensions of travel demands jointly (Garret and Wachs, 1996). Further, the aforementioned inconsistency of the sequential approach may also introduce extra error into the uncertainty analysis. In this study, the combined travel-destination-mode-route choice model originally proposed by Oppenheim (1995) is adopted to overcome the problems associated with the sequential four-step procedure. Thus, the combined travel demand model (CTDM) provides a viable avenue for modeling and predicting multi-dimensional travel demands and equilibrium flows on congested networks, while overcoming the inconsistency issues associated with the sequential four-step procedure and behavioral inconsistency between different steps. This feature makes it possible to develop an analytical sensitivity-based uncertainty analysis for assessing the uncertainty associated with a combined travel demand model.

(2) An analytical sensitivity-based method recently developed by Yang and Chen (2009) for the CTDM will be adopted to develop the uncertainty analysis methodology instead of the time-consuming sampling-based methods (e.g., Monte Carlo, Quasi- Monte Carlo, Latin Hypercube sampling, etc.). For the sampling-based methods, it is unclear how many samples are sufficient for conducting the uncertainty analysis of a travel demand model. In addition, the computational burden could be prohibitively expensive for practical applications. Since the CTDM is formulated as a nonlinear programming (NLP) problem and the uniqueness of solution can be guaranteed under the commonly adopted assumptions, the sensitivity analysis method of NLP can be employed and customized to derive the sensitivity expressions of the output variables (e.g., multi-dimensional travel demand, traffic flow, and travel cost) with respect to perturbations from various input variables (e.g., number of zonal potential travelers) and parameters (e.g., attractiveness of travel choices) in the combined model. In addition, the sensitivity analysis of the network-level performance measures such as the total system travel time and total vehicle mile traveled can also be obtained using matrix manipulation and differential chain rule. Based the above derivative information and a given variance-covariance matrix of inputs/parameters, we can estimate the variance-covariance matrix of outputs, the confidence intervals of outputs, and correlation between outputs and inputs/ parameters.

(3) Uncertainties stemming from input data and model parameters can be treated separately so that the individual and collective effects of uncertainty on the outputs can be clearly assessed and quantified. This capability provides valuable information on improving the quality of model estimation by reducing uncertainty. The approach allows analysis of input data and model parameter uncertainties in a systematic way for addressing the travel demand forecasts quality and the implications and significance of the relevant information for model improvement and infrastructure investment.

The remainder of this paper is organized as follows. Travel demand models are discussed in Section 2. Section 3 provides the sensitivity analysis of the combined travel demand model, which is followed by an uncertainty analysis in Section 4. Section 5 presents some numerical examples to demonstrate the sensitivity-based uncertainty analysis of the combined travel demand model. Finally, some concluding remarks are given in Section 6.

2. Travel demand forecasting models

This section presents the travel demand forecasting models. Notation is listed first, which is followed by discussions on the sequential four-step and combined travel demand models, and the CTDM formulation.

2.1 Notation

β Parameters in the combined travel demand model: β_r, β_m, β_d, and β_t are positive parameters associated with the variances of the random components in the route, mode, destination and travel choices, respectively

β' Rescaled parameters β, where $\frac{1}{\beta'_m}=\frac{1}{\beta_m}-\frac{1}{\beta_r}$; $\frac{1}{\beta'_d}=\frac{1}{\beta_d}-\frac{1}{\beta_m}$; $\frac{1}{\beta'_t}=\frac{1}{\beta_t}-\frac{1}{\beta_d}$

Assume that $\beta_r > \beta_m > \beta_d > \beta_t$, so that the rescaled parameters are all positive

$\delta_{ijr}^{a_m}$ Link-route incidence indicator, 1 if link a_m is on route r from origin i to destination j on mode m, 0 otherwise

τ A scalar attached to travel time in the utility function (value of time)

a_m Link index of mode m

c_{a_m} Fixed travel cost on link a of mode m

$t_{a_m}(\cdot)$ Travel time function for link a of mode m

$g_{a_m}(\cdot)$ Generalized link travel cost function for link a of mode m, $g_{a_m}(\cdot)=\tau t_{a_m}(\cdot)+c_{a_m}$

g_{ijmr} Generalized cost of taking route r on mode m between origin i and destination j, $g_{ijmr}=\sum_{a_m} g_{a_m}\delta_{ijr}^{a_m}$

h_x	Constant term in the utility specification, which can be specified as a linear function of socioeconomic characteristics: h_i is the traveling attractiveness of origin i; h_{ij} is the attractiveness of destination j from origin i; h_{ijm} is the attractiveness of mode m between i and j
N_i	Number of potential travelers from origin i
P_{xy}	(Unconditional) joint probability of x and y; for example, P_{ijmr} is the probability that a traveler in origin i travels to destination j on mode m through route r
$P_{y\|x}$	Conditional probability of choosing y given x; for example, $P_{r\|ijm}$ is the probability of taking route r given that a traveler in origin i has chosen to travel to destination j on mode m
T_i	Number of travelers from origin i
T_{i0}	Number of non-travelers in origin i
T_{ij}	Number of travelers from origin i to destination j
T_{ijm}	Number of travelers using mode m from origin i to destination j
T_{ijmr}	Number of travelers taking route r on mode m from origin i to destination j for the travel purpose of interest
U	Direct utility of choices
$\tilde{W}$	Expected received utility of choices

2.2 Sequential four-step and combined travel demand models

The conventional travel demand forecasting model uses a sequential four-step procedure: trip generation, trip distribution, mode choice, and traffic assignment (Ortuzar and Willumsen, 2001). Alternatively, the four-step procedure can be viewed in two stages: 1) various characteristics of the travelers and land use (and, to a certain extent, the transportation system) are calibrated, validated, and evaluated to produce a non-equilibrated measure of travel demand (i.e., mode-specific trip tables), and 2) the trip tables are loaded onto the transportation network in the traffic assignment step to yield equilibrium flows and travel times (McNally, 2000a). In both stages, the results of one step (or stage) act as inputs to next step. However, it should be recognized that the sequential four-step procedure suffers from inconsistent consideration of travel times and congestion effects amongst various steps since it is not based on a single unifying rationale that would explain or legitimize all dimensions of travel demands jointly (Garret and Wachs, 1996; Oppenheim, 1995). The models used in different steps have different rationales along with different assumptions (e.g., Gravity model in the trip distribution step, Logit model in the modal split step, and user equilibrium (UE) model in the traffic assignment step).

Although there are some efforts to remedy this inconsistency (e.g., by introducing a feedback mechanism to ensure the level of service is consistent between the trip distribution and traffic assignment steps), generally there is no guarantee that it will always converge to give stable

and accurate results. Typically, the feedback is implemented by averaging (or weighing) successive four-step solutions. As mentioned by Zhang and Boyce (2000), two issues need to be considered in the feedback:

(a) *Determination of the weight parameter*: Since there is no objective function in the four-step procedure, it is not possible to search for the optimal weight through the one-dimensional optimization. Usually, the method of successive averages (MSA) is adopted despite its known weakness of obtaining highly accurate solution due to its slow convergence characteristics;

(b) *Selection of variables for the feedback control*: The feedback variable could be trip table, travel time, volume, or speed. The consistency can be achieved, but only if the feedback is performed in certain ways (Boyce *et al*., 1994; Boyce, 2002).

With a carefully selected feedback variable and the MSA scheme, theoretically we may be able to guarantee the convergence/consistency. However, it is well known that the MSA scheme suffers from the *sub-linear convergence rate*. The stepsizes are quite small after the first few iterations, slowing down the convergence significantly. Hence, the feedback method with MSA will not be able to obtain stable and accurate solutions within an acceptable computational effort for real networks. The inferior performance of the feedback procedure has also been demonstrated by Bar-Gera and Boyce (2003) in the comparison with the origin-based algorithm and Evans algorithm for solving the combined distribution and assignment model. In addition, Zhang and Boyce (2000) concluded that progresses in improving travel forecast may be neither adequate nor efficient with "feedback". For other weaknesses and limitations of the conventional four-step procedure, interested readers may refer to McNally (2000b) and Boyce (2002).

Motivated by the problems associated with the sequential four-step procedure, researchers proposed alternative models for travel demand forecasting. Such models are referred to as combined or integrated models. Evans (1976) formulated a combined model, which integrates trip distribution and traffic assignment, as a constrained convex optimization problem. Florian and Nguyen (1978) extended the model to include modal split using the fact that an entropy distribution model implies a logit modal-split model. Florian *et al*. (2002) provided a variational inequality formulation for a multi-class multi-mode travel demand model with hierarchical logit structure. Boyce *et al*. (1983) proposed a unified approach by using an entropy function as a general measure of the dispersion of choices across routes, modes, destinations, or locations. A family of combined models of location, destination, mode, and

route choices could be derived following this approach (see Boyce and Daskin (1997) for a travel forecasting model that integrates travelers' route choice, mode choice and destination choice; Boyce and Bar-Gera (2001) for a nonlinear programming formulation that combines origin, destination, mode, and route choices without mode interactions; and Bar-Gera and Boyce (2003) for a fixed point formulation). Safwat and Magnanti (1988) proposed a simultaneous transportation equilibrium model (STEM), which can simultaneously predict the trip generation, distribution, modal split and trip assignment. The STEM can be formulated as an equivalent convex optimization problem, and it is behaviorally richer than other models. However, travelers' route choice behavior (traffic assignment) was characterized by the UE model (i.e., a deterministic network equilibrium model), whereas the destination choice behavior (trip distribution) was governed by the logit model (i.e., a random utility model). There exist behavioral inconsistencies between these two travel choice dimensions. Lam and Huang (1992) formulated a multi-class model by combining trip distribution and traffic assignment. Oppenheim (1995) proposed a combined travel demand model (CTDM), which combines the travel-destination-mode-route choice based on the random utility theory. The approaches to estimate the model parameters were also provided, so that it is possible for the model to be used in practice. Zhou *et al*. (2009) provided alternative formulations, including mathematical programming (*MP*) formulation and variational inequality (*VI*) formulations, for the CTDM that integrates trip generation, trip distribution, modal split, and traffic assignment using the random utility theory framework. Several algorithms have also been proposed in the literature for solving various combined travel demand model formulations (e.g., the partial linearization algorithm by Evans (1976) for solving the combined distribution and assignment (CDA) problem as a constrained convex optimization problem; the origin-based algorithm by Bar-Gera and Boyce (2003) for solving the CDA problem; the improved origin-based algorithm by Xu *et al*. (2008) by adopting the modified origin-destination flow update strategy proposed by Huang and Lam (1992) to enhance its computational efficiency; and the Block Gauss-Seidel decomposition approach coupled with the method of successive averages by Florian *et al* (2002) for solving the variational inequality formulation).

In this paper, we employ the CTDM originally proposed by Oppenheim (1995) as the combined travel demand model for the sensitivity-based uncertainty analysis. The CTDM defines a traveler as a consumer of urban trip and reflects his/her budget constraint choices. The solution of the model is proved to correspond to individual as well as collective utility maximization. With its sound behavioral rationale, the CTDM is a viable avenue with

behavioral consistency for modeling and predicting multi-dimensional travel demands and equilibrium flows in transportation networks. It overcomes not only the inconsistency of the sequential four-step procedure but also the behavioral inconsistency between different models/approaches used in other combined travel demand model (i.e., deterministic and stochastic models). More importantly, the logit-based probability expression for all travel choice dimensions ensures that the solution to CTDM is unique. This nice feature makes it possible to develop an analytical sensitivity-based uncertainty analysis approach. To sum up, the motivations of adopting the CTDM in the uncertainty analysis include: (1) a single unifying rationale (i.e., utility maximization theory), and (2) elegant closed-form probability expressions for all travel choice dimensions and a unique solution to the CTDM formulation. However, the proposed framework of uncertainty analysis could also be applied to other combined demand models with an equivalent convex programming formulation.

2.3 Combined travel demand model

Following Oppenheim (1995), each traveler's decision process is assumed to have the following top-down structure:

- Given an individual at location i, a given time period (hour, day, etc.), and an activity (e.g., shopping, work, recreation, etc.), a potential traveler first decides whether to travel or not. $P_{t|i}$ is the probability that a potential traveler makes a trip in the study time period.
- Given the choice made at the first level, the conditional probability that an individual will choose destination j to conduct the activity is $P_{j|i}$.
- Given the outcomes from the first two decisions, the conditional probability that an individual will choose mode m (for traveling from i to j) to conduct the activity is $P_{m|ij}$.
- Given the outcomes from the preceding decisions, the conditional probability that an individual will choose route r (for traveling from i to j on mode m) to conduct the activity is $P_{r|ijm}$.

The above hierarchical structure can be represented as in Figure 1. Following the random utility theory, the probabilities at each stage are calculated by the multinomial logit choice function. This "nested" structure is the basis for constructing the combined travel-destination-mode-route choice model as a mathematical programming given below. For example, at the first stage (i.e., decision to travel) the number of trips from origin i (T_i) is calculated by

multiplying the conditional probability ($P_{t|i}$) of making a trip with the number of potential travelers from origin i (N_i). The conditional probability ($P_{t|i}$) is calculated as follows:

$$P_{t/i} = \frac{e^{\beta_t (h_i + \tilde{W}_{t/i})}}{1 + e^{\beta_t (h_i + \tilde{W}_{t/i})}}, \forall i, \tag{1}$$

where β_t is the parameter associated with travel choice; h_i is the attractiveness of making a trip at origin i; and $\tilde{W}_{t/i}$ is the expected received utility of traveling from origin i, which depends on the utility at the next stage (i.e., destination choice):

$$\tilde{W}_{t/i} = b_i + \frac{1}{\beta_d} \ln \sum_j e^{\beta_d (h_{ij} + \tilde{W}_{j/i})}, \forall i, \tag{2}$$

where b_i is the budget (i.e., time and money) of an individual spent on traveling from origin i; β_d is the parameter associated with destination choice; h_{ij} is the attractiveness of traveling from origin i to destination j. The expected received utility at the destination choice stage also depends on the utility at the mode choice stage, and so on until it reaches the last stage (i.e., route choice).

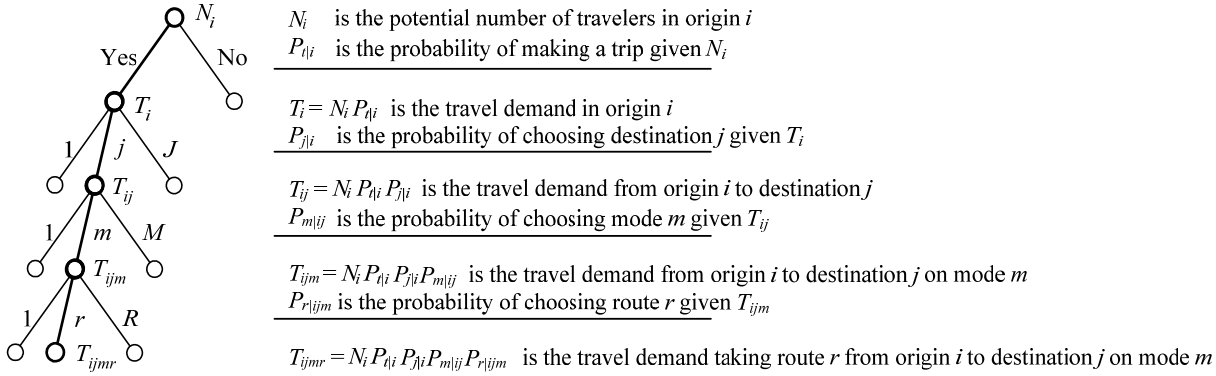


Figure 1 Hierarchical structure of a combined travel demand model

Hence, the probability that an individual takes route r on mode m from origin i to destination j can be obtained by multiplying the conditional probability at each stage in a "*nested*" structure starting from the route choice stage until the decision to travel stage as follows:

$$\begin{aligned} P_{ijmr} &= P_{t|i} \cdot P_{j|i} \cdot P_{m|ij} \cdot P_{r|ijm} \\ &= \frac{e^{\beta_t (h_i + \tilde{W}_{t|i})}}{1 + e^{\beta_t (h_i + \tilde{W}_{t|i})}} \frac{e^{\beta_d (h_{ij} + \tilde{W}_{j|i})}}{\sum_j e^{\beta_d (h_{ij} + \tilde{W}_{j|i})}} \frac{e^{\beta_m (h_{ijm} + \tilde{W}_{m|ij})}}{\sum_m e^{\beta_m (h_{ijm} + \tilde{W}_{m|ij})}} \frac{e^{-\beta_r g_{ijmr}}}{\sum_r e^{-\beta_r g_{ijmr}}}, \quad \forall i, j, m, r. \end{aligned} \tag{3}$$

It is worth noting that for keeping the notation system compact, we use the same index between the denominator and numerator for each probability expression. The mathematical programming formulation for the combined travel-destination-mode-route choice model can be constructed as follows (Oppenheim, 1995):

$$\begin{aligned}\min U_{TDMR}(T_i, T_{i0}, T_{ij}, T_{ijm}, T_{ijmr}) = & \sum_m \sum_{a_m} \int_0^{\sum_{ijr} T_{ijmr} \delta_{ijr}^{a_m}} g_{a_m}(\omega) d\omega - \sum_{ijm} h_{ijm} T_{ijm} - \sum_{ij} h_{ij} T_{ij} - \sum_i h_i T_i \\ & + \frac{1}{\beta_r} \sum_{ijmr} T_{ijmr} \ln T_{ijmr} + \frac{1}{\beta'_m} \sum_{ijm} T_{ijm} \ln T_{ijm} + \frac{1}{\beta'_d} \sum_{ij} T_{ij} \ln T_{ij} \\ & + \frac{1}{\beta'_t} \sum_i T_i \ln T_i + \frac{1}{\beta_t} \sum_i T_{i0} \ln T_{i0}\end{aligned} \tag{4}$$

subject to:

$$\sum_r T_{ijmr} = T_{ijm}, \ \forall i, j, m, \tag{5}$$

$$\sum_m T_{ijm} = T_{ij}, \ \forall i, j, \tag{6}$$

$$\sum_j T_{ij} = T_i, \ \forall i, \tag{7}$$

$$T_i + T_{i0} = N_i, \ \forall i, \tag{8}$$

$$T_{i0} > 0, \ T_i > 0, \ T_{ij} > 0, \ T_{ijm} > 0, \ T_{ijmr} > 0, \ \forall i, j, m, r, \tag{9}$$

where U_{TDMR} is a modified negative direct utility of a representative traveler (R.T.), which is defined to represent the utility maximizing choices at the aggregated demand level. Conceptually, this objective function integrates the utility maximization of individual travelers and the congestion effect (i.e., multiple travelers) by combining multi-dimensional discrete choice concept into the network equilibrium problem. Structurally, this objective function is constructed using a similar way as Fisk (1980)'s model for the logit-based stochastic user equilibrium problem. The first four terms are related to the direct utility of route choice, mode choice, destination choice and travel choice, respectively. The remaining five terms are the "entropy" terms associated with the logit-based probabilistic choices (i.e., route choice, mode choice, destination choice, travel choice, and no travel choice). Constraints (5)–(8) are the conservation constraints. Constraint (9) ensures the solutions are positive.

The Karush-Kuhn-Tucker (KKT) conditions for optimality with respect to T_{ijmr}, T_{ijm}, T_{ij}, and T_i indeed give the probabilities of choosing route ($P_{r|ijm}$), mode ($P_{m|ij}$), destination ($P_{j|i}$), and

decision to travel ($P_{t|i}$), respectively, as given in Eq. (3). Hence, the above mathematical programming formulation indeed represents the structure of traveler decision process as illustrated in Figure 1. In addition, the above formulation is a strictly convex programming, guaranteeing the solution uniqueness. For the detailed proof on the existence, equivalency, and uniqueness of the solution to the above mathematical programming formulation, interested readers may refer to Oppenheim (1995) and Yang *et al.* (2013).

2.4 Solution algorithm for solving CTDM

In this section, we provide an algorithm for solving the CTDM given in Eqs (4) – (9). The algorithm is based on the partial linearization method (Evans, 1976), which is a descent algorithm for continuous optimization problems (Patriksson, 1994). A search direction is obtained from the solution of a convex auxiliary problem, which is defined by an approximation of the objective function in Eq. (4) through a first-order approximation of the first term (i.e., link cost integral term). A (possibly inexact) line search is made in the direction obtained with respect to the objective function, and the resulting step size defines a new solution with a reduced objective value. The search direction and line search steps are iteratively performed until some convergence criterion is reached. It is important to recognize that the demand of each stage in the CTDM is solved simultaneously by using the partial linearization algorithm. The traffic forecasts resulting from this multi-dimensional choice problem are consistent. That is, T_i, T_{i0}, T_{ij}, T_{ijm}, and T_{ijmr} are consistent with the traveler's expected received utility. By contrast, in the sequential four-step procedure, consistency is not always guaranteed even with a "*feedback*" mechanism.

The key steps of the partial linearization algorithm for solving the CTDM are as follows.

Step 0 *Initialization*: Set an initial solution **T**(0) with $T_i^n, T_{i0}^n, T_{ij}^n, T_{ijm}^n, T_{ijmr}^n = 0,\ \forall i, j, m, r$;

$v_{a_m}^n = 0,\ \forall\ m, a_m$, and let n=0.

Step 1 Update link travel time $t_{a_m}^n = t_{a_m}\left(v_{a_m}^n\right), \forall\ m, a_m$, and link generalized cost

$g_{a_m}^n = c_{a_m} + \tau t_{a_m}^n, \forall\ m, a_m$.

Step 2 *Search Direction*: Find the minimum cost route based on the generalized link cost;

update the route set; calculate the route cost $g_{ijmr}^n = \sum_{a_m} g_{a_m}^n \left(\delta_{ijr}^{a_m}\right)^n, \forall i, j, m, r$; and then

---

solve a partially linearized subproblem to obtain an auxiliary solution $\mathbf{H}(n)=\left\{H_i^n, H_{i0}^n, H_{ij}^n, H_{ijm}^n, H_{ijmr}^n\right\}$. The resultant search direction is $\mathbf{H}(n)$ - $\mathbf{T}(n)$.

Step 3 *Line Search*: Solve $\alpha(n)=\arg\min_{0\le\alpha\le 1} Z\left[\mathbf{T}(n)+\alpha\left(\mathbf{H}(n)-\mathbf{T}(n)\right)\right]$.

Step 4 *Update*: Let $\mathbf{T}(n+1)=\mathbf{T}(n)+\alpha(n)\cdot\left[\mathbf{H}(n)-\mathbf{T}(n)\right]$.

Step 5 *Termination Criterion*: If some termination criterion is satisfied, then terminate; otherwise, set n:=n+1 and go to Step 1.

The main computational efforts in this algorithm are in Step 2 and Step 3. In Step 2, a column generation method is used to update the route sets for each origin-destination (O-D) pair on each mode after finding the shortest path. The expected received utilities and conditional probabilities are calculated from the bottom (the route choice stage) to the top (the travel choice stage) of the hierarchical structure of the CTDM while the auxiliary demand of each stage is calculated from the top to the bottom. The main difference of the partial linearization algorithm with the Frank-Wolfe algorithm (LeBlanc *et al*., 1975) is that the auxiliary demand is determined by the logit choice function instead of an all-or-nothing loading mechanism to determine the descent direction. For the line search in Step 3, the objective function is very complex. To find the optimal (exact) step size, objective function evaluations or its derivatives need to be computed multiple times via golden section method or bisection method. This is a time-consuming step due to the high dimension of solution variables. Also, the line search is conducted on an approximate descent direction. The exact line search on the approximate descent direction may not be so efficient. Instead, inexact line search methods are recommended. Detailed implementation steps are provided in Yang *et al*. (2013).

3. Sensitivity analysis

Sensitivity analysis is an effective way to quantitatively analyze the behavior of demand and flow pattern with respect to the perturbations of network characteristics. Dafermos and Nagurney (1984) performed sensitivity analysis for the asymmetric network equilibrium problem to predict the change of traffic pattern upon the change in the traffic demand and link cost function. Tobin and Friesz (1988) proposed a sensitivity analysis method for the fixed-demand network equilibrium problem. Since the path flow solution of the user equilibrium problem is not unique, the standard sensitivity analysis method for either nonlinear

programming problem (Fiacco, 1983) or variational inequality problem (Tobin, 1986) cannot be used directly to derive the sensitivity information. A restricted problem, which is equivalent to the traffic equilibrium problem, was developed. The restricted problem has the uniqueness properties, so that the existing sensitivity analysis method can be applied. Yang (1997) extended Tobin and Friesz's (1988) approach for the elastic-demand network equilibrium problem. Qiu and Magnanti (1989) proposed a general approach for the sensitivity analysis for variational inequalities and demonstrated the approach for traffic assignment problems. Yang and Bell (2007) revisited the gradient-based sensitivity analysis approach of user equilibrium and developed a simple but general approach for calculating the derivatives of equilibrium link flows. The sensitivity analysis of stochastic user equilibrium (SUE) was also derived for both logit-based SUE (Ying and Miyagi, 2001) and probit-based SUE (Clark and Watling, 2000; 2002). Yang *et al*. (2001) derived the explicit derivative expressions of the logit-based SUE and applied them to solve the simultaneous estimation of O-D matrix and travel-cost coefficients. Davis (1994) developed exact local solution based on the derivatives of the logit-based SUE problem for solving the continuous network design problem. Liu *et al*. (2010) developed a dimension-reduced sensitivity analysis method for the SUE problem with smooth or non-smooth cost functions. Boyles (2012) developed bush-based sensitivity analysis for subnetwork diversion, while Du *et al*. (2012) derived the sensitivity expressions for the equilibrium trip distribution-assignment model with variable destination costs. Recently, Yang and Chen (2009) studied the sensitivity analysis for the CTDM using the standard sensitivity analysis method for nonlinear programming problem due to the uniqueness of solution.

In this section, we present the key results of sensitivity analysis for the CTDM developed by Yang and Chen (2009) for the purpose of uncertainty analysis. Let ε denote the vector of perturbations associated with the attractiveness of choices, link cost function, and number of potential travelers. After formulating the Lagrangian function L for the CTDM (4)-(9), we can derive its Hessian matrix as well as the gradient vector. Under some commonly-used assumptions for guaranteeing the perturbed problem is regular or $M(\varepsilon)$ is non-singular (see Fiacco (1983), Yang and Chen (2009) on how the CTDM satisfies these conditions), we have:

$$\nabla_{\varepsilon} y(\varepsilon) = -M(\varepsilon)^{-1} N(\varepsilon), \tag{10}$$

where $y = (T_i, T_{i0}, T_{ij}, T_{ijm}, T_{ijmr}, \pi_{ijm}, \mu_{ij}, \lambda_i, \varphi_i)$; $\pi_{ijm}, \mu_{ij}, \lambda_i, \varphi_i$ are the Lagrange multipliers associated with constraints (5)–(8);

$$M(\varepsilon)=\begin{bmatrix}
\nabla^2_{T_i}L & & & & 0 & 0 & 0 & -I & I \\
& \nabla^2_{T_{i0}}L & & & & 0 & 0 & 0 & I \\
& & \nabla^2_{T_{ij}}L & & & 0 & -I & \Gamma^T & 0 \\
& & & \nabla^2_{T_{ijm}}L & & -I & \Phi^T & 0 & 0 \\
0 & & & & \nabla^2_{T_{ijmr}}L & \Lambda^T & 0 & 0 & 0 \\
0 & 0 & 0 & -I & \Lambda & 0 & \cdots & & 0 \\
0 & 0 & -I & \Phi & 0 & \vdots & & & \\
-I & 0 & \Gamma & 0 & 0 & & & & \vdots \\
I & I & 0 & 0 & 0 & 0 & & \cdots & 0
\end{bmatrix}, \quad (11)$$

$$N(\varepsilon)=\left[\nabla_{T_i,\varepsilon}L \quad \nabla_{T_{i0},\varepsilon}L \quad \nabla_{T_{ij},\varepsilon}L \quad \nabla_{T_{ijm},\varepsilon}L \quad \nabla_{T_{ijmr},\varepsilon}L \quad 0 \quad 0 \quad 0 \quad -\nabla_\varepsilon N_i\right]^T, \quad (12)$$

where Λ is the path/O-D mode incidence matrix; Φ is the mode/O-D incidence matrix; Γ is the O-D/origin incidence matrix. Note that the first entropy term in Eq. (4) makes $\nabla^2_{T_{ijmr}}L$ positive definite (the sum of a positive semi-definite matrix and a positive definite matrix is a positive definite matrix). This is different from the UE case, where the Hessian matrix with respect to route flow is positive semi-definite. In addition, $\nabla^2_{T_i}L$, $\nabla^2_{T_{i0}}L$, $\nabla^2_{T_{ij}}L$, and $\nabla^2_{T_{ijm}}L$ are all diagonal and positive definite matrices. Hence, the whole Hessian matrix of L is positive definite. Also, the large matrix at the top right corner (above the zero matrices) contains linearly independent columns. Then, it is easy to prove that **M** is non-singular or invertible. For demonstration purposes, an invertibility proof is provided in the appendix for a simple case with one O-D pair (no destination choice) and one single car mode (no mode choice). The CTDM under this case then reduces to the Logit-based SUE model. This proof follows the approach of Yang and Bell (2007) and it is different from the one provided in Yang and Chen (2009).

For a given perturbation ε, we may estimate $y(\varepsilon)$ using the first-order Taylor series approximation:

$$y(\varepsilon)\approx y(0)+\varepsilon\left[\nabla_\varepsilon y(\varepsilon)\right]^T. \quad (13)$$

From Eq. (10), we can obtain the derivatives of decision variables $y=(T_i,T_{i0},T_{ij},T_{ijm},T_{ijmr},\pi_{ijm},\mu_{ij},\lambda_i,\varphi_i)$ with respect to a given perturbation ε. For other outputs, such as link flows, total travel time, total vehicle miles traveled, etc., the derivatives can be calculated using matrix manipulation and differential chain rule.

$$\frac{\partial v_{a_m}}{\partial \varepsilon_i} = \sum_{ijr} \frac{\partial T_{ijmr}}{\partial \varepsilon_i} \delta_{ijr}^{a_m} , \tag{14}$$

$$\frac{\partial t_{a_m}}{\partial \varepsilon_i} = \left(t_{a_m}\right)_{v_{a_m}} \frac{\partial v_{a_m}}{\partial \varepsilon_i} + \left(t_{a_m}\right)_{\varepsilon_i} , \tag{15}$$

where $\partial T_{ijmr} / \partial \varepsilon_i$ can be obtained from Eq. (10); $\left(t_{a_m}\right)_{v_{a_m}}$ is the derivative of explicit link cost function with respect to link flow, which can be calculated given the link cost function; $\left(t_{a_m}\right)_{\varepsilon_i}$ is the derivative of explicit link cost function with respect to perturbation, which can be either an input or a parameter. $\left(t_{a_m}\right)_{\varepsilon_i}$ is an "apparent" derivative, while $\partial t_{a_m} / \partial \varepsilon_i$ is the "true" partial derivative (Ying and Miyagi, 2001). The derivatives of network performance indices, such as total travel time (*TTT*) and total vehicle mileage (*TVM*), w.r.t. perturbations can also be calculated as:

$$\frac{\partial TTT}{\partial \varepsilon_i} = \sum_m \sum_{a_m} \left(\frac{\partial v_{a_m}}{\partial \varepsilon_i} t_{a_m} + \frac{\partial t_{a_m}}{\partial \varepsilon_i} v_{a_m} \right), \tag{16}$$

$$\frac{\partial TVM}{\partial \varepsilon_i} = \sum_m \sum_{a_m} h_{a_m} \left(\frac{\partial v_{a_m}}{\partial \varepsilon_i} \right), \tag{17}$$

where h_{a_m} is the length of link *a* on mode *m*. These derivatives provide useful information for different applications (Yang and Chen, 2009).

Note that the above sensitivity analysis method is conducted in the route domain based on a pre-generated or pre-enumerated route set. It has both disadvantage and advantage. The disadvantage is that it requires enumeration of a set of routes and it can pose a problem for large networks. Using Dial's link-based stochastic loading method without route enumeration would be an advantage in conducting the sensitivity analysis. However, route enumeration has its own appeal as more behaviorally realistic routes could be generated according to recent advances in route generation strategies (Bekhor *et al*., 2006; Bovy *et al*., 2007; Prato and Bekhor, 2006, 2007). These pre-generated or pre-enumerated routes explicitly account for the route overlapping problem, which is not considered in Dial's stochastic loading method (Bekhor *et al*., 2008; Cascetta *et al*., 1996, 1997). Hence, the link-based stochastic loading method may give biased results if the routes are not truly independent.

To reduce the computational burden, we may substitute the route-based sensitivity analysis by the link-based sensitivity analysis for the general unconstrained formulation of CTDM (Zhou *et al*., 2009). The work by Ying and Miyagi (2001) for the logit-based SUE and Ying and Yang (2005) for the bi-modal logit-based SUE can be extended to the sensitivity analysis for the general unconstrained formulation of CTDM. However, this substitution does not change the uncertainty analysis framework to be presented in Section 4.

4. Uncertainty analysis

In this section, we present an analytical uncertainty analysis method based on the sensitivity analysis of the CTDM. As mentioned in the introduction, there are three steps in the uncertainty analysis. In this paper, we focus on the last two steps while assuming the characteristics of model inputs/parameters are given. The second step, i.e., uncertainty propagation, is to estimate the probability distributions of outputs given that of model inputs/parameters. The third step, i.e., characterization of output uncertainty, is to investigate the confidence level of outputs, the relationship between the outputs and inputs as well as the relationship between the outputs and parameters.

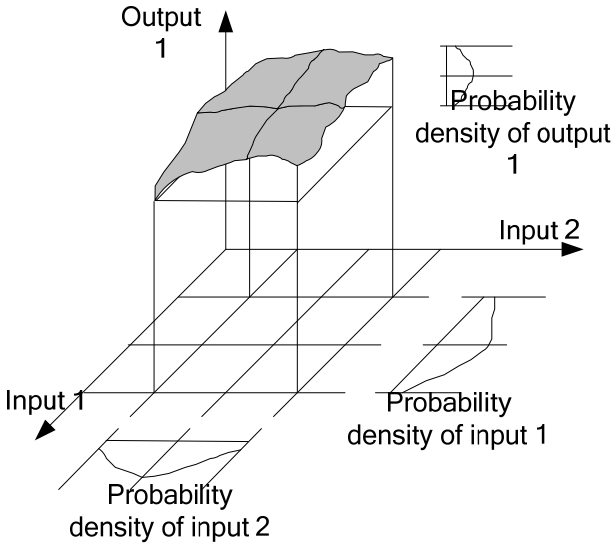


Figure 2 Propagation of uncertainties

Figure 2 graphically depicts the concept of uncertainty propagation. Each point of output 1 is characterized by a probability density function (PDF), which depends on the PDFs of input 1 and input 2. Two possible approaches to investigate the uncertainty propagation are the

sampling based methods and the analytical sensitivity-based method. Sampling based methods (e.g., Monte Carlo sampling, Latin Hypercube sampling, etc.) are a flexible approach that can be used to simulate any systems. However, these methods require a relatively high computational effort and are limited to investigate a small number of perturbations. Also, the simulation adds an element of non-reproducibility (Bell *et al*., 1999). On the other hand, studies have addressed the importance of sensitivity analysis in identifying the possible errors in the travel demand model (Robbins, 1978; Bonsall *et al*, 1977). Leurent (1998) conducted the sensitivity and error analysis for the dual criteria traffic assignment model. The sensitivity analysis is more effective and can investigate the uncertainty stemming from inputs and parameters simultaneously or separately. However, since sensitivity analysis is locally valid for minor perturbations of inputs and parameters, its application is limited to the problems that satisfy the necessary conditions to conduct the sensitivity analysis.

In Zhao and Kockelman (2002), the Monte Carlo sampling method was employed to perform the uncertainty analysis in the sequential four-step procedure. The aforementioned inconsistency of the sequential four-step procedure introduces extra uncertainty of the model. Since there is no unified mathematical formulation for the four-step procedure, the sensitivity analysis method cannot be used. In the numerical example, the authors identified 118 inputs and parameters for the case study. However, due to the heavy computational burden of repeatedly solving the sequential four-step procedure, only 100 samples were generated in the Monte Carlo simulation for the uncertainty analysis. It is not clear whether the number of samples is sufficient for conducting the uncertainty analysis of the sequential four-step procedure. In addition, to investigate the relationship between inputs and outputs, they conducted a linear regression analysis to obtain the correlation between inputs and outputs. However, it is difficult to separate the impact of inputs and the model inconsistencies on the outputs. Pradhan and Kockelman (2002) and Krishnamurthy and Kockelman (2003) also used the same sampling-based approach to examine the uncertainty propagation of an integrated land use-transportation model. To investigate the sensitivity of outputs and inputs, they standardized the coefficients obtained from the regression as follows:

$$\beta_{o,i}^{std} = \beta_{o,i} \cdot \frac{\sigma_o}{\sigma_i}, \tag{18}$$

where $\beta_{o,i}^{std}$ is the standardized coefficient between output o and input i; $\beta_{o,i}$ is the coefficient between output o and input i by linear regression; σ_o is the standard deviation of output o

obtained from simulation; σ_i is the standard deviation of input *i*. The standardized coefficient represents the change in the output variable caused by a change in the input variable. The accuracy of estimating the standardized coefficient depends on the accuracy of the sampling method (σ_i) and the linear regression ($\beta_{o,i}$).

In this paper, we use the analytical sensitivity-based uncertainty analysis for investigating the uncertainty propagation of the combined travel demand model proposed by Oppenheim (1995). Using the derivative information and Eq. (13), we can estimate the variance-covariance matrix of outputs given the inputs variance as follows:

$$S_{output} = \nabla_{\varepsilon} y \cdot S_{input} \cdot \left(\nabla_{\varepsilon} y\right)^{T}, \tag{19}$$

where S_{input} is the given variance-covariance matrix of inputs. By assuming the normality of outputs, we can also estimate the confidence intervals of outputs. Furthermore, the covariance of outputs and inputs is

$$S_{output,input} = \nabla_{\varepsilon} y \cdot S_{input}. \tag{20}$$

Let s_i be the *i*-th diagonal element of S_{output}, which is the variance of output *i*; s_j be the *j*-th diagonal element of S_{input}, which is the variance of input *j*; s_{ij} be the element of $S_{output,input}$ at row *i* and column *j,* which is the covariance of output *i* and input *j*. The correlation of output *i* and input *j* can be obtained as follows:

$$r_{ij} = \frac{s_{ij}}{\sqrt{s_i}\sqrt{s_j}}. \tag{21}$$

The correlation of outputs and inputs provides useful information for the critical input analysis. Different from the sampling-based method, the correlation is obtained analytically and directly from the CTDM without the need to conduct a post analysis using linear regression equations. The same analysis can also be conducted for model parameters (separate from input uncertainty).

Remark 1: In Section 2, we assume transit and car networks are independent, and all links in each network have no interaction (i.e., separable link travel time functions) in the CTDM. This assumption is used to construct an equivalent convex programming formulation of the CTDM, which is more computationally and theoretically tractable compared to other advanced demand modeling tools. However, this is not a limitation of the uncertainty

analysis framework proposed in this paper. The framework is still applicable for travel demand models using non-separable link travel time functions with asymmetric interactions. The CTDM with link or mode interactions can be formulated as a variational inequality (VI) problem or a fixed point problem. Accordingly, the sensitivity analysis method for VI problem could be adopted and then embedded into the uncertainty analysis framework. Therefore, the proposed framework of uncertainty analysis is general, which is not limited to separable link travel time functions.

Remark 2: The sampling-based methods (e.g., Monte Carlo simulation) are indeed capable of providing the full probability distribution of output uncertainty, rather than the variance and coefficient of variation only. However, it is difficult to determine a suitable sample size for the simulation method to yield stable solution. In addition, its computational burden could be prohibitively expensive for practical applications. From this viewpoint, we may consider the relationship between the sampling-based methods and the analytical method as a tradeoff between information richness and computational burden.

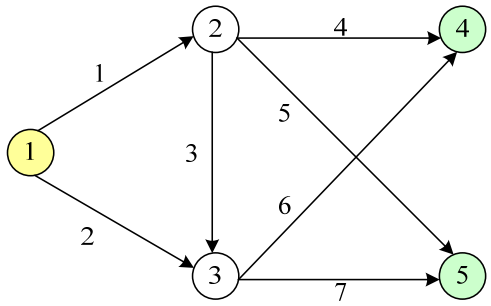
5. Numerical examples

In this section, we present numerical examples to demonstrate the features of the sensitivity-based uncertainty analysis method for the CTDM. The example network shown in Figure 3 is used to illustrate the equilibrium solution, sensitivity analysis, and uncertainty analysis of the CTDM in detail. This network consists of five nodes, seven links and two O-D pairs (from 1 to 4 and from 1 to 5). There are two modes in the network, i.e., car and transit (subscription 'c' for car and 't' for transit). The transit network has the same topology as the car network while they are assumed to be independent. The number of potential travelers (N_1) and attractiveness of zone 1 (h_1) are 200 and 5.0, respectively. Other attractiveness parameters are as follows: h_{14}=3.5, h_{15}=3.8, h_{14c}=3.5, h_{14t}=3.6, h_{15c}=3.8, h_{15t}=3.4. Parameters associated with the route, mode, destination and travel choices, i.e., β_r, β_m, β_d, and β_t, are set as 2.0, 1.0, 0.5, and 0.2, respectively. The fixed monetary cost of traveling on link a of mode m (c_{a_m}) is set at zero for all links in both networks, and the value of time (τ is set at 1.0. Thus, the general link cost equals the link travel time. The link travel time function for both car and transit networks are as follows:

$$t_{a_c}(v_{a_c}) = t_{a_c}^0 \left[1 + \alpha_c \left(v_{a_c} / C_{a_c} \right)^{\gamma_c} \right], \tag{22}$$

$$t_{a_t}(v_{a_t}) = t_{a_t}^0 + \alpha_t \left(v_{a_t} / C_{a_t} \right)^{\gamma_t}, \tag{23}$$

where v_{a_c} and v_{a_t} are link flows for car and transit; $t_{a_c}^0$ and $t_{a_t}^0$ are free-flow travel times for car and transit; C_{a_c} and C_{a_t} are link capacities for car and transit. Here, α_c and γ_c are set at 0.15 and 4.0, and α_t and γ_t are set at 0.06 and 2.0. The free-flow travel time and capacity of each link are given in Table 1. For simplicity, we set the values of link lengths equal to the values of free-flow link travel times.



O-D pair	Route	Link sequences
(1, 4)	1	1-4
	2	1-3-6
	3	2-6
(1, 5)	4	1-5
	5	1-3-7
	6	2-7

Figure 3 Example network

Table 1 Link performance parameters

Link	$t_{a_c}^0$	C_{a_c}	$t_{a_t}^0$	C_{a_t}
1	4.0	25.0	4.0	25.0
2	5.2	25.0	5.2	25.0
3	1.0	15.0	1.0	15.0
4	5.0	15.0	5.0	15.0
5	5.0	15.0	5.0	15.0
6	4.0	15.0	4.0	15.0
7	4.0	15.0	4.0	15.0

5.1 Equilibrium solution

Figure 4 shows the choice probability, demand, and expected received utility (i.e., the log-sum term) at different choice levels. As expected, the equilibrium solution pattern satisfies the logit-type choice probability in Eq. (3). For example,

$$P_{1/14c} = \frac{\exp(-2 \cdot 10.15)}{\exp(-2 \cdot 10.15) + \exp(-2 \cdot 10.53) + \exp(-2 \cdot 10.06)} = 0.38$$

$$\tilde{W}_{c|14} = \frac{1}{2} \ln\left[\exp(-2 \cdot 10.15) + \exp(-2 \cdot 10.53) + \exp(-2 \cdot 10.06)\right] = -9.66$$

$$P_{c/14} = \frac{\exp\left[1 \cdot (3.5 - 9.66)\right]}{\exp\left[1 \cdot (3.5 - 9.66)\right] + \exp\left[1 \cdot (3.6 - 9.01)\right]} = 0.32$$

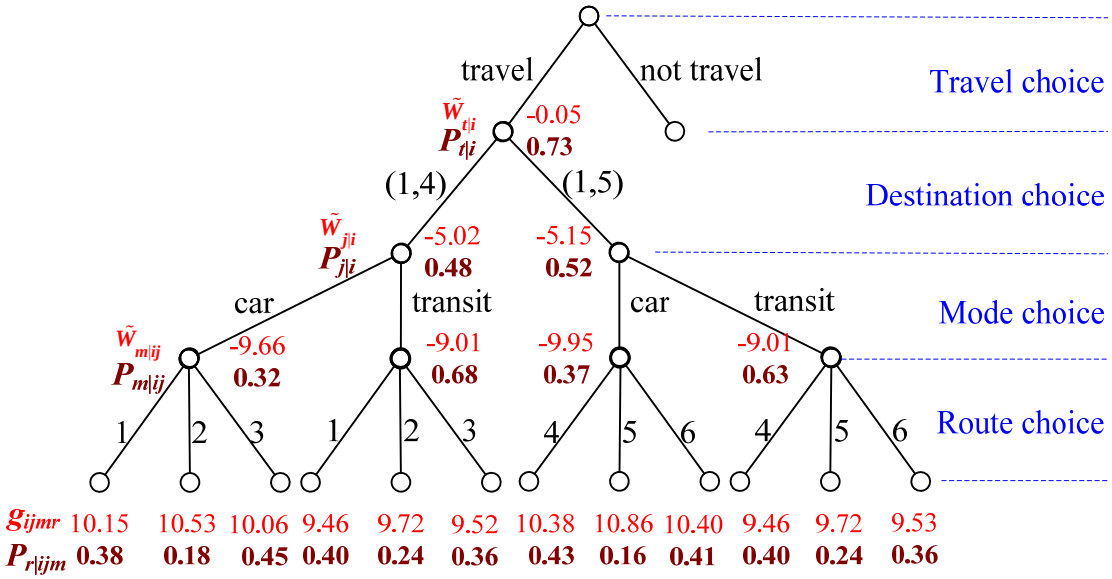
$$\tilde{W}_{4|1} = \frac{1}{1} \ln\left(\exp\left[1 \cdot (3.5 - 9.66)\right] + \exp\left[1 \cdot (3.6 - 9.01)\right]\right) = -5.02$$

$$P_{4/1} = \frac{\exp\left[0.5 \cdot (3.5 - 5.02)\right]}{\exp\left[0.5 \cdot (3.5 - 5.02)\right] + \exp\left[0.5 \cdot (3.8 - 5.15)\right]} = 0.48$$

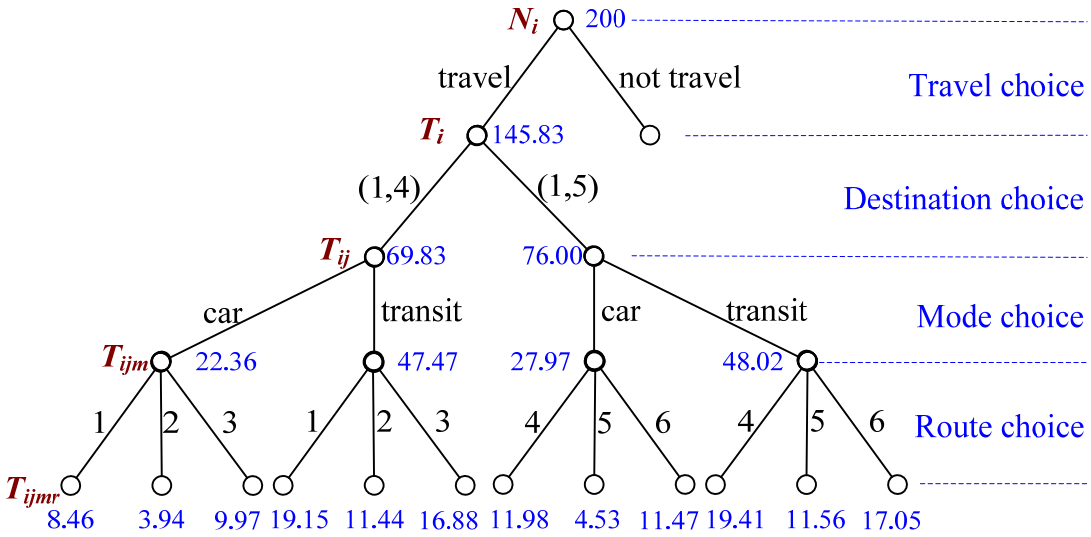
$$\tilde{W}_{t|1} = \frac{1}{0.5} \ln\left(\exp\left[0.5 \cdot (3.5 - 5.02)\right] + \exp\left[0.5 \cdot (3.8 - 5.15)\right]\right) = -0.05$$

$$P_{t/1} = \frac{\exp\left[0.2 \cdot (5 - 0.05)\right]}{\exp\left[0.2 \cdot (5 - 0.05)\right] + 1} = 0.73$$

In addition, all demand variables are positive and the conservation constraints are satisfied at all levels as shown in Figure 4(b).



(a) Choice probability and expected received utility



(b) Equilibrium demand

Figure 4 Multi-dimensional equilibrium demand and choice probability pattern

We should point out that the above multi-dimensional demand patterns are consistent along with the tree structure. In other words, T_i, T_{i0}, T_{ij}, T_{ijm}, and T_{ijmr} are consistent with the traveler's expected received utility at the corresponding choice stage. In contrast, consistency may not always be guaranteed even with a "feedback" mechanism in the sequential four-step forecasting procedure.

5.2 Sensitivity analysis results

For demonstration purposes, we concentrate the analysis on twelve outputs, i.e., x=[T_1, T_{10}, T_{14}, T_{14c}, T_{14t}, T_{14c1}, T_{14c2}, T_{14c3}, v_{1c}, v_{1t}, TTT, TVM]^T, which are the production from zone 1, the number of non-travelers from zone 1, O-D demand from zone 1 to zone 4, O-D demands from zone 1 to zone 4 by car and by transit, flows on three routes between O-D pair (1, 4) using car network, flows on link 1 in car and transit networks, total travel time (TTT), and total vehicle miles (TVM) traveled, respectively.

Sensitivity analysis is an effective way to quantitatively analyze the behavior of demand and flow pattern with respect to (w.r.t.) perturbations of model inputs and parameters. Table 2 presents the derivatives of the twelve selected outputs w.r.t. eight selected inputs (i.e., the number of potential travelers from zone 1, and the link capacities in car network). We can observe that the sum of the derivatives of T_1 and T_{10} w.r.t. inputs is equal to zero except for N_1, which is equal to 1.0. This is consistent with the conservation constraint $T_1+T_{10}=N_1$. Also,

the sum of the derivatives of T_{14c} and T_{14t} w.r.t. inputs is equal to that of T_{14} due to the conservation constraint $T_{14c} + T_{14t} = T_{14}$. This sensitivity conservation also happens for the route choice conservation constraint $T_{14c1} + T_{14c2} + T_{14c3} = T_{14c}$. Additionally, the number of potential travelers of zone 1 (i.e., N_1) has a more significant impact on the trip production of this zone (i.e., T_1 and T_{10}), the travel demand between zone 1 and zone 4 (i.e., T_{14}), as well as the two network-wide performance measures. The link capacities in car network have more influences on the O-D demand by car (i.e., T_{14c}) compared with that by transit (i.e., T_{14t}). The car flow on link 1 is substantially impacted by this link capacity.

Table 2 Derivatives of outputs with respect to *inputs*

	N_1	C_{1c}	C_{2c}	C_{3c}	C_{4c}	C_{5c}	C_{6c}	C_{7c}
T_1	0.676	0.085	0.043	0.000	0.007	0.022	0.035	0.051
T_{10}	0.324	-0.085	-0.043	0.000	-0.007	-0.022	-0.035	-0.051
T_{14}	**0.334**	0.070	0.034	-0.001	0.046	-0.144	0.171	-0.222
T_{14c}	**0.058**	0.341	0.172	-0.002	0.088	-0.155	0.352	-0.190
T_{14t}	**0.276**	-0.271	-0.138	0.001	-0.042	0.010	-0.181	-0.032
T_{14c1}	**0.031**	0.295	0.047	-0.011	0.122	-0.130	-0.198	-0.076
T_{14c2}	**-0.003**	0.154	-0.078	0.013	-0.030	-0.064	0.215	0.039
T_{14c3}	**0.031**	-0.108	0.203	-0.004	-0.004	0.040	0.335	-0.154
v_{1c}	0.046	0.844	-0.093	0.009	0.035	0.104	0.033	0.059
v_{1t}	0.306	-0.280	-0.144	0.002	-0.023	-0.071	-0.116	-0.165
TTT	7.462	-0.169	-0.260	0.046	-0.054	-0.103	-0.078	0.053
TVM	6.158	0.669	0.443	0.016	0.062	0.196	0.348	0.481

Similarly, we can also obtain the derivatives of outputs w.r.t. parameters. Here, for demonstration purposes, we only consider the uncertainty of twelve selected parameters as listed in Table 3. The first four parameters are the attractiveness of zone 1, O-D pair (1, 4), and O-D pair (1, 4) by car and transit, respectively; the next four parameters are associated with the travel, destination, mode, and route choices; and the last four parameters are related to the link cost functions. Table 3 presents the derivatives of the selected outputs w.r.t. the selected parameters.

Among the above twelve selected parameters, the critical parameters are β_t, β_d, α_c and α_t. Interestingly, γ_c and γ_t are not critical parameters in this example even though they are the exponent parameters in the polynomial link cost functions. The critical parameter for both *TTT* and TVM is β_t. One percent perturbation of β_t will lead to a change of 20 and 16.5 units in *TTT* and *TVM*, respectively. Based on these derivatives, we need to put more efforts and

resources to enhance the calibration and validation of parameters β_t, β_d, α_c and α_t. It is worth pointing out that these results are only applicable to this example setting. In addition, among the four attractiveness parameters, T_1 is more sensitive to h_1, T_{14} is more sensitive to h_{14}, T_{14c} and T_{14t} are more sensitive to h_{14c} and h_{14t}, respectively. Also, the derivatives of T_1 w.r.t. h_1, T_{14} w.r.t. h_{14}, T_{14c} w.r.t. h_{14c}, and T_{14t} w.r.t. h_{14t} are all positive. These can be explained by the hierarchical tree-structure of the CTDM. When it goes down from the top (i.e., travel choice stage) of the hierarchical structure to the mode choice stage, the demand pattern becomes more sensitive to the attractiveness at the corresponding choice stage.

Table 3 Derivatives of outputs with respect to *parameters*

	h_1	h_{14}	h_{14c}	h_{14t}	β_t	β_d	β_m	β_r	α_c	γ_c	α_t	γ_t
T_1	**7.346**	3.627	0.630	2.997	181.851	-20.381	-3.860	-1.809	-8.207	-0.006	-44.079	-1.781
T_{10}	-7.346	-3.627	-0.630	-2.997	-181.851	20.381	3.860	1.809	8.207	0.006	44.079	1.781
T_{14}	3.627	**16.138**	3.769	12.369	89.788	-14.927	-1.781	-0.898	-0.608	0.048	-18.625	-0.740
T_{14c}	0.630	3.769	**7.386**	-3.616	15.590	-2.920	-4.046	0.054	-23.790	0.057	27.822	1.122
T_{14t}	2.997	12.369	-3.616	**15.985**	74.198	-12.007	2.264	-0.953	23.182	-0.009	-46.448	-1.862
T_{14c1}	0.335	2.323	4.403	-2.081	8.295	-1.661	-2.232	0.456	-6.953	-0.131	14.822	0.598
T_{14c2}	-0.038	0.296	0.336	-0.041	-0.935	-0.002	0.112	-0.677	-7.465	-0.254	-1.637	-0.066
T_{14c3}	0.332	1.151	2.646	-1.495	8.230	-1.257	-1.926	0.275	-9.371	0.442	14.638	0.591
v_{1c}	0.496	0.41	1.995	-1.586	12.283	-1.433	-2.546	-0.477	-37.325	-0.681	21.767	0.881
v_{1t}	3.322	1.45	-2.104	3.556	82.261	-9.155	1.075	-1.565	27.112	0.017	-136.973	-7.393
TTT	80.944	36.860	18.608	18.261	2004.097	-223.570	-63.076	-23.585	22.114	-0.323	405.862	16.340
TVM	66.810	32.996	5.818	27.188	1654.130	-185.397	-35.218	-16.253	-73.453	0.087	-382.486	-15.074

Note that the sensitivity information can be used to estimate the equilibrium solutions without the need to resolve the CTDM. Specifically, we can use the above derivatives and the first-order Taylor series approximation to estimate the equilibrium solutions under small perturbations of inputs or parameters. For demonstration purposes, Table 4 shows the exact and estimated solutions for the perturbations of δN_1=10 and $\delta\beta_t$=0.01. Recall that N_1 and β_t have a large derivative value as shown in Table 2 and Table 3. The exact solutions are obtained by resolving the perturbed CTDM. From Table 4, we can see that the estimated solutions are fairly close to the exact solutions even with 5%-perturbations.

Table 4 Estimated and exact solutions for perturbed input and parameter

Solution variable	Unperturbed solution	δN_1=10 (i.e., 200×5%)			$\delta\beta_t$=0.01 (i.e., 0.2×5%)		
		Exact	Estimated	Difference (Exact-Estimated)	Exact	Estimated	Difference (Exact-Estimated)
T_1	145.827	152.575	152.585	-0.010	147.621	147.645	-0.024
T_{10}	54.173	57.425	57.415	0.010	52.379	52.355	0.024
T_{14}	69.829	73.182	73.165	0.016	70.720	70.726	-0.007
T_{14c}	22.358	22.918	22.938	-0.019	22.509	22.514	-0.005
T_{14t}	47.470	50.263	50.228	0.035	48.211	48.212	-0.002
T_{14c1}	8.455	8.755	8.763	-0.008	8.535	8.538	-0.003
T_{14c2}	3.936	3.899	3.901	-0.002	3.926	3.926	0.000
T_{14c3}	9.967	10.264	10.273	-0.009	10.048	10.050	-0.002
v_{1c}	28.896	29.343	29.356	-0.014	29.017	29.019	-0.003
v_{1t}	61.560	64.607	64.620	-0.012	62.371	62.382	-0.011
TTT	1432.011	1506.850	1506.632	0.218	1451.807	1452.052	-0.245
TVM	1323.514	1384.904	1385.094	-0.191	1339.834	1340.056	-0.222

5.3 Uncertainty analysis results

(1) Uncertainty from Inputs

We assume the inputs are independently and normally distributed in order to simplify the characterization of inputs uncertainty. The mean of each input is the value given in the unperturbed condition (ε=0). The coefficient of variation (CoV) of inputs is set at 0.30. By setting CoV at 0.30, one can obtain statistically significant results (Zhao and Kockelman, 2002). Again, for demonstration purposes, we consider the eight inputs in Table 2.

Using Eq. (19), we can obtain the variance-covariance matrix of outputs, from which we can further calculate their standard deviation (SD) and CoV. One of the advantages of using the analytical sensitivity-based uncertainty analysis method is that the complex CTDM only needs to be solved once; while in the sampling-based method, the complex CTDM needs to be repeatedly solved many times according to the required sample size. Table 5 lists the estimated SD and CoV of selected outputs. One can see that the CoVs of most outputs are similar to that of inputs (i.e., 0.30). This is different from the results in Zhao and Kockelman (2002). In Table 5, we only consider the uncertainty stemming from the selected *inputs*. Also, the structure of the CTDM is different from the sequential four-step procedure. The sequential four-step procedure could be one reason that amplifies the uncertainty while the CTDM solves the multi-dimensional choice problem simultaneously. On the other hand, the CoV of link flows is not larger than that of the inputs. This is consistent with the results reported in

Zhao and Kockelman (2002) and Leurent (1998) due to the equilibrium nature of the traffic assignment step. The traffic assignment in the CTDM is equivalent to the logit-based stochastic user equilibrium (SUE). Due to the perceived error in the SUE model, the magnitude of uncertainty reduction of the SUE model is less than that of the user equilibrium (UE) model. In addition, since the CoV of a weighted sum of all independent random variables is less than the weighted average CoV of such variables, the CoV of *TVM* (i.e., 0.28) is less than the weighted average of link flow CoVs (i.e., 0.32). However, *TTT* is the sum of the product of link flows and link costs. The CoV of *TTT* is larger than that of *TVM*.

Table 5 Uncertainty of outputs due to *inputs uncertainty*

Solution variable	Mean	SD	CoV	90% confidence interval	
				5%	95%
T_1	145.83	40.56	0.28	79.11	212.54
T_{10}	54.17	19.47	0.36	22.15	86.19
T_{14}	69.83	20.08	0.29	36.80	102.86
T_{14c}	22.36	4.92	0.22	14.27	30.45
T_{14t}	47.47	16.72	0.35	19.96	74.98
T_{14c1}	8.46	3.16	0.37	3.26	13.65
T_{14c2}	3.94	1.67	0.42	1.19	6.69
T_{14c3}	9.97	3.02	0.30	5.00	14.94
v_{1c}	28.90	6.96	**0.24**	17.44	40.35
v_{1t}	61.56	18.54	**0.30**	31.07	92.05
TTT	1432.01	447.73	**0.31**	695.49	2168.53
TVM	1323.51	369.54	**0.28**	715.62	1931.41

By using the above estimated SD and further assuming the normality condition, the confidence intervals of outputs are easily calculated. The 90% confidence intervals of the selected outputs are also provided in Table 5. Again, these confidence intervals are obtained from the sensitivity-based approximation while avoiding the computationally demanding simulations. Considering the existence of inputs or parameters uncertainty, using the mean, SD, and confidence interval simultaneously provides a more complete uncertainty characterization of model outputs. With the standard deviations and confidence intervals of outputs, we can conduct hypothesis tests to obtain a statistically significant evaluation of network enhancement schemes.

With the sensitivity of performance measures w.r.t. link capacities and also the estimated variances of performance measures, we can identify the critical links in the studied network.

From Eq. (19), the variance of performance measures explicitly includes both the uncertainty of each link capacity and the partial derivatives of performance measures w.r.t. link capacities. Thus, it can be used to identify the most critical link that affects the system performance. Note that a weak link with higher capacity variability may not necessarily be the critical link. In contrast, a critical link must be one that is both important (i.e., substantial impact on system performance) and weak (i.e., large capacity variability) (Nicholson and Du, 1997). The critical links should be the prime candidates for strengthening, rather than those that are merely weak. Specifically, we can define the critical index of a link as follows:

$$\eta_{a_m} = \left(\frac{\partial M}{\partial C_{a_m}} \right)^2 \frac{\sigma_{C_{a_m}}^2}{\sigma_M^2}, \ \forall a_m . \tag{24}$$

The critical index indicates the proportion of overall uncertainty of performance measure M contributed by the uncertainty of link capacity C_{a_m}. For demonstration purposes, we use the *TTT* and *TVM* as the system performance measures. To eliminate the effect of N_1, we recalculate the variance of *TTT* and *TVM* by only considering the uncertainty of seven link capacities in car network. The critical index values w.r.t. *TTT* and *TVM* are shown in Figure 5. We can see link 1 and link 2 are more critical for both *TTT* and *TVM*. Note that both link capacities have the same standard deviation (i.e., 25×0.30). The critical index of different links only differs in the partial derivatives. From Table 2, link 1 and link 2 have larger partial derivatives w.r.t. both *TTT* and *TVM*. However, their ranking is different when using different system performance measures. From this viewpoint, it is necessary to explicitly consider multiple performance measures in critical link identification, especially for conflicting performance measures.

Furthermore, we can identify the critical inputs relative to the output uncertainty by the correlation of inputs and outputs. From Eq. (21), we can obtain the correlation of inputs and outputs as shown in Table 6. One can see that most selected outputs (e.g., trip production of zone 1, demand of O-D pair (1, 4), *TTT*, and *TVM*) are strongly correlated to the number of potential travelers of zone 1 (i.e., N_1). Compared to N_1, the link capacities in car network have a weaker correlation with the selected outputs. Among others, link 1 and link 2 are more correlated with the outputs, especially for those associated with mode choice and route choice. This result further verifies the importance of these two links as shown in Figure 5. In conclusion, the correlation results in Table 6 provide the model users insights into the

relationship between the model inputs and outputs, regardless of whether they are dependent or not, and how positively or negatively correlated if they are dependent.

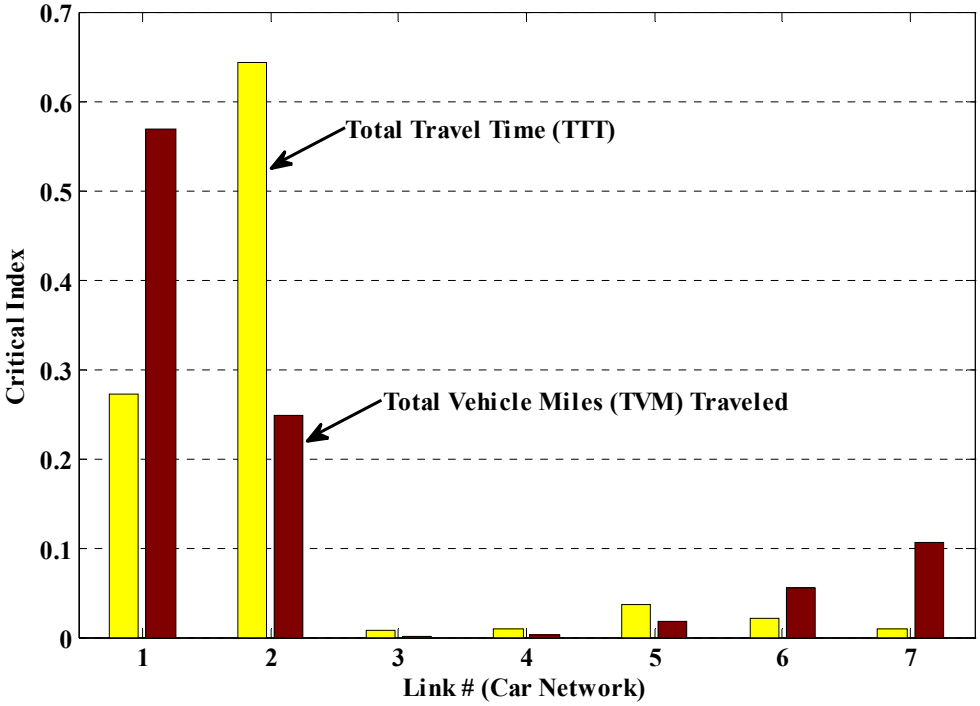


Figure 5 Critical index values of car network links with respect to *TTT* and *TVM*

Table 6 Correlation of outputs with *inputs*

Correlation	N_1	C_{1c}	C_{2c}	C_{3c}	C_{4c}	C_{5c}	C_{6c}	C_{7c}
T_1	1.000	0.016	0.008	0.000	0.001	0.002	0.004	0.006
T_{10}	0.999	-0.033	-0.017	0.000	-0.002	-0.005	-0.008	-0.012
T_{14}	0.997	0.026	0.013	0.000	0.010	-0.032	0.038	-0.050
T_{14c}	0.707	0.520	0.262	-0.002	0.081	-0.141	0.322	-0.174
T_{14t}	0.989	-0.122	-0.062	0.000	-0.011	0.003	-0.049	-0.009
T_{14c1}	0.585	0.699	0.112	-0.016	0.173	-0.185	-0.282	-0.108
T_{14c2}	-0.125	0.692	-0.351	0.034	-0.080	-0.173	0.578	0.105
T_{14c3}	0.608	-0.267	0.503	-0.006	-0.006	0.059	0.500	-0.229
v_{1c}	0.396	0.909	-0.100	0.006	0.023	0.067	0.021	0.038
v_{1t}	0.990	-0.113	-0.058	0.000	-0.006	-0.017	-0.028	-0.040
TTT	1.000	-0.003	-0.004	0.000	-0.001	-0.001	-0.001	0.001
TVM	1.000	0.014	0.009	0.000	0.001	0.002	0.004	0.006

(2) Uncertainty from Parameters

For the uncertainty analysis from parameters, we assume the parameters are also independently and normally distributed. The mean of each parameter is the value given in the unperturbed condition (ε=0). We also set the CoV of all parameters at 0.30. Given the partial

derivatives of selected outputs w.r.t. parameters (in Table 3), we can estimate the SD, CoV, and 90% confidence interval of outputs as shown in Table 7. Since $T_1 + T_{10} = N_1$, the variances of T_1 and T_{10} are exactly equal. From the travel choice stage to the mode choice stage, the CoV seems to increase, while the CoV of link flows drops due to the equilibrium nature of the traffic assignment step. Also, due to the different network-wide aggregation manners, the CoV of *TVM* is smaller than that of *TTT*.

Table 7 Uncertainty of outputs due to *parameters uncertainty*

Solution variable	Mean	SD	CoV	90% confidence interval	
				5%	95%
T_1	145.83	16.72	***0.11***	118.32	173.34
T_{10}	54.17	16.72	0.31	26.66	81.68
T_{14}	69.83	23.36	0.33	31.40	108.26
T_{14c}	22.36	9.82	0.44	6.21	38.50
T_{14t}	47.47	22.98	0.48	9.66	85.28
T_{14c1}	8.46	5.81	0.69	0.00*	18.02
T_{14c2}	3.94	0.78	0.20	2.66	5.21
T_{14c3}	9.97	3.65	0.37	3.96	15.97
v_{1c}	28.90	3.64	**0.13**	22.91	34.88
v_{1t}	61.56	10.06	**0.16**	45.01	78.11
TTT	1432.01	182.49	**0.13**	1131.81	1732.21
TVM	1323.51	152.03	**0.11**	1073.42	1573.60

*: The actual value is -1.11 due to the normality assumption and the large CoV. We truncate negative flows to zero.

We compare the CoVs of outputs due to the input uncertainty (in Table 5) and due to the parameter uncertainty (in Table 7). Except for the travel choice step, the impact of *parameter* uncertainty on output uncertainty of each choice step is generally higher than that of *input* uncertainty. In other words, to improve the confidence level of the estimated outputs of each choice step, improving the accuracy of parameter estimation is more effective than that of improving input estimation. We can also use the proposed approach to quantify the possible benefit of improving the quality of parameter estimation. We vary the CoVs of parameters from 0.10 to 0.50 with an interval of 0.20. The CoVs of outputs are recalculated using the same sensitivity information but with different CoVs of parameters. From Figure 6, we can observe that when the accuracy of parameter estimation increases, the accuracy of output estimation will also increase.

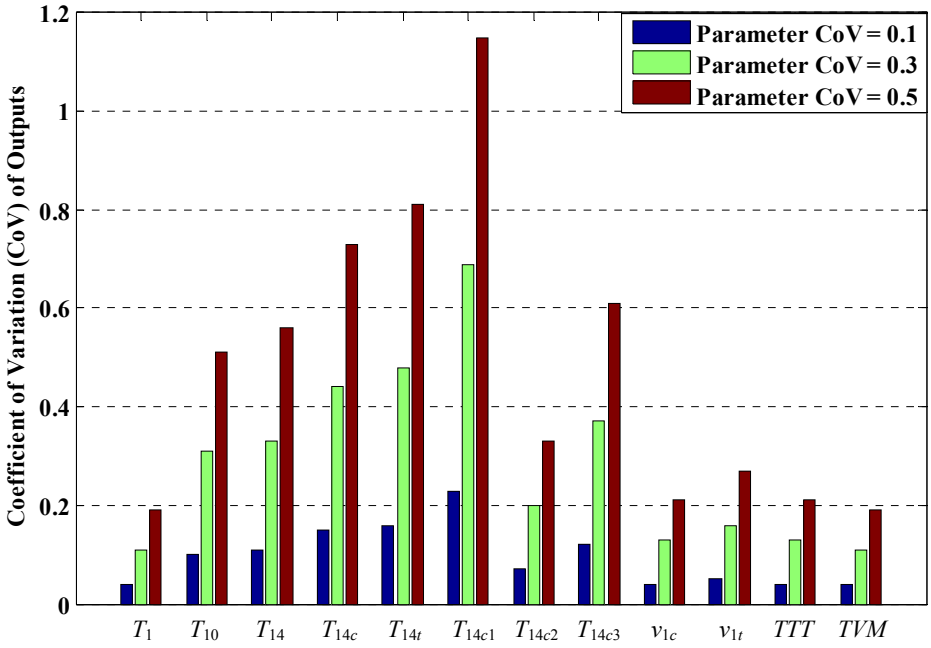


Figure 6 CoV of outputs due to *parameter uncertainty* with different CoVs

Similar to the input uncertainty analysis, we can examine the correlation of outputs with parameters as shown in Table 8. T_1 is strongly correlated to h_1, T_{14} is strongly correlated to h_{14}, T_{14c} and T_{14t} are strongly correlated to h_{14c} and h_{14t}, respectively. Also, the correlations of T_{14c} with h_{14c} and h_{14t} and T_{14t} with h_{14c} and h_{14t} indicate the mode choice competition between car and transit connecting this O-D pair. In addition, the mode-specific link flows are correlated with the corresponding mode-specific parameters in the link cost functions. The correlation matrix can help the model users to identify the critical parameters for improvement.

Table 8 Correlation of outputs with *parameters*

	h_1	h_{14}	h_{14c}	h_{14t}	β_t	β_d	β_m	β_r	α_c	γ_c	α_t	γ_t
T_1	0.659	0.228	0.040	0.194	0.652	-0.183	-0.069	-0.065	-0.022	0.000	-0.047	-0.064
T_{10}	-0.659	-0.228	-0.040	-0.194	-0.652	0.183	0.069	0.065	0.022	0.000	0.047	0.064
T_{14}	0.233	0.725	0.169	0.572	0.231	-0.096	-0.023	-0.023	-0.001	0.002	-0.014	-0.019
T_{14c}	0.096	0.403	0.790	-0.398	0.095	-0.045	-0.124	0.003	-0.109	0.007	0.051	0.069
T_{14t}	0.196	0.565	-0.165	0.751	0.194	-0.078	0.030	-0.025	0.045	0.000	-0.036	-0.049
T_{14c1}	0.086	0.420	0.795	-0.387	0.086	-0.043	-0.115	0.047	-0.054	-0.027	0.046	0.062
T_{14c2}	-0.073	0.399	0.454	-0.056	-0.072	0.000	0.043	-0.523	-0.432	-0.393	-0.038	-0.051
T_{14c3}	0.137	0.331	0.761	-0.442	0.135	-0.052	-0.158	0.045	-0.115	0.145	0.072	0.097
v_{1c}	0.204	0.118	0.576	-0.471	0.203	-0.059	-0.210	-0.079	-0.462	-0.225	0.108	0.145
v_{1t}	0.495	0.151	-0.220	0.382	0.490	-0.136	0.032	-0.093	0.121	0.002	-0.245	-0.441
TTT	0.665	0.212	0.107	0.108	0.659	-0.184	-0.104	-0.078	0.005	-0.002	0.040	0.054
TVM	0.659	0.228	0.040	0.193	0.653	-0.183	-0.069	-0.064	-0.022	0.001	-0.045	-0.059

(3) Total Uncertainty

The total uncertainty analysis is used to investigate the combined impact of inputs and parameters uncertainty on the outputs uncertainty. The CoV of both inputs and parameters is set at 0.3. Following the same approach as in the previous two sections, the CoVs of outputs are listed in Table 9. By comparing with Table 5 and Table 7, we can find that the uncertainty (in terms of both SD and CoV) of outputs stemming from both inputs and parameters uncertainty is not simply the sum of uncertainties from inputs and parameters individually. In addition, compared to the *TTT*, *TVM* has a smaller CoV. *TVM* seems to be a more reliable system performance metric given the uncertainty of both inputs and parameters.

Table 9 Uncertainty of outputs due to *both input and parameter uncertainty*

Solution variable	Mean	SD	CoV	90% confidence interval	
				5%	95%
T_1	145.83	43.87	0.30	73.66	217.99
T_{10}	54.17	25.66	0.47	11.96	96.39
T_{14}	69.83	30.81	0.44	19.15	120.51
T_{14c}	22.36	10.98	0.49	4.30	40.42
T_{14t}	47.47	28.42	0.60	0.71	94.23
T_{14c1}	8.46	6.62	0.78	0.00*	19.34
T_{14c2}	3.94	1.84	0.47	0.90	6.97
T_{14c3}	9.97	4.74	0.48	2.17	17.76
v_{1c}	28.90	7.86	0.27	15.97	41.82
v_{1t}	61.56	21.09	0.34	26.86	96.26
TTT	1432.01	483.49	0.34	636.66	2227.36
TVM	1323.51	399.59	0.30	666.19	1980.84

*: The actual value is -2.43 due to the normality assumption and the large CoV. We truncate negative flows to zero.

To investigate the output uncertainty at each travel choice step of the CTDM, we calculate the average CoV of outputs at each step using the derivatives of all outputs w.r.t. the eight selected inputs (Table 2) and twelve parameters (Table 3). The results are graphically depicted in Figure 7. The average CoV of outputs at each step is increasing as the choice step moves from the top to the bottom according to the hierarchical structure of the CTDM shown in Figure 4. However, the average CoV of link flows is smaller than that of mode-specific O-D demands. This is consistent with the results reported in Zhao and Kockelman (2002) and Leurent (1998) due to the equilibrium nature of the traffic assignment step. The traffic assignment in the CTDM is equivalent to the logit-based stochastic user equilibrium (SUE). Due to the perceived error in the SUE model, the magnitude of uncertainty reduction of the SUE model is less than that of the user equilibrium (UE) model, which is used in Zhao and

Kockelman (2002). We should point out that the average CoV values at different travel choice steps are not the same as the uncertainty propagation in the sequential four-step procedure (Zhao and Kockelman, 2002). In the CTDM, the results of each step are calculated simultaneously. There is essentially no uncertainty propagation due to the integrated travel choice steps. The figure is only used to illustrate the relative uncertainty magnitudes at different aggregate levels due to the input and parameter uncertainty.

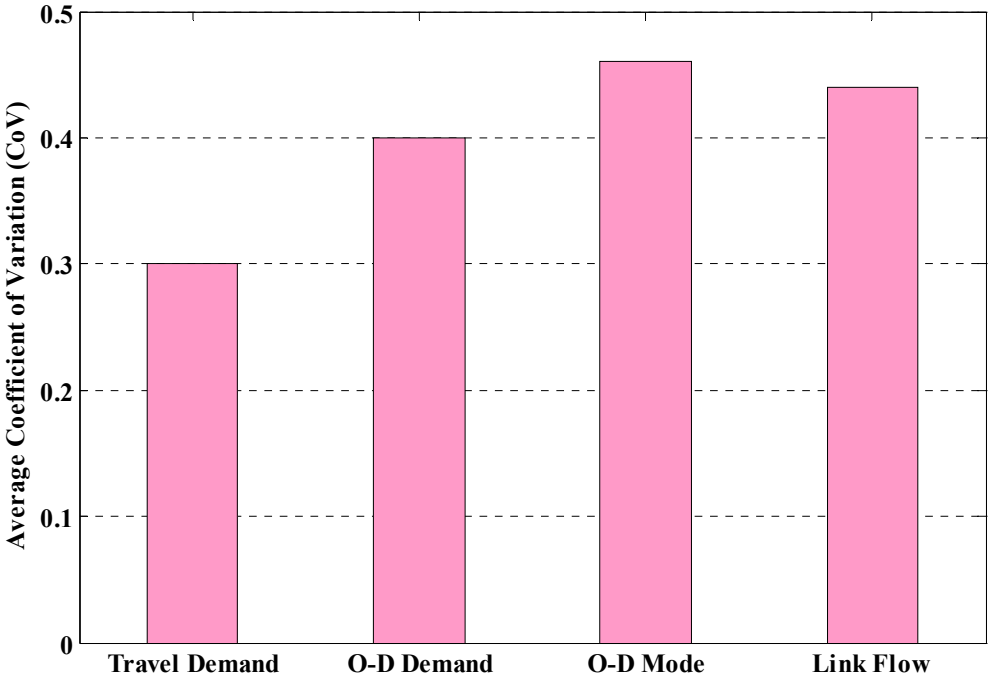


Figure 7 Output uncertainty at each travel choice step

6. Conclusions

In this paper, we proposed a systematic framework for the uncertainty analysis of a combined travel demand model (CTDM). The CTDM is based on the random utility theory of user behavior, which is behaviorally richer than the sequential four-step procedure. The CTDM can be formulated as an equivalent convex optimization problem, which makes it possible to conduct the sensitivity analysis. We employed the analytical sensitivity-based method for the uncertainty analysis of the CTDM, which requires significantly less computational efforts than the sampling-based methods. Furthermore, uncertainties stemming from inputs and parameters can be treated separately so that the individual and collective effects of uncertainty on the outputs can be clearly assessed and quantified. The numerical results indicated that at each disaggregate choice step except for the travel choice step, the impact of parameter uncertainty on the output uncertainty is generally more important than that of input uncertainty. This information enables planners to effectively allocate the limited resources for

input data collection and parameter estimation of key variables. Using the sensitivity-based uncertainty analysis, we can also estimate the possible benefits of improving the parameter accuracy.

In this paper, we only used the mean and variance to characterize the input uncertainty. This is justified for normal distribution. However, for asymmetric distributions (e.g., lognormal distribution), we need higher-order moment information (e.g., skewness and kurtosis) to enrich the characterization. How to make use of higher-order moment information to enhance the quality of uncertainty analysis will be a valuable future research direction. In addition, the semi-deviation measure could also be used to quantify uncertainty since variance is a symmetric statistic and gives equal weight to deviations below and above the mean without considering the risks associated with extreme outcomes (Li *et al*., 2012). These uncertainty analysis methods in turn could be used in evaluating transportation network reliability, vulnerability, and flexibility measures (Chen *et al*., 1999, 2002, 2007b, 2013; Chen and Kasikitwiwat, 2011), modeling network equilibrium under uncertainty (Zhou and Chen, 2008; Chen and Zhou, 2010; Chen *et al*., 2011b; Xu *et al*., 2013), and developing network design under uncertainty (Chootinan *et al.,* 2005; Chen *et al*., 2007a, 2010, 2011a; Yim *et al*., 2011; Chen and Xu, 2012).

Acknowledgements

The authors are grateful to three anonymous referees for their constructive comments and suggestions to improve the quality and clarity of the paper. The authors also wish to express their thanks to Professor Hai Yang for his valuable comments on the sensitivity analysis formulation. This research was supported by the Oriental Scholar Professorship Program sponsored by the Shanghai Ministry of Education in China to Tongji University, National Natural Science Foundation of China (71171147), Fundamental Research Funds for the Central Universities, and the China Scholarship Council.

Appendix: Invertibility Proof

For the Logit-based SUE model, $\mathbf{M} = \begin{bmatrix} \nabla^2_{f_k^w} L & \mathbf{\Lambda}^T \\ \mathbf{\Lambda} & \mathbf{0} \end{bmatrix}$, where $\nabla^2_{f_k^w} L$ is the Hessian matrix of Lagrangian function L with respect to route flow f_k^w (route k between O-D pair w), and $\mathbf{\Lambda}$ is the OD-route incidence matrix. $\nabla^2_{f_k^w} L = \Delta^T \nabla_{\mathbf{v}} \mathbf{t} \Delta + \frac{1}{\theta} diag\left(1/f_k^w\right)$ is positive definite, since the sum of a positive semi-definite matrix (i.e., $\Delta^T \nabla_{\mathbf{v}} \mathbf{t} \Delta$) and a positive definite matrix (i.e., $diag\left(1/f_k^w\right)$) is a positive definite matrix.

To prove **M** is non-singular (or invertible), it suffices to prove that all columns of **M** are linearly independent. Consider a nonzero vector $\boldsymbol{\lambda} = \left[\tilde{\boldsymbol{\lambda}}^T, \hat{\boldsymbol{\lambda}}^T\right]^T \neq \mathbf{0}$, where $\tilde{\boldsymbol{\lambda}}$ is a column vector with the number of elements equal to the *number of routes* and $\hat{\boldsymbol{\lambda}}$ is a column vector with the number of elements equal to the *number of O-D pairs*.

Let $\mathbf{M}\boldsymbol{\lambda} = \mathbf{0}$, we have

$$\left(\nabla^2_{f_k^w} L\right)\tilde{\boldsymbol{\lambda}} + \mathbf{\Lambda}^T \hat{\boldsymbol{\lambda}} = \mathbf{0}, \tag{A1}$$

$$\mathbf{\Lambda}\tilde{\boldsymbol{\lambda}} = \mathbf{0}. \tag{A2}$$

Multiplying both sides of Eq. (A1) (from the left side) by $\tilde{\boldsymbol{\lambda}}^T$ yields:

$$\tilde{\boldsymbol{\lambda}}^T \left(\nabla^2_{f_k^w} L\right)\tilde{\boldsymbol{\lambda}} + \tilde{\boldsymbol{\lambda}}^T \mathbf{\Lambda}^T \hat{\boldsymbol{\lambda}} = \tilde{\boldsymbol{\lambda}}^T \left(\nabla^2_{f_k^w} L\right)\tilde{\boldsymbol{\lambda}} + \left(\mathbf{\Lambda}\tilde{\boldsymbol{\lambda}}\right)^T \hat{\boldsymbol{\lambda}} = 0. \tag{A3}$$

Substituting Eq. (A2) into Eq. (A3) yields:

$$\tilde{\boldsymbol{\lambda}}^T \left(\nabla^2_{f_k^w} L\right)\tilde{\boldsymbol{\lambda}} = 0. \tag{A4}$$

Since $\nabla^2_{f_k^w} L$ is positive definite, Eq. (A4) implies that $\tilde{\boldsymbol{\lambda}} = \mathbf{0}$.

From $\tilde{\boldsymbol{\lambda}} = \mathbf{0}$ and Eq. (A1), we have $\mathbf{\Lambda}^T \hat{\boldsymbol{\lambda}} = \mathbf{0}$. Since $\mathbf{\Lambda}^T$ (i.e., OD-route incidence matrix) contains linearly independent columns and is of full column rank equal to the number of O-D pairs, we thus conclude $\hat{\boldsymbol{\lambda}} = \mathbf{0}$. To sum up, we have $\tilde{\boldsymbol{\lambda}} = \mathbf{0}$ and $\hat{\boldsymbol{\lambda}} = \mathbf{0}$. These results contradict our assumption of a nonzero vector $\boldsymbol{\lambda} = \left[\tilde{\boldsymbol{\lambda}}^T, \hat{\boldsymbol{\lambda}}^T\right]^T \neq \mathbf{0}$. Therefore, matrix **M** is invertible.

References

Bar-Gera, H., Boyce, D., 2003. Origin-based algorithms for combined travel forecasting models. *Transportation Research Part B* 37 (5), 405-422.

Bekhor, S., Ben-Akiva, M. E., Ramming, M. S. 2006. Evaluation of choice set generation algorithms for route choice models. *Annals of Operations Research* 144, 235–247.

Bekhor, S., Toledo, T., Prashker, J.N., 2008. Effects of choice set size and route choice models on path-based traffic assignment. *Transportmetrica* 4(2), 117-133.

Bell, M.G.H., Cassir, C., Iida, Y., Lam, W.H.K., 1999. A sensitivity based approach to network reliability assessment. In: Ceder, A. (Ed.), *Proceedings of the 14th International Symposium on Transportation and Traffic Theory*. Oxford, Pergamon, 283-300.

Bonsall, P.W., Champerwone, A.F., Mason, A.C., Wilson, A.G., 1977. Transport modeling: sensitivity analysis and policy testing. *Progress in Planning* 7 (3), 153-237.

Bovy, P. H. L., Fiorenzo-Catalano, S. 2007. Stochastic route choice set generation: behavioral and probabilistic foundations. *Transportmetrica* 3(3) 173–189.

Bowman, J.L., Gopinath, D., Ben-Akiva, M., 2002. Estimating the probability distribution of a travel demand forecast. Working paper, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology.

Boyce, D.E., 2002. Is the sequential travel forecasting procedure counterproductive? *ASCE Journal of Urban Planning and Development* 128 (4), 169-183.

Boyce, D., Bar-Gera, H., 2001. Network equilibrium models of travel choices with multiple classes. In: Lahr, M.L., Miller, R.E. (Eds.), *Regional Science Perspectives in Economic Analysis*. Elsevier Science, Oxford, 85-98.

Boyce, D., Chon, K.S., Lee, Y.J., Lin, K.T., 1983. Implementation and computational issues for combined models of location, destination, mode, and route choice. *Environment and Planning A* 15 (9), 1219-1230.

Boyce, D., Daskin, M.S., 1997. Urban transportation. In: ReVelle, C., McGarity, A. (Eds.), *Design and Operation of Civil and Environmental Engineering Systems*. John Wiley & Sons, Inc., New York, 277-341.

Boyce, D., Lupa, M., Zhang, Y., 1994. Introducing feedback into four-step travel forecasting procedure versus equilibrium solution of combined model. *Transportation Research Record* 1443, 65-74.

Boyles, S.D., *2012.* Bush-based sensitivity analysis for approximating subnetwork diversion. *Transportation Research Part B* 46 (1), 139-155.

Brattin, W.J., Barry, T.M., Chiu, N., 1996. Monte Carlo modeling with uncertain probability density functions. *Human and Ecological Risk Assessment* 2 (4), 820-840.

Cascetta, E., Nuzzolo, A., Russo, F., Vitetta, A., 1996. A modified logit route choice model overcoming path overlapping problems: specification and some calibration results for interurban networks. In: *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, Leon, France, 697-711.

Cascetta, E., Russo, F., Vitetta, A., 1997. Stochastic user equilibrium assignment with explicit path enumeration: comparison of models and algorithms. In: *Proceedings of International Federation of Automatic Control Transportation Systems*, Chania, Greece, 1078-1084.

Chen, A., Kasikitwiwat, P., 2011. Modeling network capacity flexibility of transportation networks. *Transportation Research Part A* 45(2), 105-117.

Chen, A., Kasikitwiwat, P., Yang, C., 2013. Alternate capacity reliability measures for transportation networks. *Journal of Advanced Transportation* 47(1), 79-104.

Chen, A., Kim, J., Lee, S., Kim, Y., 2010. Stochastic multi-objective models for network design problem under demand uncertainty. *Expert Systems with Applications* 37(2), 1608-1619.

Chen, A., Kim, J., Zhou, Z., Chootinan, P., 2007a. Alpha reliable network design problem. *Transportation Research Record* 2029, 49-57.

Chen, A., Yang, C., Kongsomsaksakul, S., Lee, M., 2007b. Network-based accessibility measures for vulnerability analysis of degradable transportation networks. *Network and Spatial Economics* 7(3), 241-256.

Chen, A., Xu, X., 2012. Goal programming approach to solving the network design problem with multiple objectives and demand uncertainty. *Expert Systems with Applications* 39(4), 4160-4170.

Chen, A., Yang, H., Lo, H.K., Tang, W., 1999. A capacity related reliability for transportation networks. *Journal of Advanced Transportation* 33(2), 183-200.

Chen, A., Yang, H., Lo, H.K., Tang, W., 2002. Capacity reliability of a road network: an assessment methodology and numerical results. *Transportation Research Part B* 36(3), 225-252.

Chen, A., Zhou, Z., 2010. The α-reliable mean-excess traffic equilibrium model with stochastic travel times. *Transportation Research Part B* 44(4), 493-513.

Chen, A., Zhou, Z., Chootinan, P., Ryu, S., Yang, C., Wong, S.C., 2011a. Transport network design problem under uncertainty: A review and new developments. *Transport Reviews* 31(6), 743-768.

Chen, A., Zhou, Z., Lam, W.H.K., 2011b. Modeling stochastic perception error in the mean-excess traffic equilibrium model with stochastic travel times. *Transportation Research Part B* 45(10), 1619-1640.

Chootinan, P., Wong, S.C., Chen, A., 2005. A reliability-based network design problem. *Journal of Advanced Transportation* 39(3), 247-270.

Clark, S.D., Watling, D.P., 2000. Probit based sensitivity analysis for general traffic networks. *Transportation Research Record* 1733, 88–95.

Clark, S.D., Watling, D.P., 2002. Sensitivity analysis of the probit-based stochastic user equilibrium assignment model. *Transportation Research Part B* 36 (7), 617-635.

Dafermos, S., Nagurney, A., 1984. Sensitivity analysis for the asymmetric network equilibrium problem. *Mathematical Programming* 28(2), 174-184.

Davis, G., 1994. Exact solution of the continuous network design problem via stochastic user equilibrium assignment. *Transportation Research Part B* 28(1), 61-75.

de Jong, G., Daly, A., Pieters, M., Miller, S., Plasmeijer, R., Hofman, F., 2007. Uncertainty in traffic forecasts: literature review and new results for The Netherlands. *Transportation* 34(4), 375-395.

Du, M., Cheng, L., Rakha, H., 2012. Sensitivity analysis of combined distribution-assignment model with applications. *Journal of the Transportation Research Board* 2284, 10-20.

Evans, S., 1976. Derivation and analysis of some models for combining trip distribution and assignment. *Transportation Research* 10 (1), 37-57.

Fiacco, A.V., 1983. *Introduction to sensitivity and stability analysis in nonlinear programming*. Academic Press, New York.

Fisk, C., 1980. Some developments in equilibrium traffic assignment. *Transportation Research Part B* 14(3), 243-255.

Florian, M., Nguyen, S., 1978. A combined trip distribution model split and trip assignment model. *Transportation Research* 12 (4), 241-246.

Florian, M., Wu, J.H., He, S., 2002. A multi-class multi-mode variable demand network equilibrium model with hierarchical logit structures. In: Gendreau, M., Marcotte, P.

(Eds.), *Transportation and network analysis: Current trends*. Kluwer, Dordrecht, The Netherlands, 119–133.
Garret, M., Wachs, M., 1996. *Transportation Planning on Trial*. Sage, Thousand Oaks, California.
Huang, H.J., Lam, W.H.K., 1992. Modified Evans' algorithms for solving the combined trip distribution and assignment problem. *Transportation Research Part B* 26 (4), 325–337.
Krishnamurthy, S., Kockelman, K.M., 2003. Propagation of uncertainty in transportation land use models: Investigation of DRAM-EMPAL and UTPP predictions in Austin, Texas. *Transportation Research Record* 1831, 219-229.
Lam, W.H.K., Huang, H.J., 1992. A combined trip distribution and assignment model for mulitple user classes. *Transportation Research Part B* 26 (4), 275-287.
LeBlanc, L.J., Morlok, E.K., Pierskalla, W.P., 1975. An efficient approach to solving the road network equilibrium traffic assignment problem. *Transportation Research* 9 (5), 309-318.
Leurent, F., 1998. Sensitivity and error analysis of the dual criteria traffic assignment model. *Transportation Research Part B* 32 (3), 189-204.
Li, Z.C., Lam, W.H.K., Wong S.C., Sumalee, A., 2012. Environmentally sustainable toll design for congested road networks with uncertain demand. *International Journal of Sustainable Transportation* 6(3), 127-155.
Liu, J., Ma, S., Huang, C., Ma, S., 2010. A dimension-reduced method of sensitivity analysis for stochastic user equilibrium assignment model. *Applied Mathematical Modelling* 34, 325-333
McNally, M.G., 2000a. The four-step model. In: Hensher, D.A., Button, K.J. (Eds.), *Handbook of Transport Modelling*. Pergamon, Oxford, 35-52.
McNally, M.G., 2000b. The activity-based approach. In: Hensher, D.A., Button, K.J. (Eds.), *Handbook of Transport Modelling*. Pergamon, Oxford, 53-69.
Nicholson, A., Du, Z.P., 1997. Degradable transportation systems: an integrated equilibrium model. *Transportation Research Part B* 31(3), 209-223.
Oppenheim, N., 1995. *Urban Travel Demand Modeling*. John Wiley & Sons, Inc., New York.
Ortuzar, J.D., Willumsen, L.D., 2001. *Modeling Transport* (3rd ed.). John Wiley & Sons, Inc., New York.
Patriksson, M., 1994. *The Traffic Assignment Problem: Models and Methods*. VSP, Utrecht, The Netherlands.
Pradhan, A., Kockelman, K.M., 2002. Uncertainty propagation in an integrated land use-transportation modeling framework: output variation via UrbanSim. *Transportation Research Record* 1805, 128-135.
Prato, C.G., Bekhor, S. 2006. Applying branch-and-bound technique to route choice set generation. *Transportation Research Record 1985*, 19–28.
Prato C. G., Bekhor, S. 2007. Modeling route choice behavior: How relevant is the composition of choice set? *Transportation Research Record 2003*, 64–73.
Qiu, Y., Magnanti, T.L., 1989. Sensitivity analysis for variational inequalities defined on polyhedral sets. *Mathematics of Operations Research* 14 (3), 410-432.
Rai, S.N., Krewski, D., Bartlett, S., 1996. A general framework for the analysis of uncertainty and variability in risk assessment. *Human and Ecological Risk Assessment* 2(4), 972-989.
Rasouli, S., Timmermans, H., 2012. Uncertainty in travel demand forecasting models: literature review and research agenda. *Transportation Letters* 4, 55-73.
Robbins, J., 1978. Mathematical modeling – the errors of our ways. *Traffic Engineering and Control* 19 (1), 32-35.

Safwat, K.N.A., Magnanti, T.L., 1988. A combined trip generation, trip distribution, mode split, and trip assignment model. *Transportation Science* 22 (1), 14-30.

Ševčíková, H., Raftery, A.E., Waddell, P.A., 2007. Assessing uncertainty in urban simulations using Bayesian melding. *Transportation Research Part B* 41(6), 652-669.

Tobin, R.L., 1986. Sensitivity analysis for variational inequalities. *Journal of Optimization Theory and Applications* 48 (1), 191-204.

Tobin, R.L., Friesz, T.L., 1988. Sensitivity analysis for equilibrium network flows. *Transportation Science* 22 (4), 242-250.

Vose, D., 2000. *Risk Analysis a Quantitative Guide*. John Wiley & Sons Ltd., England.

Waller, S.T., Schofer, J.L., Ziliaskopoulos, A.K., 2001. Evaluation with traffic assignment under demand uncertainty. *Transportation Research Record* 1771, 69-74.

Xu, X., Chen, A., Cheng, L., 2013. Assessing the effects of stochastic perception error under travel time variability. *Transportation* 40(3), 525-548.

Xu, M., Chen, A., Gao, Z., 2008. An improved origin-based algorithm for solving the combined distribution and assignment problem. *European Journal of Operational Research* 188(2), 354-369.

Yang, C., Chen, A., 2009. Sensitivity analysis of the combined travel demand model with applications. *European Journal of Operational Research* 198(3), 909-921.

Yang, C., Chen, A., Xu, X., 2013. Improved partial linearization algorithm for solving the combined travel-destination-mode-route choice problem. *Journal of Urban Planning and Development* 139(1), 22-32.

Yang, H., 1997. Sensitivity analysis for the elastic-demand network equilibrium problem with applications. *Transportation Research Part B* 31 (1), 55-70.

Yang, H., Bell, M.G.H., 2007. Sensitivity analysis of network traffic equilibria revisited: the corrected approach. In: Heydecker, B. (Ed.), *Mathematics in Transport*, Elsevier, 373-411.

Yang, H., Meng, Q., Bell, M.G.H., 2001. Simultaneous estimation of the origin-destination matrices and travel-cost coefficient for congested networks in a stochastic equilibrium. *Transportation Science* 35, 107-123.

Yim, K., Wong, S.C., Chen, A., Wong, C.K., Lam, W.H.K., 2011. A reliability-based land use and transportation optimization model. *Transportation Research Part C* 19(2), 351-362.

Ying, J.Q., Miyagi, T., 2001. Sensitivity analysis for stochastic user equilibrium network flows – a dual approach. *Transportation Science* 35 (2), 124-133.

Ying, J.Q., Yang, H., 2005. Sensitivity analysis of stochastic user equilibrium flows in a bi-modal network with application to optimal pricing, *Transportation Research Part B* 39, 769-795.

Zhang, Y., Boyce, D., 2000. Variable comparison for introducing "feedback" in travel demand model: application in New York metropolitan area. In: *Proceedings of the Second International Conference on Transportation and Traffic Studies*, Beijing, China.

Zhao, Y, Kockelman, K.M., 2002. The propagation of uncertainty through travel demand models: an exploratory analysis. *Annals of Regional Science* 36 (1) 145-163.

Zhou, Z., Chen, A., 2008. Comparative analysis of three user equilibrium models under stochastic demand. *Journal of Advanced Transportation* 42(3), 239-263.

Zhou, Z., Chen, A., Wong, S.C., 2009. Alternative formulations of a combined trip generation, trip distribution, modal split, and traffic assignment model. *European Journal of Operational Research* 198(1), 129-138.