The HKU Scholars Hub   The University of Hong Kong   香港大學學術庫

Title	Reliable network design under supply uncertainty with probabilistic guarantees
Author(s)	Szeto, WY; Wang, AB
Citation	Transportmetrica A: Transport Science, 2016, v. 12 n. 6, p. 504-532
Issued Date	2016
URL	http://hdl.handle.net/10722/226308
Rights	This is an Accepted Manuscript of an article published by Taylor & Francis in Transportmetrica A: Transport Science on 24 Mar 2016, available online: http://wwww.tandfonline.com/10.1080/23249935.2016.1154625; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

Reliable network design under supply uncertainty with probabilistic guarantees

W.Y. Szeto[1] and Anny B. Wang
Department of Civil Engineering, The University of Hong Kong, Hong Kong, P. R. China
[1] Phone: (852) 28578552, Fax: (852) 25595337, Email: ceszeto@hku.hk

Abstract

This paper proposes a bi-level risk averse network design model for transportation networks with heterogeneous link travel time distributions. The objective of the network design is to minimize the total system travel time budget, which consists of the mean total system travel time and a safety margin. The design is achieved by selecting optimal link capacity expansions subject to a fixed expansion budget. Users' selfish behavior and risk attitude are captured in the lower-level traffic assignment constraints, in which travelers select routes to minimize their own path travel time budget. The properties of the design problem are analyzed analytically and numerically. The analysis shows that despite the lack of knowledge of travel time distributions, the probabilities that the actual total system travel time and the actual path travel time are respectively within the optimal total system travel time budget and the minimum path travel time budget under optimal design have lower bounds. The lower bounds are related to the system manager's and travelers' risk aversion. The optimal total system travel time budget is proven to be bounded below even when the link expansion budget is unlimited.

Key words: Reliability-based network design, total system travel time budget, probabilistic guarantee

1. Introduction

The network design problem (NDP) involves determining the link capacity expansions and/or link additions in a transport network (e.g., Steenbrink, 1974; Adbulaal and LeBlanc, 1979; Dantzig et al., 1979; LeBlanc and Boyce, 1986; Marcotte, 1986; Ben-Ayed et al., 1988; Suh and Kim, 1992; Friesz et al., 1992; Davis, 1994; Yang, 1997; Yang and Bell, 1998; Meng et al., 2001; Meng and Yang, 2002; Yang and Wang, 2002; Chiou, 2005; Szeto and Lo, 2005; Gao et al., 2007; Chiou, 2008; Xu et al., 2009; Mathew and Sharma, 2009; Ng and Waller, 2009b; Wang and Lo, 2010; Szeto et al., 2010; 2015; Miandoabchi et al. 2012a,b, 2013, 2015; Szeto et al., 2014; Jiang and Szeto, 2015; Khooban et al., 2015; Liu and Wang, 2015; Wang et al., 2015a; Haas and Bekhor, 2016). Comprehensive reviews have been conducted by Magnanti and Wong (1984), Yang and Bell (1988), and Farahani et al. (2013). The classical NDP does not account for uncertainty. However, the performance of a transportation network is often uncertain because travel time can be affected by, for example, bad weather, road accidents, and special events. To accurately evaluate and improve the network performance, it is vital to incorporate this uncertainty into the design problem. Hence, transport NDPs under uncertainty have been proposed and examined in the literature.

Studies of the transport NDP under uncertainty mainly consider either demand uncertainty or supply uncertainty. Chen et al. (2011) conducted a detailed review and classified existing network design models into six classes, depending on the criterion for hedging against uncertainty: the expected value model, mean-variance model, chance-constraint model, probability model, min-max model, and alpha-reliable model. The expected value model minimizes the expected performance measure or objective value such as total travel time, social welfare, and profit (e.g., Chen and Yang, 2004; Chen and Subprasom, 2007; Chow and Regan, 2011). The mean-variance model optimizes both the expected value and variance/standard deviation of performance measures using either the weighted sum or the Pareto optimal approach (e.g., Ng and Waller, 2009a; Sumalee et al, 2009). The chance-constraint model optimizes the design objective function subject to problem-specific chance constraints (e.g., Waller and Ziliaskopoulos, 2001; Chen and Yang, 2004; Li et al., 2014; Wang et al., 2015b). The probability model maximizes the probabilistic measures of network performance (e.g., Chootinan et al., 2005; Sumalee et al., 2006; Yim et al., 2011). The min-max model optimizes the worst-case performance (e.g., Yin et al. 2009). The alpha reliable model (e.g., Chen et al. 2007) is considered to be a variant of the chance-constrained model, because only the confidence level is user-defined. It is different from the chance-constrained model, in which both the confidence level α and an aspiration level of the performance measure are user-defined. This alpha reliable model can also be considered to be a relaxed min–max model in which the confidence level α is used to control the probability of the occurrence of acceptable network performance.

The above studies mainly consider demand uncertainty, which can be due to special events, population characteristics, etc. Some studies treat supply uncertainty as capacity degradation (due to weather conditions, traffic incidents, and road maintenance) and incorporate it into the design (Lo and Tung, 2003; Dimitriou and Stathopoulos, 2008; Li et al., 2008). However, these studies usually assume that the link capacity follows a specific distribution (e.g., uniform distribution) or the distribution is known, which may not always be true in some applications. Moreover, for some rare events such as earthquakes, there may not be enough data to calibrate the distribution. The network may change during the time that it would take to accumulate adequate data to estimate the distribution, making the estimated distribution obsolete. For these applications and events, alternative design approaches are required.

Ng and Waller (2009a) proposed a new convex mean variance model with an implicit chance constraint. This approach only assumes that the travel time distributions are symmetric. Unlike traditional mean-variance models, the model yields a one-sided confidence interval for the total system-wide travel time, which has an a priori specified confidence level. Despite its great contribution, their approach can be further improved for the following reasons.

1) Ng and Waller (2009a) assume that travelers select their route based on the shortest expected travel time. However, previous studies (Jackson and Jucker, 1982; Abdel-Aty et al., 1995) have shown that the reliability or variation of route travel time affects travelers' route choice. Their route choice depends on how they make a trade-off between (expected) travel cost and its uncertainty (Yin et al., 2004). They may reserve extra travel time to ensure they arrive at their destination on time (Lo et al., 2006). Ng and Waller (2009a), however, did not consider this risk-averse route choice behavior.

2) The chance constraint is loose given a one-sided confidence interval, which may give an inaccurate evaluation of the total system travel time reliability.

3) The parameters used to define the travel time variation for each link are identical, which may be too restrictive and fail to capture the variation accurately.

This study also considers networks with unknown link travel time distributions, but the distributions for different links are allowed to be heterogeneous as long as the mean and variance of the link travel time are known. A reliability-based system optimal design objective based on the total system travel time budget (TSTTB) is formulated. A non-negative parameter R, which reflects the risk aversion of the system manager and the tradeoff between the mean and standard deviation of the total system travel time (TSTT), is incorporated into the design objective. The reliability of the TSTT in the design is guaranteed by an implicit chance constraint, in the sense that the probability that the actual TSTT is within the optimal TSTTB is bounded below. To capture the selfish behavior and risk attitude of travelers, a reliability-based traffic assignment proposed by Lo et al. (2006) is captured in the model.

The major contributions of this study include the following:

- It proposes a new definition of the TSTTB for networks with heterogeneous distributed link travel times whose distributions are not known explicitly;
- It introduces a novel extension of network design problems that considers the risk attitude of travelers, risk-aversion behavior of the system manager, and networks with unknown link travel time distributions.
- It derives analytical results of the studied problem.

The key results of this study are summarized as follows:

- a higher risk aversion of system manager does not yield to a higher travel time variation and a lower mean TSTT at optimality;
- the derived lower bound for the probability that the actual TSTT is not greater than the optimal TSTTB is tighter than the bound in the literature under certain conditions and is proven to be related to the level of risk aversion of the system manager; the derived bound is proven to be the tightest under certain conditions;
- the derived lower bound for the probability that the actual path travel time is within the users' minimum path travel time budget under optimal design is proven to be related to the level of risk aversion of users, and is proven to be the tightest under certain conditions;
- the optimal TSTTB with link capacity expansions is proven to be bounded below even if the construction budget is unlimited;
- links with lower uncertainty may have a higher priority for improvement, and the variation in the TSTT might be smaller if travelers chose routes selfishly; and
- the risk-averse behavior of the system manager and the travelers may have opposite effects on link capacity expansions.

The paper is organized as follows. Section 2 defines link travel time uncertainties and formulates a bi-level reliability-based network design problem that minimizes the TSTTB. Section 3 proves that the probability that the actual TSTT is within the optimal TSTTB has a lower bound, which is tight under some conditions. Section 4 analyzes the design parameters and suggests useful properties. Finally, Section 5 presents the conclusion.

2. Formulation

Consider a transport network $G(N, A)$ with multiple origin-destination (O-D) pairs, where N is the set of nodes and A is the set of links. The travel demand of each class n for each O-D pair $rs \in RS$, $d_{rs,n}$, is given and fixed, where RS is the set of O-D pairs and traverses the paths between them. The set of paths between an O-D pair is denoted as P_{rs} and $P = \bigcup_{rs \in RS} P_{rs}$. Let m be the number of traveler classes. Define $M = \{1, 2, ..., m\}$. For clarity, the notations used in this paper are summarized in Table 2.1.

Table 2.1 Notations

$\mathbb{Z}_+$	Set of positive integers
v_a	Flow on link $a \in A$
$t_a(v_a)$	Travel time on link $a \in A$ given a link flow v_a
t_a^0	Free-flow travel time for link $a \in A$
c_a	Original capacity of link $a \in A$
C_a	Total capacity of link $a \in A$ after network improvement
α_a, β_a	Positive parameters in the Bureau of Public Road function of link $a \in A$
y_a	Capacity increment or expansion for link $a \in A$
ε_a	Non-negative function related to the maximum travel time error predicted by the travel time function of link $a \in A$
s_a	A positive parameter in ε_a
e_a	A positive coefficient in ε_a
H_a	A random variable associated with link $a \in A$ over $\left[-\kappa_a^L, \kappa_a^U \right]$ where κ_a^L and κ_a^U are the non-negative numbers not larger than one (i.e., $0 \leq \kappa_a^L \leq 1, 0 \leq \kappa_a^U \leq 1$) and one of them equals one.
Ω_a	Standard deviation of H_a
R	Parameter representing the level of system manager's risk-aversion, $R \geq 0$
B	Construction budget $B > 0$
$f_{p,n}$	Traffic flow of the n-th class travelers on path $p \in P$
$b_{p,n}$	Travel time budget of the n-th class travelers on path $p \in P$
$E(\cdot)$	Expected value of $\cdot$
λ_n	Parameter representing the level of users' risk-aversion
$\sigma(\cdot)$	Standard deviation of $\cdot$
δ_p^a	Binary indicator variable that equals 1 if link a is on path p,

	and 0 otherwise
$\pi_{rs,n}$	Minimum travel time budget of the n-th class travelers between O-D pair $rs \in RS$
u_a	Maximum allowable capacity increase of link $a \in A$

The Bureau of Public Road (BPR) function is generally used to predict the travel times for a link. The BPR function for link a can be expressed as

$$t_a(v_a) = t_a^0 \cdot \left(1 + \alpha_a \left(\frac{v_a}{c_a}\right)^{\beta_a}\right), \quad \forall a \in A, \tag{1}$$

where v_a is the link flow, t_a^0 is the free-flow travel time, c_a is the original capacity, and α_a and β_a are the positive coefficients of the BPR function for link a. For link expansion, the extra capacity, y_a, is directly added to the original capacity so that the total link capacity after expansion is $C_a = c_a + y_a$. Hence, the travel time after link expansion is

$$t_a(v_a, y_a) = t_a^0 \cdot \left(1 + \alpha_a \left(\frac{v_a}{c_a + y_a}\right)^{\beta_a}\right), \quad \forall a \in A. \tag{2}$$

2.1 Travel time uncertainty

Network supply uncertainty commonly refers to the variations in link capacities caused by disruptions to the network such as accidents, road maintenance, weather, and traffic management and control (Lo et al., 2006; Chen and Zhou, 2010). The classical approach to capturing supply uncertainty in a network model is to assume that each link capacity follows a certain distribution and that the link travel time distribution can be derived. Because a link capacity variation leads to a variation in the link travel time, an alternative approach is to focus directly on modeling travel time variations due to supply uncertainty (e.g., Ng and Waller, 2009a; Chen and Zhou, 2010; Szeto and Wang, 2015).

Under supply uncertainty, the actual link travel time may deviate from the predicted travel time (obtained from the BPR function). In this paper, the actual link travel time, T_a, $a \in A$, is modeled as a random variable based on Assumption 1:

Assumption 1.

$$T_a = t_a(v_a, y_a) \cdot [1 + \varepsilon_a \cdot H_a], \quad \forall a \in A, \tag{3}$$

where ε_a is a non-negative function; and H_a is a random variable independent of $t_a(v_a, y_a)$ and ε_a. The random variable H_a is distributed over $\left[-\kappa_a^L, \kappa_a^U\right]$ with a mean of zero, where either of the two non-negative numbers κ_a^L and κ_a^U equals one and the other is smaller than or equal to one. The variance of H_a is finite and denoted as Ω_a^2. The covariance between any two distinct random variables, $Cov[H_a, H_{a'}]$, is also finite.■

The two supports of the random travel time $T_a, a \in A$, are expressed by $t_a(v_a, y_a) \cdot \left[1 - \kappa_a^L \varepsilon_a\right]$ and $t_a(v_a, y_a) \cdot \left[1 + \varepsilon_a \kappa_a^U\right]$, because the range for the random variable H_a is $\left[-\kappa_a^L, \kappa_a^U\right]$. The

lower support $t_a(v_a, y_a) \cdot \left[1 - \kappa_a^L \varepsilon_a \right]$ is the travel time in the best case, and $\kappa_a^L \varepsilon_a$ is the maximum relative error below the mean travel time. The upper support $t_a(v_a, y_a) \cdot \left[1 + \varepsilon_a \kappa_a^U \right]$ is the travel time in the worst case, and $\kappa_a^U \varepsilon_a$ is the maximum relative error above the mean travel time. A larger value of the two terms, $\varepsilon_a \kappa_a^L$ and $\varepsilon_a \kappa_a^U$, represents the maximum relative error predicted by the BPR function. Because either κ_a^L or κ_a^U equals one and the other is smaller than or equal to one, the maximum relative error predicted by the BPR function is simply reflected by the parameter ε_a. Moreover, because the actual travel time is larger than zero, the value of $\left[1 - \kappa_a^L \varepsilon_a \right]$ should be larger than zero. Hence, $0 \leq \kappa_a^L \varepsilon_a < 1$.

The random variable H_a, $a \in A$, in (3) reflects the distribution of T_a. In this study, H_a may not be identical for each link, meaning that the distributions of random travel times for different links may be heterogeneous. The standard deviation of H_a, Ω_a reflects the dispersion of the random variable value around the mean. By definition, the range of Ω_a is $[0,1]$ (see Appendix A for the properties of Ω_a). A larger Ω_a indicates that the dispersion is larger, and vice versa. The covariance between any two distinct random variables H_a and $H_{a'}$, $Cov[H_a, H_{a'}]$, reflects their correlation. Ng and Waller (2009a) introduced a similar assumption, but with the restriction that the link travel time distribution is symmetric, the value of Ω_a must be identical for all links, $\kappa_a^L = \kappa_a^U = 1$, and travel time covariances are zero.

The travel time uncertainty is simultaneously reflected by ε_a and Ω_a. A larger value of ε_a or a larger value of Ω_a indicates that the travel time is more uncertain and the worst-case travel time is longer, and vice versa. In a transportation network, the value of Ω_a is calibrated or estimated based on empirical travel time data. The function ε_a is assumed to be affected by two factors: 1) non-congestion-related uncertainty caused by, for example, bad weather, road maintenance, or capacity degradation; and 2) congestion-related uncertainty (Ng and Waller, 2009a). The non-congestion-related uncertainty is expressed by a non-negative constant s_a, whose value is calibrated based on empirical travel time data. The congestion-related uncertainty is a function of road congestion, $h_a(v_a/C_a)$, which is an increasing function of the degree of congestion. Without travel time data, the specific form of $h_a(v_a/C_a)$ cannot be determined empirically and hence is not provided in this study. For illustration purposes, we set $h_a(v_a/C_a) = e_a v_a/C_a$ (Ng and Waller, 2009a), in which e_a is a non-negative coefficient. Thus,

$$\varepsilon_a = s_a + e_a \cdot \frac{v_a}{C_a}, \ \forall a \in A . \tag{4}$$

Based on Assumption 1, the expectation and variance of the link travel time are

$$E[T_a] = t_a(v_a) + \varepsilon_a(v_a) t_a(v_a) \cdot E[H_a] = t_a(v_a), \text{ and} \tag{5}$$

$$Var[T_a] = (t_a \varepsilon_a)^2 Var[H_a] = t_a^2 \varepsilon_a^2 \Omega_a^2, \ \forall a \in A . \tag{6}$$

The travel time covariance between any two links is

$$Cov[T_a, T_{a'}] = t_a \varepsilon_a \cdot t_{a'} \varepsilon_{a'} \cdot Cov[H_a, H_{a'}], \ \forall a, a' \in A . \tag{7}$$

The total system travel time (TSTT) is defined as the sum of all link travel times. Because the link travel times are all random variables, the TSTT is a compound random variable which is expressed by

$$\widehat{TSTT} = \sum_{a \in A} T_a v_a = \sum_{a \in A} t_a(v_a, y_a) \cdot [1 + \varepsilon_a \cdot H_a] \cdot v_a \,. \tag{8}$$

The expectation and variance of the TSTT are

$$E\left[\widehat{TSTT}\right] = \sum_{a \in A} t_a v_a \,, \text{ and} \tag{9}$$

$$Var\left[\widehat{TSTT}\right] = \sum_{a \in A} \left(\varepsilon_a t_a v_a \Omega_a\right)^2 + \sum_{a,a' \in A, a \neq a'} \left(\varepsilon_a t_a v_a\right)\left(\varepsilon_{a'} t_{a'} v_{a'}\right) Cov\left[H_a, H_{a'}\right] \,. \tag{10}$$

2.2 Design objective

From the system manager's point of view, the objective is to maximize the system performance, which means minimizing the TSTT when there is no travel time uncertainty. However, for a network with supply uncertainty, it is essential that the system manager is risk averse to account for the TSTT variation and to make an appropriate tradeoff between the mean performance and performance variation. The TSTT variation may be large, such that the mean TSTT is much smaller than the actual TSTT and does not reflect the actual, very bad performance. To address this issue, Szeto and Wang (2015) introduced the concept of TSTTB which is defined as

TSTTB = [Expected TSTT] + [Safety margin].

To model the safety margin in the above equation, there are at least two options. The first option is to model it as the standard deviation or weighted standard deviation of the TSTT. The second is to model it as the variance of the TSTT, which is generally less preferable because the variance and expectation use different units. We adopt the first option here, specifically the weighted version, to account for the risk aversion of the system manager. For this purpose, a non-negative factor R is multiplied by the standard deviation to model the safety margin. A larger R leads to a larger safety margin and indicates that the system manager is more risk averse.

The definition of the TSTTB is consistent with the definition of the total travel time budget proposed by Chen et al. (2007). It is simply formulated using a different approach. Chen et al. (2007) defined the total travel time budget using the chance constraint approach without explicitly using the mean and variance of the TSTT, both of which are used to define the TSTTB in this paper. Their study needs the distribution of travel time to define the confidence level or reliability guarantee. In our study, we do not provide the confidence level corresponding to a one-sided confidence interval, because the distribution of travel time is not known explicitly. Instead, we provide a lower bound of the confidence level. The formulation approach taken in this paper has the following advantages. First, the concept of the TSTTB is easy to understand as its mathematical structure is similar to that of the travel time budget. Second, this paper lends novel insights into the TSTTB, including an interpretation of the R value, the concept of the system manager's safety margin, and the concept of the largest acceptable TSTT.

Based on the preceding discussion, we define the TSTTB as

$$TSTTB = \sum_{a \in A} t_a v_a + R \sqrt{\sum_{a \in A} \left(\varepsilon_a t_a v_a \Omega_a\right)^2 + \sum_{a,a' \in A, a \neq a'} \left(\varepsilon_a t_a v_a\right)\left(\varepsilon_{a'} t_{a'} v_{a'}\right) Cov\left[H_a, H_{a'}\right]} \,. \tag{11}$$

Then, the design objective is to minimize the TSTTB, i.e.,

$$\min TSTTB \,. \tag{12}$$

2.3 Reliability-based user equilibrium constraints

A reliable network design should consider both the selfish behavior and the risk attitude of travelers. *This study assumes that travelers are selfish but may have different attitudes towards risk. The travelers are divided into $m \in \mathbb{Z}_+$ classes based on their levels of risk aversion/seeking. The travelers gain information on travel time expectations and variations based on their past experiences and factor them into their route-choice considerations, and settle into a long-term habitual equilibrium pattern.* To formulate the equilibrium, we adopt the path travel time budget proposed by Lo et al. (2006) and adopted by others (e.g., Shao et al., 2006; Siu and Lo, 2008), which is defined as:

Path travel time budget = [Expected value of path travel time] + [Safety margin].

Note that Lo et al. (2006) considered that the path travel times follow normal distributions. This assumption is relaxed in this study. By definition, the path travel time budget is

$$b_{p,n} = \sum_{a \in A} t_a \delta_p^a + \lambda_n \sqrt{\sum_{a \in A} \left(\varepsilon_a t_a \Omega_a\right)^2 \delta_p^a + \sum_{a,a' \in A, a \neq a'} \left(\varepsilon_a t_a\right)\left(\varepsilon_{a'} t_{a'}\right) Cov\left[H_a, H_{a'}\right] \delta_p^a \delta_p^{a'}}, \tag{13}$$

$$\forall p \in P, \forall n \in M.$$

The factor λ_n reflects the level of risk aversion of the *n*-th class travelers: the larger the value, the greater the risk aversion. Moreover, a positive value of λ_n means that the *n*-th class travelers are risk averse; a zero value of λ_n means that the travelers are risk neutral; and a negative value of λ_n means that the travelers are risk seeking.

Travelers are assumed to select the route that minimizes their own path travel time budget. At equilibrium, all travelers with the same risk attitude between the same O-D pair have the same minimum travel time budget. Mathematically, the long-term habitual equilibrium condition can be formulated as follows:

$$f_{p,n}\left(b_{p,n} - \pi_{rs,n}\right) = 0, \ \forall p \in P_{rs}, \forall rs \in RS, \forall n \in M, \text{ and} \tag{14}$$

$$\left(b_{p,n} - \pi_{rs,n}\right) \geq 0, \ \forall p \in P_{rs}, \forall rs \in RS, \forall n \in M. \tag{15}$$

2.4 Budget and capacity constraints

The most popular approach to formulating the link capacity expansion constraints is to set an upper bound on the total construction cost, because the government usually provides a fixed budget, denoted as B, for link improvements. This constraint can be expressed as

$$\sum_{a \in A} \hat{F}(y_a) \leq B, \tag{16}$$

where $\hat{F}(y_a)$ represents the construction cost spent on link *a*.

The construction cost for each link is assumed to be a linear function of the capacity improvement, i.e.,

$$\hat{F}(y_a) = M_a \cdot y_a, \ \forall a \in A, \tag{17}$$

where M_a is the construction cost per unit capacity increase.

To reflect the situation in which there is no more space to further widen the link, the link capacity expansion is bounded above by u_a, i.e.,

$$0 \leq y_a \leq u_a, \ \forall a \in A. \tag{18}$$

2.5 Traffic assignment constraints

The constraints include path flow-link flow conservation, demand-path flow conservation, and the non-negativity constraints:

$$\sum_{n=1}^{m} \sum_{rs \in RS} \sum_{p \in P_{rs}} f_{p,n} \delta_p^a = v_a, \quad \forall a \in A, \tag{19}$$

$$\sum_{p \in P_{rs}} f_{p,n} = d_{rs,n}, \quad \forall rs \in RS, \forall n \in M, \text{ and} \tag{20}$$

$$f_{p,n} \geq 0, \quad \forall p \in P_{rs}, \forall rs \in RS, \forall n \in M. \tag{21}$$

2.6 The model and its properties

Denote $\mathbf{y} = (y_a)_{a \in A}$ as a feasible link capacity expansion vector that satisfies constraints (16) and (18). Denote $\mathbf{v}^{RUE}(\mathbf{y}) = (v_a^{RUE})_{a \in A}$ as a RUE link flow vector for a given link capacity expansion vector. The reliable network design problem involves determining the link capacity expansion $\mathbf{y}$ and the corresponding RUE link flow pattern $\mathbf{v}^{RUE}(\mathbf{y})$ to minimize the TSTTB, subject to reliability-based UE constraints (14)-(15), budget and capacity expansion constraints (16)-(18), and traffic assignment constraints (19)-(21).

The design problem can be viewed as a bi-level optimization problem. The lower level problem is to find the link flow pattern that satisfies the multi-class RUE constraints and traffic assignment constraints. The upper level problem is to find the optimal link expansion solution that minimizes the TSTTB subject to budget and capacity constraints.

Some basic properties of the design problem are summarized as follows.

- The design objective function is continuous and differentiable in terms of link flows and capacities. The feasible solution set is a continuous set.
- The link travel time is convex in terms of its link flow. Thus, the mean TSTT is convex in terms of link flows. The standard deviation of the TSTT is also convex with respect to link flows because ε_a, $\forall a \in A$, is a convex function of v_a (see expression (4)).
- The second derivative of the travel time function with respect to its link capacity is $\beta_a(\beta_a + 1)t_a^0 \alpha_a \left(v_a / (c_a + y_a) \right)^{\beta_a}$, which is always positive. Thus, the link travel time is a convex function of its link capacity. Therefore, ε_a, $\forall a \in A$, is a convex function of its link capacity. As a result, the mean and standard deviation of the TSTT are both convex in terms of link capacities and hence the link capacity expansion $\mathbf{y}$.
- The optimal link capacity expansions and corresponding link flow pattern might not be unique, because the solution set is non-convex.

Proposition 1. *The optimal TSTTB is non-decreasing in terms of R. If all link travel times are stochastic, the optimal TSTTB is strictly increasing in terms of R.*

Proof. Consider two cases, $R = R'$ and $R = R''$, in which $R' < R''$. The lower-level constraints remain the same in these two cases, indicating that the feasible regions of link capacity

expansions and link flows are not altered by the value of R. The objective function, however, is affected. For a pair of *feasible* link capacity expansions and link flows, $\left(\mathbf{y}, \mathbf{v}^{RUE}(\mathbf{y})\right)$, denote $TSTTB_{R=R'}\left(\mathbf{y}, \mathbf{v}^{RUE}(\mathbf{y})\right)$ as the TSTTB when $R=R'$, and $TSTTB_{R=R''}\left(\mathbf{y}, \mathbf{v}^{RUE}(\mathbf{y})\right)$ as the TSTTB when $R=R''$. The following always holds:

$$TSTTB_{R=R'}\left(\mathbf{y}, \mathbf{v}^{RUE}(\mathbf{y})\right) \leq TSTTB_{R=R''}\left(\mathbf{y}, \mathbf{v}^{RUE}(\mathbf{y})\right), \tag{22}$$

because the non-negative safety margin in the TSTTB for the case of $R=R''$ is larger than or equal to that for the case of $R=R'$, and the mean TSTTs for the two cases are the same. The equality sign holds only when the standard deviation of TSTT is zero (which is the case when all used links have no travel time uncertainty). Denote $\left(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')\right)$ and $\left(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')\right)$ as the optimal link capacity expansions and link flows for $R=R'$ and $R=R''$, respectively. Inequality (22) also holds for $\left(\mathbf{y}, \mathbf{v}^{RUE}(\mathbf{y})\right) = \left(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')\right)$, and hence we have

$$TSTTB_{R=R'}\left(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')\right) \leq TSTTB_{R=R''}\left(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')\right). \tag{23}$$

Because $\left(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')\right)$ is the optimal solution for the case of $R=R'$, other feasible solutions for this case, including $\left(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')\right)$, cannot give a TSTTB that is smaller than the optimal TSTTB for the case of $R=R'$. Hence, we have

$$TSTTB_{R=R'}\left(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')\right) \leq TSTTB_{R=R'}\left(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')\right). \tag{24}$$

Combining inequalities (23) and (24), we have

$$TSTTB_{R=R'}\left(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')\right) \leq TSTTB_{R=R''}\left(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')\right).$$

Moreover, if *all link travel times are stochastic*, the standard deviation of TSTT must be positive for any given feasible link flows and capacity expansions. Then, inequality (23) is strict, i.e., $TSTTB_{R=R'}\left(\mathbf{y}, \mathbf{v}^{RUE}(\mathbf{y})\right) < TSTTB_{R=R''}\left(\mathbf{y}, \mathbf{v}^{RUE}(\mathbf{y})\right)$. Together with inequality (24), the following is obtained: $TSTTB_{R=R'}\left(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')\right) < TSTTB_{R=R''}\left(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')\right)$.

This completes the proof.∎

Remark: Proposition 1 implies *that the optimal TSTTB is the smallest if and only if R equals zero.* However, when R equals zero, the standard deviation of the TSTT at optimality is the largest, as implied by Proposition 2.

Proposition 2. *As R gets larger, the mean of the TSTT at optimality is non-decreasing, while the standard deviation of the TSTT is non-increasing.*

Proof. See **Appendix B**.

Remark 1: Proposition 2 indicates that when R is larger, the standard deviation of the TSTT at the optimal TSTTB is always better (or at least not worse). However, when R tends to infinity, the standard deviation of the TSTT at optimality does not tend to zero because the travel demand

is non-zero, at least one link is used by travelers (i.e., the link flow is non-zero), and the congestion-related link travel time variance is larger than zero. Thus, the standard deviation of the TSTT must be larger than zero.

Remark 2: When R tends to infinity, the mean TSTT does not tend to infinity because the travel demands are fixed and the link flows are bounded. The mean link travel times, which are obtained by BPR-type functions based on link flows, are also bounded. Thus, the mean TSTT is also bounded.

Proposition 3. *The optimal TSTTB is bounded below, given any construction budget.*

Proof. Consider an extreme case in which the budget is unlimited. All links will receive infinite large capacity improvements. Recall that the link capacity C_a , $a \in A$, affects the degree of uncertainty ε_a and the mean travel time t_a. As C_a reaches infinity, the factor ε_a equals s_a (the congestion-related uncertainty is reduced to 0), and t_a equals t_a^0. Even though the link capacities are infinitely large, travelers still have to bear the minimum travel costs (the free-flow travel times) and the non-congestion-related uncertainties always exist. Thus, the optimal TSTTB is bounded below.∎

Remark: The value of the lower bound is apparently affected by the free-flow travel times and the non-congestion-related travel time uncertainties. A feasible lower bound is the optimal TSTTB for the following minimization program: $\min_{\mathbf{v}} \sum_{a \in A} t_a^0 v_a$, subject to the demand conservation constraint, link flow-path flow conservation constraint, and non-negativity constraints. This program calculates the minimum TSTTB when no congestion effect exists and travelers are perfectly cooperative. The lower bound of the optimal TSTTB is an estimate of the system performance under perfect link capacity expansion (all link capacities are improved to infinity), and its existence reveals that the system performance cannot be infinitely improved.

2.7 Solution method

The upper level problem or the whole design problem can be solved by traditional metaheuristics, such genetic algorithm and simulated annealing, or derivative-based nonlinear optimization algorithms such as generalized reduced gradient algorithm, sequential quadratic programming algorithm, and augmented Lagrangian method. A comprehensive review on the solution methods for continuous and other network design problems is given by Farahani et al. (2013). Derivative-based nonlinear optimization algorithms are efficient but they may not be able to obtain global optimal solutions to the proposed problem because of non-convex solution sets. Directly applying the above metaheuristics to solve the proposed problem may not be able to obtain good solutions quickly because the nice properties of the problem, such as differentiability, have not been used. Further research on developing efficient and accuracy solution methods is required.

The lower level problem, which is to solve the constraints (14) and (15) given a feasible link expansion solution $\mathbf{y}$, can be solved by the following variational inequality problem: to find a path flow pattern $\mathbf{f}^{RUE}(\mathbf{y}) = \left(f_{p,n}^{RUE} \right)_{p \in P, n \in M}$, such that for any feasible path flow pattern

$\mathbf{f}'(\mathbf{y}) = \left(f'_{p,n} \right)_{p \in P, n \in M}$, the following holds: $\displaystyle \sum_{n=1}^{m} \sum_{p \in P} \left(f'_{p,n} - f_{p,n}^{RUE} \right) b_{p,n}(\mathbf{f}^{RUE}(\mathbf{y})) \geq 0, \mathbf{f}'(\mathbf{y}) \in \Psi$, where

Ψ is the set of feasible path flow patterns. Efficient solution algorithms which solve the above VI problem are introduced in, for example, the study of Chen et al. (2011), which is based on Dial (2006)'s flow swapping concept between the largest travel time budget path and the minimum travel time budget path, O-D decomposition, and column generation. Existing projection algorithms developed for variational inequalities or fixed point problems (see e.g., Han et al. 2015) can also be used to solve for solutions to the lower level problem, assuming the convergence criterion is satisfied. When solving the VI problem, it is always required to solve a reliable shortest travel time sub-problem. Chen et al. (2012; 2013) presented reviews for this subproblem and proposed algorithms for solving it.

3. The probabilistic guarantees

3.1 The actual TSTT

In order to quantify the reliability of the design, it is necessary to evaluate the probability that the actual TSTT is within the optimal TSTTB. Because the travel time distributions are unknown, it is difficult to express the probability explicitly, or to calculate it. Nevertheless, it is possible to derive a lower bound for the probability (i.e., the level of probabilistic guarantee on the actual TSTT) for the purpose of judging the reliability of the design. By probabilistic guarantee, we mean a performance guarantee associated with a probability. A guarantee of 90% means that the probability that the performance meets the required standard is at least equal to 90%. In this section, the performance measure is the actual TSTT.

For the ease of presentation, we denote the TSTTB at $(\mathbf{v}^{RUE}(\mathbf{y}^*), \mathbf{y}^*)$ as $TSTTB^*$; the corresponding random travel time on link a as T_a^*, the corresponding mean link travel time under optimal design $t_a(v_a^{RUE}, y_a^*)$ as t_a^*, and the corresponding ε_a value as ε_a^*. Note that $y_a^* = 0$ for all links also represent the case before link capacity expansion.

Because the distribution of the TSTT is not known explicitly, we discuss two cases separately. One case is that TSTT follows a symmetric distribution. Another case is that TSTT follows asymmetric (or general) distribution.

Theorem 1. *The probability that the actual TSTT is within the optimal TSTTB has a sharp lower bound of 50% for $0 \leq R < 1$ if TSTT follows a symmetric distribution.*

Proof. Denote the standard deviation of TSTT at $(\mathbf{y}, \mathbf{v}(\mathbf{y}))$ as $F(\mathbf{v}(\mathbf{y}), \mathbf{y})$. Because the TSTTB is larger than or equal to the mean TSTT at optimality, the following must hold:

$$\Pr\left(\sum_{a \in A} T_a^* v_a^{RUE} \leq \sum_{a \in A} t_a^* v_a^{RUE} + R \cdot F(\mathbf{v}^{RUE}(\mathbf{y}^*), \mathbf{y}^*) \right) \geq \Pr\left(\sum_{a \in A} T_a^* v_a^{RUE} \leq \sum_{a \in A} t_a^* v_a^{RUE} \right). \qquad (25)$$

The right side of (25) is the probability of a random variable being smaller than or equal to its mean. Because TSTT follows a symmetric distribution, the right side of (25) is equal to 50% if TSTT is a continuous random variable and is larger than or equal to 50% if TSTT is a discrete random variable. Thus, the left side of (25) is larger than or equal to 50%.

To prove that 50% is the tightest lower bound of the probability on the left side of inequality (25), we just need to find a case with explicit travel time distributions in which the probability of $\sum_{a\in A} T_a^* v_a^{RUE} \leq \sum_{a\in A} t_a^* v_a^{RUE} + R \cdot F(\mathbf{v}^{RUE}(\mathbf{y}^*), \mathbf{y}^*)$ is indeed equal to 50% for $0 \leq R < 1$. Consider the following case.

A network with two nodes, A and B, is connected by two parallel links, denoted as Links 1 and 2. The demand from A to B is $d = 5$. Given optimal link expansion, the predicted link travel time is $t_1(v_1) = 0.1 + 0.1 v_1$ and $t_2(v_2) = 1 + 0.6 v_2$. The functions ε_1 and ε_2 are constant functions and equal 0.1 and 0.5, respectively (i.e., $e_1 = e_2 = 0$; $s_1 = 0.1$; $s_2 = 0.5$). The travel times of two links are independent. The probability mass function (PMF) of the random variable H_1, $z_{H_1}(\eta_1)$, is

$$z_{H_1}(\eta_1) = \begin{cases} 0.5 & \text{if } \eta_1 = 1; \\ 0.5 & \text{if } \eta_1 = -1; \\ 0 & \text{otherwise.} \end{cases}$$

The standard deviation Ω_1 is 1 (computed based on the PMF).

- As Link 1 has a much lower travel time and variance, all users d take Link 1 to travel from A to B, i.e., $v_1^{RUE} = d$, and $v_2^{RUE} = 0$. The optimal TSTTB is

$$\sum_{a\in A} t_a^* v_a^{RUE} + R \cdot F(\mathbf{v}^{RUE}(\mathbf{y}^*), \mathbf{y}^*) = t_1(d)d + R\sqrt{(t_1(d)d\varepsilon_1\Omega_1)^2} \ .$$

- The actual TSTT is $\sum_{a\in A} T_a^* v_a^{RUE} = T_1^* d$. The PMF of $\sum_{a\in A} T_a^* v_a^{RUE}$ is identical to that of $T_1^* d$. Recall that $T_1^* = t_1(d) + t_1(d)\varepsilon_1\eta_1$ (see (3)). The mean of $T_1^* d$ is $t_1(d)d$. According to the PMF of H_1, the PMF of $T_1^* d$, $z_{T_1^* d}$, is

$$z_{T_1^* d}(\omega) = \begin{cases} 0.5, & \text{if } \omega = t_1(d)d - t_1(d)d\varepsilon_1, \\ 0.5, & \text{if } \omega = t_1(d)d + t_1(d)d\varepsilon_1, \\ 0, & \text{otherwise.} \end{cases}$$

When $0 \leq R < 1$, the probability that the $T_1^* \cdot d$ is within $t_1(d)d + R \cdot \sqrt{(t_1(d)d\varepsilon_1\Omega_1)^2}$ is

$$\Pr\left(T_1^* \cdot d \leq t_1(d)d + R \cdot \sqrt{(t_1(d)d\varepsilon_1\Omega_1)^2}\right) = \Pr\left(T_1^* \cdot d \leq t_1(d)d + R \cdot t_1(d)d\varepsilon_1\right)$$

$$= \Pr\left(T_1^* \cdot d < t_1(d)d + t_1(d)d\varepsilon_1\right)$$

$$= 0.5.$$

As $\sum_{a\in A} T_a^* v_a^{RUE} = T_1^* d$, we have

$$\Pr\left(\sum_{a\in A}T_a^*v_a^{RUE} \leq \sum_{a\in A}t_a^*v_a^{RUE} + R\cdot F(\mathbf{v}^{RUE}(\mathbf{y}^*),\mathbf{y}^*)\right)$$

$$= \Pr\left(T_1^*\cdot d \leq t_1(d)d + R\cdot\sqrt{\left(t_1(d)d\varepsilon_1\Omega_1\right)^2}\right) = 0.5.$$

This case indicates that $\Pr\left(\sum_{a\in A}T_a^*v_a^{RUE} \leq \sum_{a\in A}t_a^*v_a^{RUE} + R\cdot F(\mathbf{v}^{RUE}(\mathbf{y}^*),\mathbf{y}^*)\right)$ can indeed be as small as 50% for $0\leq R<1$, meaning that 50% is the tightest lower bound.∎

<u>Remark 1:</u> The lower bound for the probability that the actual TSTT is within the optimal TSTTB remains unchanged for $0\leq R<1$, and yet the bound is indeed the tightest according to Theorem 1.

<u>Remark 2:</u> The lower bound of the probability can be improved given different values of *R that are larger than or equal to 1*. First, the following lemma is introduced.

Lemma 1. *The probability that the actual TSTT is <u>larger than or equal to</u> the optimal TSTTB has a sharp upper bound for $R\geq 1$ if TSTT follows a symmetric distribution:*

$$\Pr\left(\sum_{a\in A}T_a^*v_a^{RUE} \geq TSTTB^*\right) \leq \frac{1}{2R^2}. \tag{26}$$

Proof. The proof of inequality (26) is given in part (a) of **Appendix C** and the tightness of the upper bound is proven in part (b) of **Appendix C**.

Theorem 2. *The probability that the actual TSTT is within the optimal TSTTB has a lower bound of $1-\dfrac{1}{2R^2}$ for $R\geq 1$ if TSTT follows a symmetric distribution*

$$\Pr\left(\sum_{a\in A}T_a^*v_a^{RUE} \leq TSTTB^*\right) \geq 1-\frac{1}{2R^2}. \tag{27}$$

Proof. Denote the standard deviation of TSTT at $(\mathbf{y},\mathbf{v}(\mathbf{y}))$ as $F(\mathbf{v}(\mathbf{y}),\mathbf{y})$. The complementary event of $\sum_{a\in A}T_a^*v_a^{RUE} \geq TSTTB^*$ is $\sum_{a\in A}T_a^*v_a^{RUE} < TSTTB^*$, the probability of which is also bounded based on Lemma 1. Following the result in Lemma 1 and replacing optimal TSTTB with its expression, we have

$$\Pr\left(\sum_{a\in A}T_a^*v_a^{RUE} < \sum_{a\in A}t_a^*v_a^{RUE} + R\cdot F(\mathbf{v}^{RUE}(\mathbf{y}^*),\mathbf{y}^*)\right) \geq 1-\frac{1}{2R^2}. \tag{28}$$

Clearly, the probability of $\sum_{a\in A}T_a^*v_a^{RUE} \leq \sum_{a\in A}t_a^*v_a^{RUE} + R\cdot F(\mathbf{v}^{RUE}(\mathbf{y}^*),\mathbf{y}^*)$ is larger than or equal to that of $\sum_{a\in A}T_a^*v_a^{RUE} < \sum_{a\in A}t_a^*v_a^{RUE} + R\cdot F(\mathbf{v}^{RUE}(\mathbf{y}^*),\mathbf{y}^*)$. Thus,

$$\Pr\left(\sum_{a \in A} T_a^* v_a^{RUE} \le \sum_{a \in A} t_a^* v_a^{RUE} + R \cdot F(\mathbf{v}^{RUE}(\mathbf{y}^*), \mathbf{y}^*)\right)$$

$$\ge \Pr\left(\sum_{a \in A} T_a^* v_a^{RUE} < \sum_{a \in A} t_a^* v_a^{RUE} + R \cdot F(\mathbf{v}^{RUE}(\mathbf{y}^*), \mathbf{y}^*)\right) \ge 1 - \frac{1}{2R^2}. \quad (29)$$

This completes the proof.∎

Lemma 2. *The probability that the actual TSTT is <u>larger than or equal to</u> the optimal TSTTB has a sharp upper bound for $R > 1$, if TSTT follows an asymmetric distribution:*

$$\Pr(\sum_{a \in A} T_a^* v_a^{RUE} \ge TSTTB^*) \le \frac{1}{R^2}. \quad (30)$$

For $0 \le R \le 1$, the upper bound is simply one.

Proof. The proof of inequality (30) is given in part (a) of **Appendix C**. Condition (30) is only meaningful when R is larger than one. When R is smaller than or equal to one, the upper bound is larger than or equal to one. However, the probability is at most one by definition. Thus, the upper bound of the left side of (30) equals one for $0 \le R \le 1$.

Theorem 3. *If TSTT follows an asymmetric distribution, the probability that the actual TSTT is within the optimal TSTTB has a lower bound:*

$$\Pr\left(\sum_{a \in A} T_a^* v_a^{RUE} \le TSTTB^*\right) \ge \begin{cases} 0, & 0 \le R \le 1, \\ 1 - \dfrac{1}{R^2}, & R > 1. \end{cases} \quad (31)$$

Proof. It follows the results in Lemma 2 and the technique used in proving Theorem 2.

<u>Remark 1</u>: The numerical values of the lower bound of the probability are listed in Table 3.1.

Table 3.1 Lower bound of the probability

		Value			
	R	0~1	1.5	2	3
Symmetric distribution	Probabilistic guarantee	50%	77.8%	87.5%	94.4%
Asymmetric distribution		0	55.6%	75.0%	88.9%

<u>Remark 2</u>: In the literature, Ng and Waller (2009a) proposed a design model based on unknown and symmetrical travel time distributions, and derived a lower bound of the probability that the actual TSTT is within the optimal objective function value. Their design objective function is indeed a special case of the TSTTB in our proposed model under the following conditions: $R = 1$ and Ω_a, $a \in A$, is identical for every link. (i.e., $\Omega_a = \Omega$.) The lower bound derived by Ng and Waller (2009a) is $\left(1 - \exp\left(-\Omega^2/2\right)\right)$. Because Ω is bounded above by one (see Appendix A), their

lower bound is at most 39%, which can be obtained by setting $\Omega = 1$. Under the above mentioned conditions, the lower bound of the probability that the actual TSTT is within the optimal TSTTB derived in our study equals 50% (because $R = 1$). (Note that the consideration of RUE constraints has no impact on the randomness of the TSTT.) The lower bound of the probability derived in this study is thus tighter.

Remark 3: The probability lower bound in Theorem 2 can certainly be improved if further information regarding the distribution of the actual TSTT is given. For example, if the actual TSTT follows a normal distribution, then the probability is 95% for $R = 1.65$.

Corollary 1. *The lower bound of the probability that the actual TSTT is within the optimal TSTTB is 1) a constant for $0 \le R < 1$, and 2) strictly increasing with respect to R for $R \ge 1$.*

Proof. According to Theorems 1 and 3, the lower bound remains the same (which is 50% for a symmetric distribution and zero for an asymmetric distribution) for $0 \le R < 1$. According to Theorems 2 and 3, for $R \ge 1$, the lower bound equals $1 - \dfrac{1}{2R^2}$ for a symmetric distribution and $1 - \dfrac{1}{R^2}$ for an asymmetric distribution, which are both strictly increasing with respect to R.∎

3.2 The users' actual path travel time

Theorem 4. *Given optimal link expansion, for the n-th class travelers, the probability of their travel time being within the minimum path travel time budget is bounded below by*

1) $1 - \dfrac{1}{2\lambda_n^2}$ for $\lambda_n \ge 1$ and 50% for $0 \le \lambda_n < 1$ if the path travel time follows a symmetric distribution;

2) $1 - \dfrac{1}{\lambda_n^2}$ for $\lambda_n \ge 1$ and 0 for $0 \le \lambda_n < 1$ if the path travel time follows an asymmetric distribution.

Proof. The proof follows similar procedures to the proofs for Theorems 1, 2, and 3.

Remark: Theorem 4 is applicable when the level of risk aversion of travelers is non-negative. The case when the level of risk aversion of travelers is negative is left for future study. Similar to the results in Section 3.1, this lower bound of the probability (i.e., the level of the probabilistic guarantee) can be improved if further information regarding the actual travel time distribution is given. For example, if the actual path travel time follows a normal distribution, the probabilistic guarantee reaches 95% when λ_n equals 1.65 (see Lo et al., 2006).

Following the results in Theorem 4, the property of the lower bound is stated in Corollary 2.

Corollary 2. *The lower bound of the probability of travelers' path travel time not greater than the minimum path travel time budget under optimal design is 1) a constant if $0 \le \lambda_n < 1$, and 2) strictly increasing with respect to λ_n if $\lambda_n \ge 1$.*

4. Numerical study

In this section, numerical examples are introduced to enhance the understanding of the design model. The numerical examples are based on a three-link network, as shown in Figure 2. The basic parameter settings are given in Table 4.1. There is only one class of travelers and one O-D pair (A, C). The demand is $d_{AC,1} = 5$. The design budget is $B = 5$. The travel times of all links are independent. All the results are obtained using SOLVER in EXCEL2013.
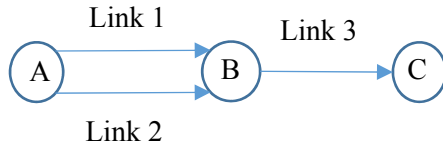
Figure 2. Network Geometry

Table 4.1 Network Parameters

Link	(tail, head)	t_a^0	c_a	α_a	β_a	M_a
1	(A,B)	5	3.0	5	4	1.2
2	(A,B)	6	3.5	8	4	1.3
3	(B,C)	10	5.0	8	4	1.3

Example 1. The objective is to demonstrate the effect of travel time variations on the design. Ω_1 varies between 0 and 1. The following parameter values are set: $R = 4$; $\lambda_1 = 2$; $s_1 = 0.5$; $s_2 = 0.1$; $s_3 = 0.1$; $e_a = 0$, $\forall a \in A$, $\Omega_2 = 0.4$; and $\Omega_3 = 0.4$. The link capacity expansions for the three links are plotted in Figure 3.
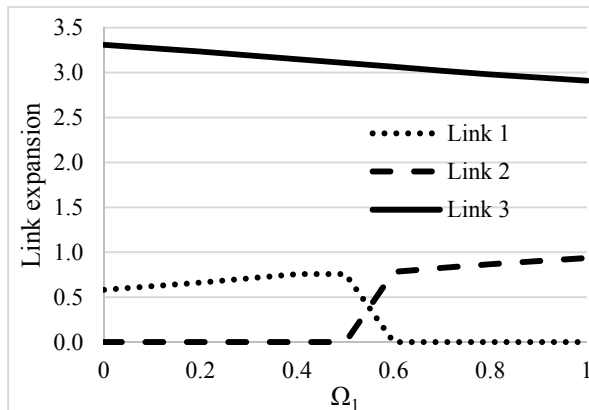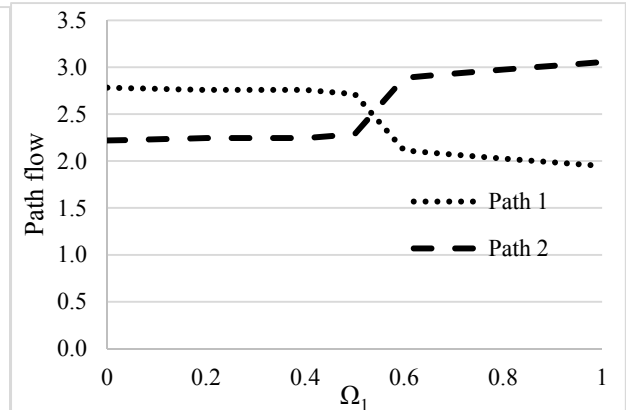

Figure 3(a). Link capacity expansions

Figure 3(b). Path flows

Observation 1. *Links with higher travel time uncertainty may not be expanded more.*

In networks with uncertainty, we generally expect links with high travel time uncertainty (i.e., more unreliable links) to be prioritized in terms of receiving more resources (i.e., higher budget). In this example, Link 3 always receives more resources than the other two links. However, Link 3 does not have a higher travel time variation than the other two links (especially compared with Link 1 when Ω_1 is large). A possible explanation for this result is that all paths pass through Link 3, making it a critical link because all travelers use it. In this case, although Link 3 is less uncertain, it still receives more resources. This result suggests that a link with higher travel time uncertainty might not have a higher priority for improvement.

The comparison between the resources spent on Links 1 and 2 is also interesting. Path 1 consists of Links 1 and 3, and Path 2 consists of Links 2 and 3. Link 3 is therefore common to both paths, and Links 1 and 2 are competing links. Three regions can be observed: 1) when Ω_1 lies between $[0,0.5]$, Link 1 receives resources but Link 2 does not; 2) when Ω_1 lies between $[0.6,1]$, Link 2 receives resources but Link 1 does not; and 3) when Ω_1 lies between $[0.5,0.6]$, Links 1 and 2 both receive resources. To better interpret and explain this result, we plot the corresponding path flows after the expansions in Figure 3(b) (note that Link 1 is on Path 1 and Link 2 is on Path 2). We also provide the following traffic assignment data *before any expansion (B = 0) and after* in Table 4.2.

Table 4.2. The traffic assignment data before and after link expansion

	Link	Before			After		
		v_a^{RUE}	$v_a^{RUE}t_a^*\Omega_a$	$\varepsilon_a^*t_a^*v_a^{RUE}\Omega_a$	v_a^{RUE}	$\varepsilon_a^*t_a^*\Omega_a$	$\varepsilon_a^*t_a^*v_a^{RUE}\Omega_a$
$\Omega_1=0$	1	2.54	0.00	0.00	2.78	0.00	0.00
	2	2.46	0.71	1.74	2.22	0.55	1.22
$\Omega_1=1$	1	2.32	7.02	16.30	1.95	5.03	9.18
	2	2.68	0.89	2.39	3.05	0.67	2.05

When Ω_1 is small (e.g., $\Omega_1=0$ in the extreme case), the standard deviation of travel time for Link 1 is small. In this case, we observe from Table 4.2 that before the link expansion, 1) more travelers select Link 1 than Link 2; 2) the standard deviation of the link travel time for Link 1 is smaller than that for Link 2 (i.e., $\varepsilon_1^*t_1^*\Omega_1$ is smaller than $\varepsilon_2^*t_2^*\Omega_2$); and 3) $\varepsilon_1^*t_1^*v_1^{RUE}\Omega_1$ equals zero when $\varepsilon_2^*t_2^*v_2^{RUE}\Omega_2$ is larger than zero (note that $\varepsilon_a^*t_a^*v_a^{RUE}\Omega_a$, $\forall a \in A$, is used to compute the standard deviation of the TSTT and TSTTB [see (11)]). The above observations indicate that more travelers select the more reliable link before the link expansion. Recall that $R=4$ in this example, meaning that it is more important to minimize the standard deviation of the TSTT than to minimize the mean TSTT. Resources should be used to improve the links that help to minimize the standard deviation of the TSTT. When Ω_1 is small, Link 1 (with lower travel time uncertainty and more travelers) should be improved instead of Link 2, because more travelers will be attracted to the improved link with lower travel time uncertainty, and fewer travelers will use the less reliable link (see the link flows of Links 1 and 2 before and after the link expansion

in Table 4.2). Table 4.2 indicates that the value of $\varepsilon_1^* t_1^* \Omega_1$ is reduced after the expansion of Link 1, and that $\varepsilon_1^* t_1^* v_1^{RUE} \Omega_1$ is lower despite the increase in v_1^{RUE}. Furthermore, although there is no expansion of Link 2, both v_2^* and $\varepsilon_2^* t_2^* \Omega_2$ are reduced because more travelers are attracted to Link 1. Thus, $\varepsilon_2^* t_2^* v_2^{RUE} \Omega_2$ is also lower after the link expansion. The result indicates that expanding the more reliable Link 1 instead of Link 2 results in lower $\varepsilon_1^* t_1^* v_1^{RUE} \Omega_1$ and $\varepsilon_2^* t_2^* v_2^{RUE} \Omega_2$, and also reduces the standard deviation of the TSTT. The result explains why, in Figures 3(a) and 3(b), only Link 1 receives resources, and why the flow on Path 1 (which equals the flow on Link 1) is larger than that on Path 2 (which equals the flow on Link 2) when Ω_1 is small.

When Ω_1 is large (e.g., $\Omega_1 = 1$ in the extreme case), the standard deviation of the travel time of Link 1 is large. In this case, we observe from Table 4.2 that before the capacity expansion, 1) more travelers select Link 2 than Link 1; 2) the standard deviation of the link travel time for Link 2 is smaller than that for Link 1 ($\varepsilon_2^* t_2^* \Omega_2$ is smaller than $\varepsilon_1^* t_1^* \Omega_1$); and 3) $\varepsilon_1^* t_1^* v_1^{RUE} \Omega_1$ is considerably larger than $\varepsilon_2^* t_2^* v_2^{RUE} \Omega_2$. These observations are the opposite of those when the travel time standard deviation of Link 1 is small. Hence, the optimal decision is to improve Link 2 but not Link 1 and the flow on Path 2 (i.e., the flow on Link 2) is larger than that on Path 1 (i.e., the flow on Link 1).

When Ω_1 is not large (consider $\Omega_1 \in [0.4, 0.6]$), the differences between the travel time variations and flows for Links 1 and 2 before capacity expansion are not as great as when Ω_1 equals one. In this case, the link expansion tends to improve both links.

Note that to perform the above analysis, we indirectly fixed the value of ε_a, $\forall a \in A$, to be a constant because the variation in congestion-related travel time is ignored (i.e., $e_a = 0$). Nevertheless, the results remain similar if we set $e_a = 0.1$, $\forall a \in A$.

Example 2. The objective is to compare how different levels of risk aversion of the system manager and travelers affect the design. Two scenarios are used, both of which use the following settings:

$e_a = 0.1$, $\forall a \in A$; $s_1 = 0.5$; $s_2 = 0.1$; $s_3 = 0.1$; $\Omega_2 = 0.4$; $\Omega_3 = 0.4$; and $\Omega_1 = 1$.

<u>Scenario 1</u>: $\lambda_1 = 2$, and the values of R change from 0 to 4.

<u>Scenario 2</u>: $R = 2$, and the values of λ_1 change from 0 to 4.

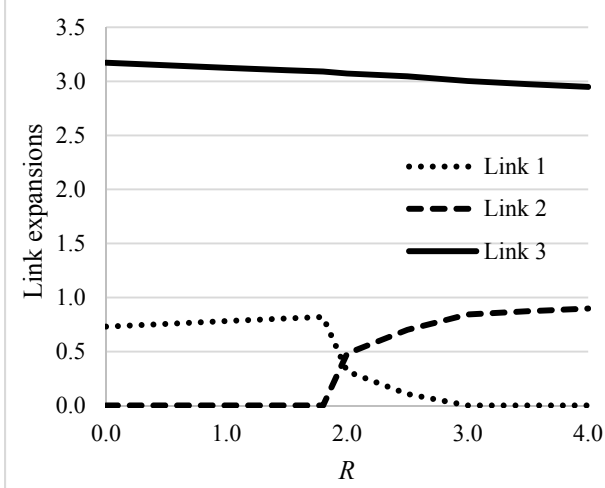The link capacity expansions for the two scenarios are plotted in Figure 4(a-d).
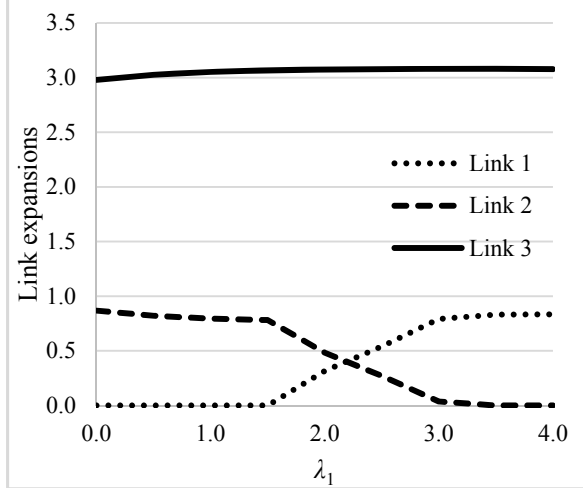
Figure 4(a). Link capacity expansions for Scenario 1
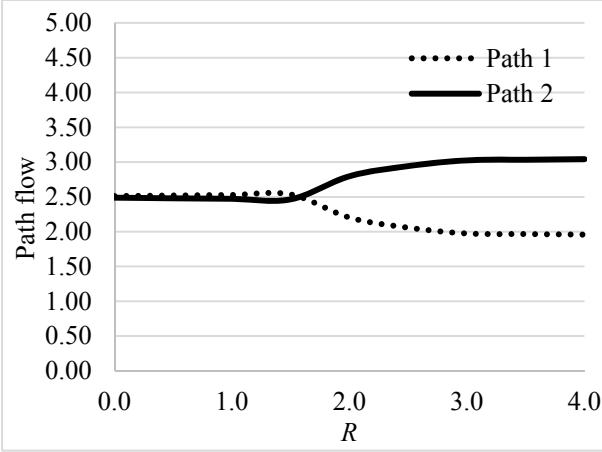

Figure 4(b). Link capacity expansions for Scenario 2
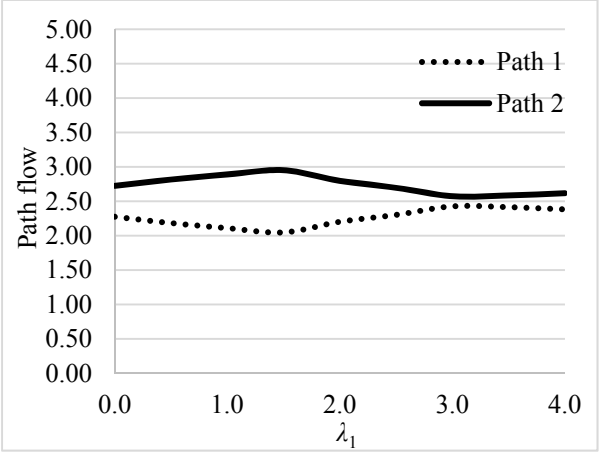

Figure 4(c). Path flows for Scenario 1

Figure 4(d). Path flows for Scenario 2

Observation 2. *The risk-averse behavior of the system manager and the travelers may have different or even opposite effects on link capacity expansions, which indicates the importance of considering the risk-aversion levels of both the system manager and the travelers in the design.*

As the factors R and λ_1 are both related to risk aversion, we may expect them to have similar effects on the link capacity expansions. In this example, we mainly compare the resources allocated to Links 1 and 2, because they are parallel and competing links. The data in Figure 4(a) indicate that when R is very small, resources are only allocated to Link 1 (the link with higher travel time uncertainty), whereas when R is large, resources are only allocated to Link 2 (the link with lower travel time uncertainty). When R is in the medium range, resources are allocated to both links. The data in Figure 4(b) indicate that when λ_1 is very small, resources are only allocated to Link 2, whereas when λ_1 is very large, resources are only allocated to Link 1. In other cases, resources are allocated to both links. Comparing these two figures, we find that the risk-aversion levels of the system manger and travelers can have opposite effects on the link capacity expansions. An increase in R leads to more allocation to Link 2 but an increase in λ_1 leads to less allocation to Link 2.

20

Another observation is that *when R is large, resources might not be allocated to the links with higher travel time uncertainty*. The factor R reflects the weight associated with the objective of minimizing the standard deviation of the TSTT in the design. It is expected that when R is large, the first priority for link expansion is to minimize the travel time variations by improving the links with the greatest travel time uncertainty. However, in Figure 4(a), when R is large (e.g., $R = 4$), Link 2 receives resources while Link 1 receives none. One possible explanation is that reducing the variation in the TSTT can be achieved either by improving uncertain links to reduce the travel time uncertainty, or by improving reliable links to divert travelers from using the links with higher travel time uncertainty. The second approach fits this example, because from Figure 4(c), we observe that the flow on Link 2 is much larger than that on Link 1 (the link with higher travel time uncertainty) when R is large.

When R is small (e.g., $R = 0$), the link expansion tends to minimize the mean TSTT as the first priority. From Figure 4(a), we observe that only Link 1 receives resources while Link 2 receives none. To better understand this result, we provide the path flows, means, and standard deviations of the travel times for both paths before any link expansion in Table 4.3.

Table 4.3. The traffic assignment data before any link expansion for Scenario 1

Path	Path flow	Mean path travel time	Path travel time standard deviation
1	2.35	104.43	22.04
2	2.65	111.75	14.72

Link 1 receives resources because the mean travel time of Path 1 is lower than that of Path 2 before any link expansion (see Table 4.3). When R is small, the link expansion minimizes the mean TSTT by improving the links on the path with lower mean travel times, i.e., Link 1.

When R is neither large nor small (see $R \in [1.8, 3]$), the differences between Links 1 and 2 are not as obvious as when R equals zero or four. Thus, both links receive resources.

The effect of λ_1, however, is quite the opposite, as observed from the results for Scenario 2. Recall that the value of R is set as two in Scenario 2, which means that the link expansion considers not only the mean TSTT, but also its standard deviation. The value of λ_1 directly reflects the route preferences of travelers. When λ_1 is small, travelers are attracted to the paths with lower mean travel times, whereas when λ_1 is large, they are attracted to more reliable paths. Table 4.4 provides the path flows and the mean path travel times and standard deviations before any link expansion, given two instances under different λ_1s.

Table 4.4. The traffic assignment data before any link expansion for Scenario 2

	Path	Path flow	Mean path travel time	Path travel time standard deviation
$\lambda_1 = 0$	1	2.53	107.78	12.64
	2	2.46	107.78	7.30
$\lambda_1 = 4$	1	2.24	102.84	10.31
	2	2.76	114.44	7.41

When λ_1 is small (e.g., $\lambda_1 = 0$), it is more efficient to improve the TSTTB by improving Link 2 on Path 2 than Link 1 on Path 1 because the differences between the mean travel times of Paths 1 and 2 at equilibrium are trivial (see Table 4.4), but the standard deviation of the travel time of Path 2 is much smaller than that of Path 1. Thus, Figure 4(b) shows that when λ_1 is small, Link 2 on Path 2 receives resources while Link 1 on Path 1 receives none.

When λ_1 is large (e.g., $\lambda_1 = 4$) before the link expansion, the number of travelers on Path 2 is larger than that on Path 1 at equilibrium (see Table 4.4). The standard deviation of path travel time of Path 2 is lower than that of Path 1, but Path 1 has a lower mean path travel time than Path 2. From Figure 4(b) we observe that Link 1 on Path 1 receives resources while Link 2 receives none, which may be because improving Link 1 (and reducing Path 1's mean travel time) is more efficient in reducing the TSTTB than improving Link 2 (and reducing Path 2's travel time variation).

When λ_1 is neither small nor large (see $\lambda_1 \in [1.5,3]$), the differences between Paths 1 and 2 are not as obvious as when λ_1 equals zero or four. Thus, both links receive resources.

Example 3. The objective is to illustrate the effect of reliability-based user equilibrium (RUE) on the design. The following parameter values are set: $e_a = 0.1$, $\forall a \in A$; $s_1 = 0.1$; $s_3 = 0.5$; $s_2 = 0.1$; $\Omega_1 = 0.4$; $\Omega_2 = 0.4$; $\Omega_3 = 1$; and $R = 2$. A design *without* considering the RUE constraints is determined and the results are recorded as benchmarks:

$$\textit{Benchmarks}: \; E[TSTT] = 171.96, \; \text{Standard deviation of } TSTT = 14.82,$$

$$TSTTB = 186.77.$$

Several designs that consider the RUE constraints are determined, in which the values of λ_1 vary from 0~4. The values of the mean TSTTs, standard deviations, and TSTTBs at optimality are divided by their corresponding benchmarks to obtain the ratios, and are presented in Figure 5.
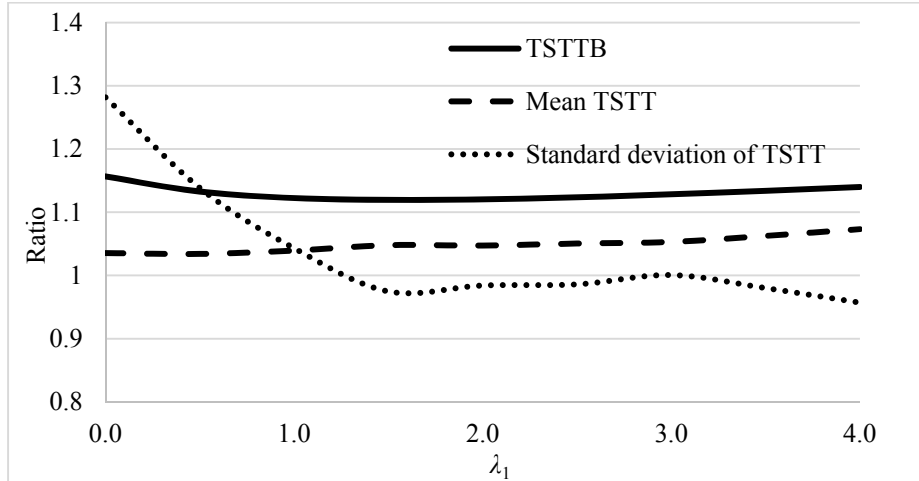

Figure 5. Ratios between the TSTTB, mean TSTT, standard deviation of TSTT, and their benchmarks

Observation 3. *When travelers' selfish routing behavior is considered, the optimal TSTTB is larger than that in the case without such consideration. However, the standard deviation of the TSTT might be smaller, depending on the risk aversion of travelers.*

It is reasonable to expect that the optimal TSTTB becomes larger when travelers choose routes that minimize their own travel times. When the RUE constraints are included in the lower level problem, the feasible solution set is smaller than that without the RUE constraints. The data in Figure 5 are consistent with this expectation. The ratio between the optimal TSTTBs with and without RUE constraints is always larger than one. Figure 5 also suggests that the ratio between the mean TSTTs with and without RUE constraints is larger than one. This indicates that the selfish routing behavior of travelers increases both the mean TSTT and the TSTTB.

In this example, the ratio between the optimal TSTTBs with and without RUE constraints is bounded above by 1.2, the ratio between the mean TSTTs is bounded above by 1.1, and the ratio between the standard deviations is bounded above by 1.3. The results suggest that the efficiency loss (the increase in the objective function value at optimality) due to the selfish routing behavior of travelers is bounded. Theoretically, because the travel demands and the cost budget are finite, the link flows and capacity improvements must be finite. Moreover, the variance of the random variable H_a, $a \in A$, which is Ω_a^2, is also finite (bounded above by one according to Appendix A). As a result, the link travel times [see (5)] and link travel time variance [see (6)] are finite. Hence, no matter whether RUE constraints are considered or not during the design process, the optimal TSTTB, the mean TSTT, and the standard deviation of TSTT must be finite. Thus, the ratios should be bounded above.

Moreover, the ratio between the travel time standard deviations with and without RUE is not only bounded above, and may also be smaller than one. The result shows that when λ_1 is larger than 1.2, the ratio is below 1 and decreases as λ_1 gets larger. This means that the standard deviation of the TSTT is smaller when travelers choose their routes selfishly. A possible explanation is as follows. When RUE is not considered, travelers are cooperative and are distributed across the network in such a way that the TSTTB is minimized. The optimal link expansion tends to minimize both the mean TSTT and its standard deviation at the same time. When RUE is considered, travelers are non-cooperative and are distributed across the network in such a way that their own travel time budgets are minimized. When λ_1 is larger, travelers are more concerned with variations in their travel times and naturally select more reliable paths. In this situation, the optimal link expansions tend to minimize the standard deviation of the TSTT prior to the mean TSTT. Thus, it is reasonable to expect that the standard deviation of the TSTT obtained with RUE constraints will be smaller than that obtained without RUE constraints. In this example, more resources are put into the link with greater travel time uncertainty (i.e., Link 1), which greatly lowers the variation in travel times. This phenomenon indicates that the selfish routing behavior of travelers is not always bad from a design point of view because at least the TSTT variation (standard deviation) can be reduced.

5. Conclusion

We propose a bi-level network design model under supply uncertainty. The distributions of random travel times are not required. The travel time distributions for different links may be heterogeneous. The upper level of the design problem aims to minimize the TSTTB, which

consists of the mean value and a safety margin. The design is achieved by selecting the optimal link capacity expansions subject to a fixed expansion budget. At the lower level, user equilibrium traffic assignment is incorporated to capture the selfish travel behavior of users. The risk aversion level of the system manager and the risk attitude of travelers are also considered. The analysis shows that despite the lack of knowledge of travel time distributions, the probabilities that the actual TSTT is within the optimal TSTTB and the actual path travel time are respectively within the users' minimum path travel time budget under optimal design have lower bounds. The lower bounds are related to the system manager's and users' risk aversion. The optimal TSTTB is proven to be bounded below even when the link expansion budget is unlimited. The optimal TSTTB is non-decreasing in terms of the level of the risk aversion of system manager. As the level gets larger, the mean of the TSTT at optimality is non-decreasing, while the standard deviation of the TSTT is non-increasing.

Our numerical studies suggest the following: (1) improving links with lower uncertainty may have a higher priority and links with higher uncertainty may not receive more resources for improvement; (2) When travelers' selfish routing behavior is considered, the optimal TSTTB is larger than in the case without such consideration. However, the standard deviation of the TSTT might be smaller, depending on the risk aversion of travelers. (3) The risk-averse behavior of the system manager and the travelers may have different or even opposite effects on link capacity expansions, which indicates the importance of considering the risk-aversion levels of both the system manager and the travelers in the design.

Our study aims at providing a new bi-level risk averse network design model for transportation networks with heterogeneous and unknown link travel time distributions and theoretical results obtained from the proposed model. To illustrate the properties of the model and key findings more clearly, only a small network is used, although the model and theoretical results can be applied to large networks. Efficient or exact solution algorithms (e.g., Wang and Lo 2010; Wang et al., 2016) for the proposed model have not been proposed and analyzed and the model has not been applied to large scale transport networks. These are left for future study. Moreover, this paper assumes that traveler follows the reliability-based user equilibrium principle. This principle assumes that all users have perfect information about the traffic conditions. In reality, this is not the case. Hence, one of the future studies is to extend the principle to consider imperfect information and develop a stochastic user equilibrium extension (e.g., Long et al., 2010, 2014; Liu and Wang, 2015; Riemann et al., 2015). Furthermore, a nontrivial lower bound for the probability of trip travel time of risk-seeking travelers and nontrivial upper bounds for the reliability of total system travel time and trip travel time have not been derived. This can be another future research direction.

Appendix A. Properties of Ω_a

The standard deviation Ω_a, $a \in A$, of the random variable H_a has a lower bound of 0 and an upper bound of 1.

Proof. If H_a is a continuous random variable, its standard deviation is $\sqrt{\int_{-1}^{1}(\eta_a - E[H_a])^2 \cdot g_{H_a}(\eta_a)d\eta_a}$, in which $g_{H_a}(\eta_a)$ is the probability density function for H_a. By

definition, Ω_a is larger than or equal to 0, which is the lower bound. Substituting 0 into $E[H_a]$, and based on the fact that $-1 \leq -\kappa_a^L \leq H_a \leq \kappa_a^U \leq 1$ (which implies $-1 \leq \eta_a \leq 1$), then the following must be true: $\Omega_a = \sqrt{\int_{-1}^1 (\eta_a - 0)^2 \cdot g_{H_a}(\eta_a) d\eta_a} \leq \sqrt{\int_{-1}^1 (1-0)^2 \cdot g_{H_a}(\eta_a) d\eta_a} = \sqrt{\int_{-1}^1 g_{H_a}(\eta_a) d\eta_a} = 1$.

If H_a takes random values from a finite data set η_{a1}, η_{a2}, ... , η_{aN}, $-1 \leq -\kappa_a^L \leq \eta_{ai} \leq \kappa_a^U \leq 1$, and each value η_{ai} has a probability of $z_{H_a}(\eta_{ai})$, the standard deviation is $\sqrt{\sum_{i=1}^N (\eta_{ai} - E[H_a])^2 z_{H_a}(\eta_{ai})}$.

Similarly, the following holds:

$$\Omega_a = \sqrt{\sum_{i=1}^N (\eta_{ai} - E[H_a])^2 z_{H_a}(\eta_{ai})} \leq \sqrt{\sum_{i=1}^N (1-0)^2 z_{H_a}(\eta_{ai})} = 1. \blacksquare$$

Appendix B. Proof of Proposition 2

Consider two cases, $R = R'$ and $R = R''$, in which $R' < R''$. Denote $(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}'))$ and $(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}''))$ as the optimal link capacity expansions and link flows for $R = R'$ and $R = R''$, respectively. Denote also ε_a' and ε_a'' as the value of ε_a at $(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}'))$ and $(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}''))$, respectively. According to Proposition 1, $TSTTB_{R=R'}(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')) \leq TSTTB_{R=R''}(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}''))$. Note that the TSTTB consists of two components which are the mean TSTT and weighted standard deviation of TSTT. For the ease of presentation, denote the standard deviation of TSTT at $(\mathbf{y}, \mathbf{v}(\mathbf{y}))$ as $F(\mathbf{y}, \mathbf{v}(\mathbf{y}))$. Then, the aforementioned inequality is equivalent to

$$\sum_{a \in A} t_a' v_a^{RUE}(y_a') + R' \cdot F(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')) \leq \sum_{a \in A} t_a'' v_a^{RUE}(y_a'') + R'' \cdot F(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')). \tag{32}$$

Consider the following three mutually exclusive cases.

Case 1: If

$$\sum_{a \in A} t_a' v_a^{RUE}(y_a') < \sum_{a \in A} t_a'' v_a^{RUE}(y_a''), \text{ and} \tag{33}$$

$$F(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')) < F(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')), \tag{34}$$

then we derive that $\sum_{a \in A} t_a' v_a^{RUE}(y_a') + R'' \cdot F(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')) < \sum_{a \in A} t_a'' v_a^{RUE}(y_a'') + R'' \cdot F(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}''))$, which contradicts the assumption that $(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}''))$ is the solution that minimizes the TSTTB when $R = R''$. Hence, this case is not true.

Case 2: If

$$\sum_{a \in A} t_a' v_a^{RUE}(y_a') > \sum_{a \in A} t_a'' v_a^{RUE}(y_a''), \text{ and} \tag{35}$$

$$F(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')) > F(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')), \tag{36}$$

then we derive that $\sum_{a \in A} t_a'' v_a^{RUE}(y_a'') + R' \cdot F(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')) < \sum_{a \in A} t_a' v_a^{RUE}(y_a') + R' \cdot F(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}'))$, which

contradicts the assumption that $(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}'))$ is the solution that minimizes the TSTTB when $R = R'$. Hence, this case is not true.

Case 3: If

$$\sum_{a \in A} t_a' v_a^{RUE}(y_a') > \sum_{a \in A} t_a'' v_a^{RUE}(y_a''), \text{ and} \tag{37}$$

$$F(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')) < F(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')), \tag{38}$$

then because $(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}'))$ is the optimal solution for $R = R'$, the following holds:

$$\sum_{a \in A} t_a' v_a^{RUE}(y_a') + R' \cdot F(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')) < \sum_{a \in A} t_a'' v_a^{RUE}(y_a'') + R' \cdot F(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')). \tag{39}$$

Let $\Delta = R'' - R' > 0$. Multiplying Δ on both sides of inequality (38), and adding it to inequality (39), the following is obtained:

$$\sum_{a \in A} t_a' v_a^{RUE}(y_a') + R'' \cdot F(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')) < \sum_{a \in A} t_a'' v_a^{RUE}(y_a'') + R'' \cdot F(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')). \tag{40}$$

Inequality (40) contradicts the assumption that $(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}''))$ is the solution that minimizes the TSTTB when $R = R''$. Hence, this case is not true.

The above discussion indicates that there is only one case that is true:

$$\sum_{a \in A} t_a' v_a^{RUE}(y_a') \le \sum_{a \in A} t_a'' v_a^{RUE}(y_a'') \text{ and } F(\mathbf{y}', \mathbf{v}^{RUE}(\mathbf{y}')) \ge F(\mathbf{y}'', \mathbf{v}^{RUE}(\mathbf{y}'')). \blacksquare$$

Appendix C. Proof of Lemmas 1 and 2

The proof is divided into two parts. Part A derives the upper bound of the probability that the actual TSTT is within the optimal TSTTB. Part B proves that the bound is sharp if TSTT follows a symmetric distribution. Denote the standard deviation of TSTT at $(\mathbf{y}, \mathbf{v}(\mathbf{y}))$ as $F(\mathbf{y}, \mathbf{v}(\mathbf{y}))$.

(a) Quoting **Tchebysheff's inequality**:

Let X be an independent random variable with expectation $E(X)$ and variance σ^2. Then, for any positive real number k that is not less than one,

$$\Pr(|X - E(X)| \ge k\sigma) \le \frac{1}{k^2}. \tag{41}$$

Set X to be the compound random variable $\widehat{TSTT}^*$, which is the random total system travel time at optimal design. Let the standard deviation be $F(\mathbf{y}^*, \mathbf{v}^{RUE}(\mathbf{y}^*))$. The following is obtained:

$$\Pr\left(\left|\widehat{TSTT}^* - E\left[\widehat{TSTT}^*\right]\right| \ge k \cdot F(\mathbf{y}^*, \mathbf{v}^{RUE}(\mathbf{y}^*))\right) \le \frac{1}{k^2}, \forall a \in A. \tag{42}$$

There are two cases: (i) the compound random variable $\widehat{TSTT}^*$ symmetrically distributed around its mean value and (ii) it is asymmetrical distributed. For case (i), the probabilities of

$\widehat{TSTT}^* - E\left[\widehat{TSTT}^*\right] \geq k \cdot F(\mathbf{y}^*, \mathbf{v}^{RUE}(\mathbf{y}^*))$ and $E\left[\widehat{TSTT}^*\right] - \widehat{TSTT}^* \geq k \cdot F(\mathbf{y}^*, \mathbf{v}^{RUE}(\mathbf{y}^*))$ are identical. Moreover, the sum of these two probabilities is bounded according to (42). Thus,

$$\Pr\left(\widehat{TSTT}^* - E\left[\widehat{TSTT}^*\right] \geq k \cdot F(\mathbf{y}^*, \mathbf{v}^{RUE}(\mathbf{y}^*))\right) \leq \frac{1}{2k^2}. \tag{43}$$

The left-side of (43) is equivalent to $\Pr\left(\widehat{TSTT}^* \geq \sum_{a \in A} t_a^* v_a^{RUE} + R \cdot F(\mathbf{y}^*, \mathbf{v}^{RUE}(\mathbf{y}^*))\right)$ if we set $k = R$,

which by definition is $\Pr\left(\widehat{TSTT}^* \geq TSTTB^*\right)$, i.e., $\Pr\left(\widehat{TSTT}^* \geq TSTTB^*\right) \leq \frac{1}{2R^2}$. For case (ii), the

sum of the probabilities of $\widehat{TSTT}^* - E\left[\widehat{TSTT}^*\right] \geq k \cdot F(\mathbf{y}^*, \mathbf{v}^{RUE}(\mathbf{y}^*))$ and

$E\left[\widehat{TSTT}^*\right] - \widehat{TSTT}^* \geq k \cdot F(\mathbf{y}^*, \mathbf{v}^{RUE}(\mathbf{y}^*))$ equals the left side of (42). Hence, we have

$$\begin{aligned} &\Pr\left(\widehat{TSTT}^* - E\left[\widehat{TSTT}^*\right] \geq k \cdot F(\mathbf{y}^*, \mathbf{v}^{RUE}(\mathbf{y}^*))\right) \leq \\ &\Pr\left(\left|\widehat{TSTT}^* - E\left[\widehat{TSTT}^*\right]\right| \geq k \cdot F(\mathbf{y}^*, \mathbf{v}^{RUE}(\mathbf{y}^*))\right) \leq \frac{1}{k^2}, \ \forall a \in A. \end{aligned} \tag{44}$$

If we set $k = R$, the left side of inequality (44), by definition, is $\Pr\left(\widehat{TSTT}^* \geq TSTTB^*\right)$, i.e.,

$\Pr\left(\widehat{TSTT}^* \geq TSTTB^*\right) \leq \frac{1}{R^2}$. Based on the conclusion of the two cases, we have

$$\Pr\left(\widehat{TSTT}^* \geq TSTTB^*\right) \leq \begin{cases} \dfrac{1}{R^2}, & \text{for general distributions,} \\ \dfrac{1}{2R^2}, & \text{for symmetric distributions.} \end{cases}$$

(b) To prove that the upper bound is sharp if TSTT follows a symmetric distribution, we just need to find a case with explicit travel time distributions in which the equality sign holds. Consider the following case.

A network with two nodes, A and B, is connected by two parallel links, denoted as Links 1 and 2. The travel times of two links are independent. The demand from A to B is $d = 5$. Given optimal link expansion, the predicted link travel times are $t_1(v_1) = 0.1 + 0.1 v_1$ and $t_2(v_2) = 1 + 0.6 v_2$. $e_1 = e_2 = 0$; $s_1 = 0.1$; $s_2 = 0.5$; then, ε_1 and ε_2 equal 0.1 and 0.5, respectively. The PMF of the random variable H_1 is $z_{H_1}(\eta_1) = 1/2R^2$, $\eta_1 = -1$; $z_{H_1}(\eta_1) = 1 - 1/R^2$, $\eta_1 = 0$; $z_{H_1}(\eta_1) = 1/2R^2$, $\eta_1 = 1$. The standard deviation Ω_1 is $1/R$ (computed based on the PMF).

- As Link 1 has a much lower travel time and variance, all users d take Link 1 to travel from A to B, i.e., $v_1^{RUE} = d$ and $v_2^{RUE} = 0$. The optimal TSTTB is

$$\sum_{a \in A} t_a^* v_a^{RUE} + R \cdot F(\mathbf{y}^*, \mathbf{v}^{RUE}(\mathbf{y}^*)) = t_1(d)d + R\sqrt{\left(t_1(d) d \varepsilon_1 \Omega_1\right)^2}.$$

- The actual TSTT is $\sum_a T_a^* v_a^{RUE} = T_1^* d$. The PMF of $\sum_a T_a^* v_a^{RUE}$ is identical to that of $T_1^* d$. Recall that $T_1^* = t_1(d) + t_1(d)\varepsilon_1\eta_1$ [see (3)]. The mean of $T_1^* d$ is $t_1(d)d$, and the standard deviation is $t_1(d)d\varepsilon_1\Omega_1$. According to the PMF of H_1, the PMF of $T_1^* d$ is

$$z_{T_1^* d}(\omega) = \begin{cases} 1/2R^2, & \omega = t_1(d)d - t_1(d)d\varepsilon_1, \\ 1 - 1/R^2, & \omega = t_1(d)d, \\ 1/2R^2, & \omega = t_1(d)d + t_1(d)d\varepsilon_1. \end{cases}$$

The probability that $T_1^* d$ is *not* within $t_1(d)d + R \cdot \sqrt{(t_1(d)d\varepsilon_1\Omega_1)^2}$ is

$$\Pr\left(T_1^* \cdot d \geq t_1(d)d + R \cdot \sqrt{(t_1(d)d\varepsilon_1\Omega_1)^2}\right) = \Pr\left(T_1^* \cdot d \geq t_1(d)d + R \cdot t_1(d)d\varepsilon_1 \cdot \frac{1}{R}\right)$$
$$= \Pr\left(T_1^* \cdot d \geq t_1(d)d + t_1(d)d\varepsilon_1\right)$$
$$= 1/2R^2.$$

As $\sum_{a \in A} T_a^* v_a^{RUE} = T_1^* d$, we have

$$\Pr\left(\sum_{a \in A} T_a^* v_a^{RUE} \geq \sum_{a \in A} t_a^* v_a^{RUE} + R \cdot F(\mathbf{y}^*, \mathbf{v}^{RUE}(\mathbf{y}^*))\right) = \Pr\left(T_1^* \cdot d \geq t_1(d)d + R \cdot \sqrt{(t_1(d)d\varepsilon_1\Omega_1)^2}\right)$$
$$= 1/2R^2.$$

This completes the proof.■

References

Abdel-Aty, M., Kitamura, R. and Jovanis, P. (1995) Investigating effect of travel time variability on route choice using repeated-measurement stated preference data. *Transportation Research Record*, 1493, 39-45.

Abdulaal, M. and LeBlanc, L.J. (1979) Continuous equilibrium network design models. *Transportation Research Part B*, 13(1), 19-32.

Ben-Ayed, O., Boyce, D.E. and Blair III, C.E. (1988) A general bilevel linear programming formulation of the network design problem. *Transportation Research Part B*, 22(4), 311-318.

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

Chen, A. and Subprasom, K. (2007) Analysis of regulation and policy of private toll roads in a build-operate-transfer scheme under demand uncertainty. *Transportation Research Part A*, 41(6), 225-247.

Chen, A. and Yang, C. (2004) Stochastic transportation network design problem with spatial equity constraint. *Transportation Research Record*, 1882, 97-104.

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

Chen, B.Y., Lam, W.H.K., Sumalee, A. and Li, Z.L. (2012) Reliable shortest path finding in stochastic networks with spatial correlated link travel times. *International Journal of Geographical Information Science*, 26(2), 365-386.

Chen, B.Y., Lam, W.H.K., Sumalee, A., Li, Q.Q., Shao, H. and Fang, Z.X. (2013) Finding reliable shortest paths in road networks under uncertainty. *Networks & Spatial Economics*, 13(2), 123-148.

Chen, B.Y., Lam, W.H.K., Sumalee, A. and Shao, H. (2011) An efficient solution algorithm for solving multi-class reliability-based traffic assignment problem. *Mathematical and Computer Modelling*, 54(5-6), 1428-1439.

Chiou, S. (2005) Bilevel programming for the continuous transport network design problem. *Transportation Research Part B*, 39(4), 361-383.

Chiou, S. (2008) A hybrid approach for optimal design of signalized road network. *Applied Mathematical Modelling*, 32(2), 195-207.

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

Chow, J.Y.J. and Regan, A.C. (2011) Network-based real option models. *Transportation Research Part B*, 45(4), 682-695.

Dantzig, G.B., Harvey, R.P., Landsowne, Z.F., Robinson, D.W. and Maier, S.F. (1979) Formulating and solving the network design problem by decomposition. *Transportation Research Part B*, 13(1), 5-17.

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

Dial, R.B. (2006) A path-based user-equilibrium traffic assignment algorithm that obviates path storage and enumeration, *Transportation Research Part B*, 40(10), 917-936.

Dimitriou, L. and Stathopoulos, A. (2008) Reliable stochastic design of road network systems. *International Journal of Industrial and Systems Engineering*, 3(5), 549-574.

Farahani, R.Z., Miandoabchi, E., Szeto, W.Y. and Rashidi, H. (2013) A review of urban transportation network design problems. *European Journal of Operational Research*, 229(2), 281-302.

Friesz, T.L., Cho, H., Mehta, N.J., Tobin, R.L. and Anandalingam, G. (1992) A simulated annealing approach to the network design problem with variational inequality constraints. *Transportation Science*, 26(1), 18-26.

Gao, Z., Sun, H. and Zhang, H. (2007) A globally convergent algorithm for transportation continuous network design problem. *Optimization and Engineering*, 8(3), 241-257.

Han, K., Szeto, W.Y., and Friesz, T.L. (2015) Formulation, existence, and computation of simultaneous route-and-departure choice bounded rationality dynamic user equilibrium with fixed or endogenous user tolerance. *Transportation Research Part B*, 79, 16-49.

Haas, I and Bekhor, S. (2016) A parsimonious heuristic for the discrete network design problem. *Transportmetrica A: Transport Science*, 12(1), 43-64.

Jackson, W. and Jucker, J. (1982). An empirical study of travel time variability and travel choice behavior. *Transportation Science*, 16(4), 460-475.

Jiang, Y. and Szeto, W.Y. (2015). Time-dependent transportation network design that considers health cost. *Transportmetrica A: Transportation Science*, 11(1), 74-101.

Khooban, Z., Farahani, R.Z., Miandoabchi, E. and Szeto, W.Y. (2015) Mixed network design using hybrid scatter search. *European Journal of Operational Research*, 247(3), 699-710.

LeBlanc, L.J. and Boyce, D.E. (1986) A bilevel programming algorithm for exact solution of the network design problem with user-optimal flows. *Transportation Research Part B*, 20(3), 259-265.

Li, H., Bliemer, M. and Bovy, P. (2008) Network reliability-based optimal toll design. *Journal of Advanced Transportation*, 42(3), 311-332.

Li, Z.C., Li, Z.K. and Lam, W.H.K. (2014) An integrated design of sustainable land use and transportation system with uncertainty in future population. *Transportmetrica A: Transport Science*, 10(2), 160-185.

Liu, H. and Wang, D.Z.W. (2015) Global optimization method for network design problem with stochastic user equilibrium. *Transportation Research Part B*, 72, 20-39.

Lo, H.K. and Tung, Y.K. (2003) Network with degradable links: capacity analysis and design. *Transportation Research Part B*, 37(4), 345-363.

Lo, H.K. and Szeto, W.Y. (2009) Time-dependent transport network design under cost-recovery. *Transportation Research Part B*, 43(1), 142-158.

Lo, H.K., Luo, X.W. and Siu, B.W.Y. (2006) Degradable transport network: Travel time budget of travelers with heterogeneous risk aversion. *Transportation Research Part B*, 40(9), 792-806.

Long, J.C., Gao, Z.Y., Zhang, H.Z. and Szeto, W.Y. (2010) A turning restriction design problem in urban road networks. *European Journal of Operational Research*, 206(3), 569-578.

Long, J.C., Szeto, W.Y. and Huang, H.J. (2014) A bi-objective turning restriction design problem in urban road networks. *European Journal of Operational Research*, 237(2), 426-439.

Magnanti, T.L. and Wong, R.T. (1984) Network design and transportation planning: Models and Algorithms. *Transportation Science*, 18(1), 1-55.

Marcotte, P. (1986) Network design problem with congestion effects: a case of bilevel programming. *Mathematical Programming*, 34(2), 142-162.

Mathew, T.V. and Sharma, S. (2009) Capacity expansion problem for large urban transportation networks. *Journal of Transportation Engineering*, 135(7), 406-415.

Meng, Q., Yang, H. and Bell, M.G.H. (2001) An equivalent continuously differentiable model and a locally convergent algorithm for the continuous network design problem. *Transportation Research Part B*, 35(1), 83-105.

Meng, Q. and Yang, H. (2002) Benefit distribution and equity in road network design. *Transportation Research Part B*, 36(1), 19-35.

Miandoabchi, E., Farahani, R.Z. and Szeto, W.Y. (2012a) Bi-objective bimodal urban road network design using hybrid metaheuristics. *Central European Journal of Operations Research*, 20(4), 583-621.

Miandoabchi, E., Farahani, R.Z., Dullaert, W. and Szeto, W.Y. (2012b) Hybrid evolutionary metaheuristics for concurrent multi-objective design of urban road and public transit networks. *Networks and Spatial Economics*, 12(3), 441-480.

Miandoabchi, E., Daneshzand, F., Szeto, W.Y. and Farahani, R.Z. (2013) Multi-objective discrete urban road network design. *Computers & Operations Research*, 40(10), 2429-2449.

Miandoabchi, E., Daneshzand, F., Farahani, R.Z. and Szeto, W.Y. (2015) Time-dependent discrete road network design with both tactical and strategic decisions. *Journal of the Operational Research Society*, 66, 894-913.

Ng, M.W. and Waller, S.T. (2009a) Reliable system optimal network design: convex mean-variance model with implicit chance constraints. *Transportation Research Record: Journal of the Transportation Research Board*, 2090, 68-74.

Ng, M.W. and Waller, T.S. (2009b) The evacuation optimal network design problem: model formulation and comparisons. *Transportation Letters—The International Journal of Transportation Research*, 1(2), 111-119.

Riemann, R, Wang, D. Z. W. and Busch, F. (2015). Optimal location of wireless charging facilities for electric vehicles: flow-capturing location model with stochastic user equilibrium. *Transportation Research Part C*, 58 Part A, 1-12.

Siu, B. and Lo, H. (2008). Doubly uncertain transport network: degradable capacity and stochastic demand. *European Journal of Operational Research*, 191(1), 166-181.

Shao, H., Lam, W.H.K. and Tam, M.L. (2006) A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand. *Networks and Spatial Economics*, 6(3), 173-204.

Steenbrink, P.A. (1974) *Optimization of Transport Network*. Wiley, New York, USA.

Suh, S. and Kim, T. (1992) Solving nonlinear bilevel programming models of the equilibrium network design problem: a comparative review. *Annals of Operations Research*, 34(1), 203-218.

Sumalee, A., Luathep, P., Lam, W.H.K. and Connors, R.D. (2009) Transport network capacity evaluation and design under demand uncertainty. *Transportation Research Record*, 2090, 93-101.

Sumalee, A., Watling, D.P. and Nakayama, S. (2006) Reliable network design problem: the case with uncertain demand and total travel time reliability. *Transportation Research Record*, 1964, 81-90.

Szeto, W.Y. and Lo, H.K. (2005) Strategies for road network design over time: robustness under uncertainty. *Transportmetrica*, 1(1), 47-63.

Szeto, W.Y., Jaber, X.Q. and O'Mahony, M. (2010) Time-dependent discrete network design frameworks considering land use. *Computer-Aided Civil and Infrastructure Engineering*, 25(6), 411-426

Szeto, W.Y., Wang, Y. and Wong, S.C. (2014) The chemical reaction optimization approach to solving the environmentally sustainable network design problem. *Computer-Aided Civil and Infrastructure Engineering*, 29(2), 140-158.

Szeto, W.Y. and Wang, A.B. (2015) Price of anarchy for reliability-based traffic assignment and network design. *Transportmetrica A: Transport Science*, 11(7), 603-635.

Szeto, W.Y., Jiang, Y., Wang, D.Z.W. and Sumalee, A. (2015) A sustainable road network design problem with land use transportation interaction over time. *Networks and Spatial Economics*, 15(3), 791-822.

Waller, S.T. and Ziliaskopoulos, A.K. (2001) Stochastic dynamic network design problem. *Transportation Research Record*, 1771, 106-113.

Wang, D.Z.W. and Lo, H.K. (2010) Global optimum of the linearized network design problem with equilibrium flows. *Transportation Research Part B*, 44(4), 482-492.

Wang, D.Z.W., Liu, H.-X and Szeto, W.Y. (2015a). A novel discrete network design problem formulation and its global optimization solution algorithm. *Transportation Research Part E* 79, 213-230.

Wang, D.Z.W., Liu, H.-X., Szeto, W.Y. and Chow, A.H.F. (2016). Identification of critical combination of vulnerable links in transportation networks – a global optimization approach. *Transportmetrica A: Transport Science*, DOI:10.1080/23249935.2015.1137373.

Wang, H., Lam, W.H.K., Zhang, X.N. and Hu, S. (2015b) Sustainable transportation network design with stochastic demands and chance constraints. *International Journal of Sustainable Transportation*, 9(2), 126-144.

Xu, T., Wei, H. and Hu, G. (2009) Study on continuous network design problem using simulated annealing and genetic algorithm. *Expert Systems with Applications*, 36(2, Part 1), 1322-1328.

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

Yang, H. and Bell, M.G.H. (1998) Models and algorithms for road network design: a review and some new developments. *Transport Reviews*, 18(3), 257-278.

Yang, H. and Wang, J.Y.T. (2002) Travel time minimization versus reserve capacity maximization in the network design problem. *Transportation Research Record*, 1783, 17-26.

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

Yin, Y., Madanat, S.M. and Lu, X.Y. (2009) Robust improvement schemes for road networks under demand uncertainty. *European Journal of Operational Research*, 198(2), 470-479.

Yin, Y., Lam, W.H.K. and Ieda, H. (2004) New technology and the modeling of risk-taking behavior in congested road networks. *Transportation Research Part C*, 12(3-4), 171-192.