JOURNAL OF ADVANCED TRANSPORTATION
J. Adv. Transp. 2013; **47**:79–104
Published online 20 February 2012 in Wiley Online Library (wileyonlinelibrary.com). **DOI**: 10.1002/atr.216

Alternate capacity reliability measures for transportation networks

Anthony Chen[1,3]*, Panatda Kasikitwiwat[2] and Chao Yang[3]

[1]*Department of Civil and Environmental Engineering, Utah State University, Logan, UT 84322-4110, U.S.A.*
[2]*Department of Civil Engineering, Kamphaeng Saen Campus, Kasetsart University, 73140, Thailand*
[3]*Key Laboratory of Road and Traffic Engineering, Tongji University, Shanghai, 201804, China*

SUMMARY

Capacity reliability is defined as the probability that the network capacity can accommodate a certain volume of traffic demand at a required service level. It is a supply-side reliability measure for assessing the adequacy of a degradable transportation network. The network capacity model used to calculate the capacity reliability measure is based on the concept of reserve capacity, which requires preserving a pre-determined origin–destination (O–D) demand pattern. In this paper, we relax this assumption by allowing a non-uniform growth in the spatial distribution of the O–D demand pattern. By using this non-uniform O–D growth approach, two network capacity models related to the concepts of ultimate capacity and practical capacity are developed to estimate alternate capacity reliability measures. Numerical results are provided to analyze the features of three capacity reliability measures for transportation networks. Copyright © 2012 John Wiley & Sons, Ltd.

KEY WORDS: network capacity; reserve capacity; traffic equilibrium; bi-level program; capacity reliability

1. INTRODUCTION

Capacity reliability analysis addresses the issue of adequate capacity planning for a highway network to accommodate the growing traffic demand. It provides a probabilistic assessment of road networks under uncertainty. It is an important tool for planning a reliable transportation system when considering everyday disturbances, such as uncertainty of route choice behavior, variations in traffic demand, and link capacity degradation caused by traffic accidents, maintenance work, or bad weather conditions. Further, capacity reliability analysis can be incorporated as an integral part of the transportation systems analysis. Potential applications include decision making for infrastructure management (e.g., prioritizing degraded/degradable roadways for repairs), improving the durability of roadways against man-made/natural disasters, and providing an indication to implement flow control (e.g., congestion pricing, traffic restraint, emission control, etc.). For further information on modeling uncertainty in traffic and transportation applications, see Ottomanelli and Wong [1].

Conventional capacity reliability analysis used the concept of reserve capacity to determine the spare capacity of isolated intersections [2–4]. This concept was then extended to a general signal-controlled road network using bi-level programming to determine the largest common multiplier that can be applied to an origin–destination (O–D) matrix subject to signal timing and saturation

*Correspondence to: Anthony Chen, Department of Civil and Environmental Engineering, Utah State University, Logan, UT 84322-4110, U.S.A. E-mail: anthony.chen@usu.edu, fengpak@ku.ac.th, tongjiyc@tongi.edu.cn

flow constraints with user-equilibrium assignment [5]. Ceylan and Bell [6] further extended the reserve capacity model of Wong and Yang [5] to consider signal coordination under a fixed-time signal-controlled road network. Sumalee *et al*. [7], on the other hand, explicitly considered stochastic demand in the reserve capacity model to design robust networks that can accommodate different levels of variation in future stochastic O–D demand. Using the concept of reserve capacity of a road network, Chen *et al*. [8] defined capacity reliability as the probability that the network capacity can accommodate a certain volume of traffic demand at a required service level. Chen *et al*. [9] provided an assessment methodology, which combines the reliability and uncertainty analysis, network equilibrium models, sensitivity analysis of the equilibrium network flow, and the expected performance measure, as well as Monte Carlo methods, to assess the capacity reliability of a degradable road network. Chen *et al*. [10] also examined the effect of route choice models on assessing network capacity reliability. In terms of applications, Sumalee and Kurauchi [11] adopted the capacity reliability concept to evaluate traffic regulation strategies after a major disaster; Lo and Tung [12] studied the problem of allocating design capacities to maximize the reserve capacity subject to the probabilistic user-equilibrium constraints and the budgetary constraint; Gao and Song [13] extended the reserve capacity of a signal-controlled road network of Wong and Yang [5] to include capacity enhancement as a network design problem (NDP); Chootinan *et al*. [14] provided an alternate capacity reliability index, which measures the probability that all of the network links are operating below their respective capacities under the day-to-day route choice variability, as a surrogate for determining the optimal design variables to maximize network capacity (or to minimize the probability of link failures) in a reliability-based NDP model; Chen *et al*. [15] embedded this new capacity reliability index to formulate a new reserve capacity model for a signal-controlled road network; Chen *et al*. [17] extended the work of Chootinan *et al*. [14] to include both capacity reliability (a supply-side measure for the planners) and travel-time reliability (a demand-side measure for the network users) as a bi-objective reliable NDP model that explicitly optimizes both supply-side and demand-side reliability measures with demand uncertainty; and Yim *et al*. [16], on the other hand, extended the reliability-based NDP model to include not only capacity enhancements of the network but also residential and job allocations in the system to form a more comprehensive reliability-based land use and transportation model for the integrated residential and job allocations and transportation NDP. For NDP under uncertainty, see Chen *et al*. [17] for a recent review and new developments and Chen *et al*. [18] for a review of multi-objective NDP models.

In the previous studies of capacity reliability analysis by Chen *et al*. [8,10,9], the concept of reserve capacity for a signal-controlled road network was used to calculate the capacity reliability measure. Although the concept of reserve capacity has provided a feasible approach to determining the maximum network capacity with a particular route choice, it is restricted to a common multiplier for all O–D pairs. Essentially, it assumes that every O–D pair will have a uniform growth or decline in its O–D demand pattern. The fixed O–D distribution pattern assumption also means that the relative ratio of the resultant total trip productions or attractions is also fixed and hence is only of limited use in undertaking comparative analysis of zonal activity allocation. The purpose of this paper is to relax this assumption by allowing non-uniform growth in the spatial distribution of the O–D demand pattern when estimating the network capacity of a transportation network. Two alternate network capacity models related to the concepts of ultimate capacity and practical capacity are applied to assess the capacity reliability of a transportation network.

This paper is organized as follows. After the introduction, Section 2 reviews the transportation network reliability measures. Section 3 describes the network capacity problems related to three different concepts of network capacity: reserve capacity, ultimate capacity, and practical capacity. Using these network capacity models, alternate capacity reliability measures are provided. Section 4 presents a capacity reliability assessment procedure, which includes Monte Carlo simulation, network capacity evaluation, and reliability analysis. A numerical example is presented in Section 5 to illustrate the features of the alternate capacity reliability measures, and concluding remarks are provided in Section 6.

DOI: 10.1002/atr

2. REVIEW OF TRANSPORTATION NETWORK RELIABILITY MEASURES

Reliability is generally defined as the probability that the system of interest has the ability to perform an intended function or goal [19]. Traditionally, transportation network reliability studies were concerned mainly with two problems: connectivity reliability and travel-time reliability [20]. Recently, the reliability of transportation networks has become an increasingly important issue because of its critical status as the most important lifeline in the restoration process following the occurrence of a disaster [21]. It has attracted many researchers to develop various indicators to assess the reliability of transportation networks (see the recently edited books, proceedings, and special issues by Lam [22], Bell and Cassir [23], Bell and Iida [24], Nicholson and Dantas [25], Sumalee and Kurauchi [26], Murray and Grubesic [27], van Zuylen [28], Kurauchi and Sumalee [29], Schmocker and Lo [30], and Levinson *et al*. [31]). These reliability indicators are summarized in Table I and described in the following sections.

2.1. Connectivity reliability

Connectivity reliability is concerned with the probability that network nodes are connected. A special case of connectivity reliability is the terminal reliability, which concerns the existence of a path between a specific O–D pair [32]. For each node pair, the network is considered successful if at least one path is operational. A path consists of a set of roadways or links that are characterized by zero–one variables denoting the state of each link (operating or failed). Capacity constraints on the links are not accounted for when determining connectivity reliability. This type of connectivity reliability analysis may be suitable for abnormal situations, such as earthquakes, but there is an inherent deficiency in the sense that it only allows for two operating states: operating at full capacity or complete failure with zero capacity. The binary state approach limits the application to everyday situations where links are operating between these two extremes. Therefore, the reliability and risk assessment results obtained through this approach may be misleading for normal conditions.

2.2. Travel-time reliability

Travel-time reliability is concerned with the probability that a trip between a given O–D pair can be made successfully within a given time interval and a specified level of service [33–34]. This measure is useful when evaluating network performance under normal daily flow variations. Bell *et al*. [34] proposed a sensitivity analysis-based procedure to estimate the variance of travel time arising from daily demand fluctuations. Asakura [35] extended the travel-time reliability measure to consider capacity degradation due to deteriorated roads. He defined travel-time reliability as a function of the ratio of travel times under the degraded and non-degraded states. This definition of reliability can be used as a criterion to define the level of service that should be maintained despite the deterioration of certain links in the network. Chen *et al*. [36–37] further examined the effect of considering different risk-taking route choice models on estimation of travel-time reliability under demand and supply variations.

2.3. Within budget time reliability

Within budget time reliability (WBTR) is concerned with the probability that travel time exceeding the travel-time budget (TTB) is less than a predefined confidence level α specified by the traveler to represent his or her risk preference [38]. TTB is a random variable, and it is defined by a travel-time reliability chance constraint. Each commuter is assumed to learn the travel-time variations through his or her daily commutes and choose a route that minimizes his or her TTB. The WBTR definition is similar to that defined by Chen and Ji [39], where the path with the minimum TTB is termed "α-reliable path". It is also similar to the equilibrium conditions of the demand driven travel-time reliability-based user-equilibrium model [40] despite that the source of travel-time variability is induced by demand uncertainty. The core idea behind the WBTR definition is based on the concept of TTB, which is defined as the average travel time

DOI: 10.1002/atr

Table I. Principal characteristics and definitions of transportation network reliability indices (modified from [42]).

Index	Uncertainty	Performance indicator	Probability definition	Reliability aspect for	
				Users	Planners
Connectivity reliability [32]	Disruption of road links	$\theta_c = 1$ if connected and $\theta_c = 0$ if disconnected	Connected and disconnected network	Minimal usefulness	Good usefulness
Travel-time (threshold based) reliability [33]	Fluctuation of daily traffic flow	Specified travel time	Travel time less than a specified value	Good usefulness	Minimal usefulness
Travel-time (level-of-service based) reliability [35]	Degradable link capacity	Specified network service level	Service level less than a specified value	Minimal usefulness	Good usefulness
Travel demand reduction reliability [21]	Degradable link capacity	Intolerable decrement rate of O–D flow	Decrement rate less than a specified value	Minimal usefulness	Good usefulness
Within budget time reliability [38]	Degradable link capacity	On-time arrival	Minimum travel-time budget required to satisfy a certain confidence level of on-time arrival	Good usefulness	Minimal usefulness
Travel demand satisfaction reliability [42]	Degradable link capacity	Demand ratio	Travel demand satisfaction ratio greater than a specified value	Minimal usefulness	Good usefulness
Encountered reliability [43]	Disruption or degradation of road links	Least cost path	Not encountering a link degradation	Good usefulness	Minimal usefulness
Capacity reliability [8,10,9]	Degradable link capacity	Required demand level	Network capacity greater than a specified value	Minimal usefulness	Good usefulness

plus an extra buffer time as an acceptable travel time, such that the probability of completing the trip within the TTB is no less than a predefined reliability threshold (or a confidence level α). In fact, the concept of TTB is analogous to the Value-at-Risk, which is by far the most widely applied risk measure in the finance area [41].

2.4. *Travel demand reduction reliability*

Travel demand reduction reliability is concerned with the probability that the decrement rate of O–D flow is less than a given intolerable value under a degradable network. Nicholson and Du [21] provided two definitions to evaluate the reduction of travel demand: O–D subsystem and system reliabilities. The decrement rate of O–D flow is defined as the ratio of the travel demand reduction due to degradation of a network to the travel demand of a non-degradable network. The O–D flow decrement rate can vary between zero (i.e., no degradation) to unity (i.e., degradation is so severe that the travel demand is zero). A system surplus was also suggested as a performance measure to assess the socioeconomic impacts of the system degradation.

2.5. *Travel demand satisfaction reliability*

Travel demand satisfaction reliability is concerned with the probability that the network can accommodate a given travel demand satisfaction ratio. Heydecker *et al*. [42] defined this ratio as the equilibrium travel demand (i.e., travel demand that can be satisfied by using the transportation network) over the latent travel demand (i.e., total travel demand that intends to use the transportation network). The latent travel demand is the sum of the equilibrium travel demand or the satisfied travel demand and non-satisfied travel demand stemming from the latent demand events, such as horse race or weekend events. The key feature of this reliability measure is that it attempts to distinguish the difference between recurrent travel demand and latent travel demand under a degradable transportation network.

2.6. *Encountered reliability*

Encountered reliability is concerned with the probability that a trip can be made successfully without encountering link degradation on the least (expected) cost path [43]. Level of information to the users is important in the encountered reliability because users will often try to avoid degraded links and links that may be degraded. In addition, different users may behave differently. Risk-averse users are concerned with avoiding disruptions and are willing to travel longer, while risk-neutral users will still travel on their preferred routes based on expected cost considerations regardless of the probability of encountering disruptions along their preferred routes.

2.7. *Capacity reliability*

Capacity reliability refers to the probability that the network capacity can accommodate a certain volume of traffic demand at a required service level [8,10,9]. Link capacities for a road network can change from time to time because of various reasons, such as the blockage of one or more lanes because of traffic accidents, and are considered random variables. The joint distribution of random link capacities can be experimentally obtained or theoretically specified. Capacity reliability explicitly considers the uncertainties associated with link capacities by treating roadway capacities as continuous quantities subject to routine degradation due to physical and operational factors. Readers may note that when the roadway capacities are assumed to take only discrete binary values (zero for total failure and one for operating at ideal capacity), then capacity reliability includes connectivity reliability as a special case.

3. NETWORK CAPACITY MODELS AND CAPACITY RELIABILITY MEASURES

Network capacity in the transportation system is an important measurement for transportation planning and management because it addresses the question of whether or not the transportation system has adequate capacity to handle continuing economic surges and traffic congestion. It has

DOI: 10.1002/atr

been used in examining various transportation problems: traffic restraint [44], road pricing [45], traffic control [5,13,6,15], and road network design [46–47,14,17].

Capacity in transportation has traditionally been measured at individual elements of the network, such as links (rail lines, road segments, waterways, etc.) and nodes (terminals, signalized intersections, etc.). These measures do not constitute the transportation network capacity as a capacity for the whole system [48]. Assuming that capacities of links are known, the maximal network capacity can be determined by the classical maximal flow problem, which is formulated as a network-programming problem to find a feasible flow that leads to the maximum flow capacity. This method has been used in communication networks [49–50], water distribution systems [51], electric power systems [52], and others to determine the maximum flow capacity of the network. For the capacity problem of freight transportation, Nijkamp *et al*. [53] stated that capacity of a transportation system has to be viewed as multidimensional constraints that include different constituents such as environmental limitations, regulations, and effective management roles. Morlok and Riddle [48] provided an operational mathematical model that considers some of these multidimensional constraints.

The approaches for communication networks, water distribution systems, electric power systems, and freight transportation systems, however, are not directly applicable to a passenger transportation network where capacity modeling characteristics are quite different for the following reasons: (1) the movement in a transportation network involves flows of people rather than pure physical commodities as treated in the classical maximal flow problem; (2) travel delay increases with increasing flow as a result of congestion and as opposed to fixed cost; (3) route choice behavior has to be considered; (4) the traditional maximal flow problem does not consider the level of service when finding the maximum throughput; however, transportation network capacity should be specified with a level of service, such as a predefined threshold of the O–D travel time; and (5) multiple O–D pairs exist, and the flows of different O–D pairs are not exchangeable or substitutable in a passenger transportation network capacity problem; thus, it is important to define the O–D demand pattern that influences the resultant value of the transportation network capacity. These characteristics make the modeling of a transportation network capacity quite complex yet an intriguing problem to solve [54]. In this section, models for estimating three network capacity concepts of a transportation network are provided, as well as the application of these network capacity models for estimating alternate capacity reliability measures.

3.1. Network capacity concepts

In this study, three capacity concepts are applied to the transportation problem to estimate the network capacity: (1) reserve capacity, (2) ultimate capacity, and (3) practical capacity. The concept of reserve capacity is based on the premise of preserving a fixed demand pattern when estimating network capacity. This concept is useful when the zonal growth information is not available (i.e., assuming all O–D pairs grow at the same rate). The ultimate capacity concept, when applied to the transportation problem, is the maximum throughput the system can handle without violating roadway and zonal capacity constraints (i.e., only consider the physical limitations of the system). The network users can choose both destination and route simultaneously to minimize their user costs. The practical capacity concept, applied to the transportation problem, is the summation of the current O–D demand and the additional demand that the system can accommodate without violating roadway and zonal capacity constraints. The additional demand or users can choose both destination and route simultaneously, while the current demand pattern is preserved. For destination choice, users choose the destination based on travel time to the destination and attractiveness measures of the destination. In system capacity, practical capacity is the maximum output at which cost does not exceed a maximum acceptable value or capacity limit, which continues to provide an acceptable level-of-service deterioration or delay.

The reserve capacity concept, although simple, requires the assumption of preserving the base O–D demand pattern while determining the largest multiplier that can be allocated to the network

DOI: 10.1002/atr

without violating any link capacity in the network. This reserve capacity model hence can only capture the changes in demand volume (i.e., changes in demand pattern are not considered). The fixed O–D distribution pattern assumption also means that the relative ratio of the resultant total trip productions or attractions is also fixed and hence is only of limited use in undertaking comparative analysis of zonal activity allocation. Another drawback of the reserve capacity model is the lack of level-of-service consideration in determining the network capacity. It simply treats the road network capacity as a maximum physical amount of flow capable of being accommodated without modeling the interaction between network capacity and level of service of a road network [54].

The two alternate capacity concepts to be used in capacity reliability analysis enable evaluating both changes in demand volume and variations in demand pattern. The practical capacity concept estimates how much more demand volume could be added to a fixed demand pattern by allowing the additional demand to deviate from the fixed demand pattern, while the ultimate capacity concept estimates the maximum network capacity by allowing all users in the network to choose both destination and route. For zonal development potential and equilibrium network capacity analysis, Yang *et al*. [54] suggested using the equilibrium trip distribution/assignment with variable destination costs (ETDA-VDC) of Oppenheim [55], which is the practical network capacity model in our paper, to study the activity characteristics of individual trip-producing zones. The ETDA-VDC model (or the practical network capacity model) incorporated a destination attractiveness measure to reflect the activity opportunities available in determining the travelers' choice of destinations and routes simultaneously for any given number of trips originating from each origin. These two non-uniform network capacity models not only relax the limitation of a common multiplier for all O–D pairs (i.e., uniform OD growth) in the reserve capacity model, but also provide a higher behavioral richness in modeling the interaction between network capacity and level of service of a road network with consideration of zonal development. For example, the ultimate capacity model can be used to determine the network capacity of a new developed city (or a new town), and the practical capacity model can be used to estimate the additional demand that an existing city can accommodate while preserving the current demand pattern without violating roadway and zonal capacity constraints. Note that the practical usefulness of the ultimate capacity model may be limited because it assumes that the roadway infrastructure is available when making land use decisions (i.e., residential and employment locations). Nevertheless, the ultimate capacity model serves as a theoretical upper bound for estimating the maximum network capacity. For our analysis, we include both alternate network capacity models to estimate the capacity reliability for transportation networks.

3.2. *Network capacity models*

Models for estimating three network capacity concepts (i.e., reserve capacity, ultimate capacity, and practical capacity) of a transportation network are provided.

3.2.1. *Notation.*

A	set of links in the network
N	set of nodes in the network
I	set of all origin nodes, $I \subseteq N$
J	set of all destination nodes, $J \subseteq N$
R	set of routes in the network
R_{ij}	set of routes between origin $i \in I$ and destination $j \in J$
i	an origin node, $i \in I$
j	a destination node, $j \in J$
a	a link in the network, $a \in A$
r	a route, $r \in R_{ij}$
Z	objective function
μ	O–D matrix multiplier for the whole network
C_a	capacity on link a

DOI: 10.1002/atr

v_a	flow on link a
$t_a(v_a)$	travel time on link a
$\bar{q}_{ij}$	existing demand between O–D pair ij
$\tilde{q}_{ij}$	additional demand between O–D pair ij
q_{ij}	total demand between O–D pair ij, $q_{ij} = \bar{q}_{ij} + \tilde{q}_{ij}$
$\mathbf{q}$	O–D demand matrix in vector form
h_r^{ij}	flow on route r between O–D pair ij associated with $\bar{q}_{ij}$
f_r^{ij}	flow on route r between O–D pair ij associated with $\tilde{q}_{ij}$
δ_{ar}^{ij}	1 if link a is on route r from origin $i \in I$ to destination$j \in J$; 0 otherwise
$\bar{o}_i$	existing trip production at origin i
$\tilde{o}_i$	additional trip production at origin i
o_i	total trip production at origin i, $o_i = \bar{o}_i + \tilde{o}_i$
$\mathbf{o}$	trip production in vector form
$o_i^{\max}$	maximum trip production at origin i (a constant)
$\bar{d}_j$	existing trip attraction at destination j
$\tilde{d}_j$	additional trip attraction at destination j
d_j	total trip attraction at destination j, $d_j = \bar{d}_j + \tilde{d}_j$
$d_j^{\max}$	maximum trip attraction at destination j (a constant)
$c_j(d_j)$	cost of destination j
θ	impedance parameter for trip distribution

3.2.2. *Reserve capacity model*

The concept of reserve capacity is defined as the largest multiplier μ applied to a given existing O–D demand matrix that can be allocated to a network without violating the link capacities C_a or exceeding a pre-specified level of service [5]. The method for estimating the network capacity uses a common multiplier to scale all O–D pairs. This network capacity model, with a uniform O–D growth, was used in Chen *et al*. [8–10] as a core component in the reliability assessment procedure for estimating the capacity reliability of a transportation network. Mathematically, finding the reserve capacity μ can be formulated as a bi-level program as follows.

$$\text{Max} \quad \mu \tag{1a}$$

subject to

$$v_a(\mu\mathbf{q}) \le C_a, \forall a \in A \tag{1b}$$

where $v_a(\mu\,\mathbf{q})$ is obtained by solving the following user-equilibrium problem:

$$Min \sum_{a \in A} \int_0^{v_a} t_a(x)dx \tag{1c}$$

subject to

$$\sum_{r \in R_{ij}} f_r^{ij} = \mu q_{ij}, \forall i \in I, j \in J \tag{1d}$$

$$v_a = \sum_{i \in I} \sum_{j \in J} \sum_{r \in R_{ij}} f_r^{ij} \delta_{ar}^{ij}, \forall a \in A \tag{1e}$$

DOI: 10.1002/atr

$$f_r^{ij} \geq 0, \forall i \in I,\ j \in J,\ r \in R_{ij} \tag{1f}$$

Route choice behavior and congestion effect are considered by the lower-level problem, while the upper-level problem determines the maximum O–D matrix multiplier in Equation (1a) subject to the roadway capacity constraints in Equation (1b). The lower-level problem is a standard network equilibrium problem [56] for a given μ value determined from the upper-level problem. The objective function in Equation (1c) is the sum of the integrals of the link performance functions. Equation (1d) is a set of flow conservation constraints. Equation (1e) is the incidence relationship, which expresses the link flows in terms of path flows. Equation (1f) represents the non-negativity condition to ensure a meaningful solution. $v_a(\mu \mathbf{q})$ represents the equilibrium link-flow pattern obtained from solving the lower-level problem for a given existing demand pattern $\mathbf{q}$ (i.e., the O–D demand matrix in vector form) uniformly scaled by μ. The largest value of μ indicates whether the current network has spare capacity or not. For example, if $\mu > 1$, then the network has a reserve (or spare) capacity amounting to $100(\mu - 1)$ per cent of the existing O–D demand matrix $\mathbf{q}$; otherwise, the network is overloaded by $100(1 - \mu)$ per cent of the existing O–D demand matrix $\mathbf{q}$. Note that alternate traffic equilibrium models can also be considered in the lower-level problem (e.g., C-logit model by Zhou *et al.* [57], reliability-based traffic equilibrium model by Chen *et al.* [58], mean-excess traffic equilibrium model by Chen and Zhou [59], and stochastic mean-excess traffic equilibrium model by Chen *et al.* [67]). For a review of traffic equilibrium models under uncertainty, see Zhou and Chen [60].

3.2.3. Network capacity model based on the ultimate capacity concept

The ultimate capacity concept is defined as the maximum throughput the system can handle without violating the roadway and zonal capacity constraints or exceeding a pre-specified level of service. This concept relaxes the common multiplier requirement of the network reserve capacity model by allowing the maximum throughput to be scaled by individual O–D pairs. The network capacity model adopted here is a variant of the network capacity and the level-of-service problem described in Yang *et al.* [54], which integrates a combined distribution–assignment model to determine the maximum zonal trip productions. This concept allows all travelers in the network to choose both destination and route simultaneously to minimize their costs. The network capacity model, with a non-uniform growth based on the ultimate capacity concept, is also a bi-level program. The upper-level problem maximizes the zonal trip productions subject to the roadway and zonal capacity constraints, while the lower-level problem is a combined trip distribution and assignment model. The bi-level program is formulated as follows.

$$\text{Max} \sum_{i \in I} o_i \tag{2a}$$

subject to

$$v_a(\mathbf{o}) \leq C_a, \forall a \in A \tag{2b}$$

$$o_i = \sum_{j \in J} q_{ij}(\mathbf{o}) \leq \mathrm{o}_i^{\max}, \forall i \in I \tag{2c}$$

$$d_j = \sum_{i \in I} q_{ij}(\mathbf{o}) \leq d_j^{\max}, \forall j \in J \tag{2d}$$

DOI: 10.1002/atr

$$o_i \geq 0, \forall i \in I \tag{2e}$$

where $q_{ij}(\mathbf{o})$ and $v_a(\mathbf{o})$ are obtained by solving the combined trip distribution–assignment problem:

$$\text{Min} \quad \sum_{a \in A} \int_0^{v_a} t_a(x)\mathrm{d}x + \frac{1}{\theta} \sum_{i \in I} \sum_{j \in J} q_{ij} \left(\ln q_{ij} - 1 \right) \tag{2f}$$

subject to

$$\sum_{j \in J} q_{ij} = o_i, \ \forall i \in I \tag{2g}$$

$$\sum_{r \in R_{ij}} f_r^{ij} = q_{ij}, \ \forall i \in I, j \in J \tag{2h}$$

$$v_a = \sum_{i \in I} \sum_{j \in J} \sum_{r \in R_{ij}} f_r^{ij} \delta_{ar}^{ij}, \ \forall a \in A \tag{2i}$$

$$q_{ij} \geq 0, \ \forall i \in I, j \in J \tag{2j}$$

$$f_r^{ij} \geq 0, \ \forall i \in I, j \in J, r \in R_{ij} \tag{2k}$$

The upper-level problem determines the maximum total trip productions from all origins in Equation (2a) subject to roadway capacity constraints in Equation (2b), maximum trip production and trip attraction constraints in Equations (2c) and (2d), and non-negativity constraints in Equation (2e). The total number of trips generated in each origin can be determined by translating the maximum development potential that is suitable and available for residential use. For commercial use, the total number of trips attracted to each destination zone can be determined by translating the maximum number of job opportunities, parking capacity, and others. The O–D demand matrix ($q_{ij}(\mathbf{o})$) and equilibrium link-flow pattern ($v_a(\mathbf{o})$) are solved by the lower-level problem. Both destination choice and route choice are simultaneously considered in the combined trip distribution–assignment model. Equations (2g) and (2h) represent the flow conservation constraints. Equation (2i) is the incidence relationship that expresses the link flows in terms of path flows. Equations (2j) and (2k) are the non-negativity conditions for O–D flows and path flows, respectively. The impedance parameter θ for trip distribution in Equation (2f) reflects the sensitivity of network users to travel time from an origin to a destination. As shown by Oppenheim [55], existence and uniqueness of solution can be proved as long as the link cost and functions and destination cost functions are strictly increasing in terms of link flows and O–D flows, respectively.

3.2.4. Network capacity model based on the practical capacity concept

Network capacity with the practical capacity concept is defined as the summation of the current O–D demand and the additional demand that the network can accommodate. This concept, using the network capacity and the level-of-service problem described in Yang *et al*. [54], allows only the additional demand to choose both destination and route based on the attractiveness measures of destinations, while the current demand pattern is preserved (i.e., only route choice is allowed). To

DOI: 10.1002/atr

choose the destination, travelers consider both cost of traveling to the destination and cost at the destination. In this model, the upper-level problem maximizes the additional zonal trip productions subject to the roadway and zonal capacity constraints, while the lower-level problem is a combined trip distribution–assignment model with variable destination costs. The bi-level program is formulated as follows.

$$\text{Max} \quad \sum_{i \in I} \tilde{o}_i \tag{3a}$$

subject to

$$v_a(\mathbf{o}) \leq C_a, \ \forall a \in A \tag{3b}$$

$$\tilde{o}_i = \sum_{j \in J} \tilde{q}_{ij}(\mathbf{o}) \leq o_i^{\max} - \bar{o}_i, \ \forall i \in I \tag{3c}$$

$$\tilde{d}_j = \sum_{i \in I} \tilde{q}_{ij}(\mathbf{o}) \leq d_j^{\max} - \bar{d}_j, \ \forall j \in J \tag{3d}$$

$$\tilde{o}_i \geq 0, \ \forall i \in I \tag{3e}$$

where $\tilde{q}_{ij}(\mathbf{o})$ and $v_a(\mathbf{o})$ are obtained by solving the combined trip distribution–assignment problem with variable destination costs:

$$\text{Min} \quad \sum_{a \in A} \int_0^{v_a} t_a(x) \mathrm{d}x + \frac{1}{\theta} \sum_{i \in I} \sum_{j \in J} \tilde{q}_{ij} \left(\ln \tilde{q}_{ij} - 1 \right) + \sum_{j \in J} \int_0^{\sum_{i \in I} \left(\tilde{q}_{ij} + \bar{q}_{ij} \right)} c_j(y) \mathrm{d}y \tag{3f}$$

subject to

$$\sum_{j \in J} \tilde{q}_{ij} = \tilde{o}_i, \ \forall i \in I \tag{3g}$$

$$\sum_{r \in R_{ij}} h_r^{ij} = \bar{q}_{ij}, \ \forall i \in I, j \in J \tag{3h}$$

$$\sum_{r \in R_{ij}} f_r^{ij} = \tilde{q}_{ij}, \ \forall i \in I, j \in J \tag{3i}$$

$$v_a = \sum_{i \in I} \sum_{j \in J} \sum_{r \in R_{ij}} \left(f_r^{ij} + h_r^{ij} \right) \delta_{ar}^{ij}, \ \forall a \in A \tag{3j}$$

$$\tilde{q}_{ij} \geq 0, \ \forall i \in I, j \in J \tag{3k}$$

$$f_r^{ij} \geq 0, \forall i \in I, j \in J, r \in R_{ij} \tag{3l}$$

$$h_r^{ij} \geq 0, \ \forall i \in I, j \in J, r \in R_{ij} \tag{3m}$$

The upper-level problem determines the maximum additional productions from all origins in Equation (3a) subject to roadway capacity constraints in Equation (3b), maximum trip production and trip attraction (excluding the current trip productions and attractions) constraints in Equations (3c) and (3d), and non-negativity constraints in Equation (3e). Similar to the ultimate capacity concept, the total number of trips generated in each origin zone can be determined by translating the maximum amount of vacant land that is suitable and available for residential use. For commercial use, the total number of trips attracted to each destination zone can be determined by translating the maximum number of job opportunities, parking capacity, and others (see Yang *et al*. [54] for further discussions on this issue using demand–supply equilibrium concept). The additional O–D demands ($\tilde{q}_{ij}(\mathbf{o})$) and equilibrium link-flow pattern ($v_a(\mathbf{o})$) are solved in the lower-level problem together with the current O–D demands (or background traffic demands), which are assigned to the network in the deterministic user-equilibrium manner. On the other hand, the additional demand or the traffic growth at each origin zone is distributed among various destination zones by a multinomial logit model, depending on the O–D travel times and the destination costs. The destination cost is an increasing function of the total number of trips attracted to destination j (i.e., $d_j = \sum_{i \in I} \left(\tilde{q}_{ij} + \bar{q}_{ij} \right)$). Equation (3f) is the objective function of the ETDA-VDC model given by Oppenheim [55]. Equations (3g) to (3i) are the flow conservation constraints for the additional O–D demands, the existing path flows, and the additional path flows, respectively. Equation (3j) is the incidence relationship that expresses the link flows in terms of path flows. Equations (3k) to (3m) are the non-negativity conditions for the additional O–D demands, the existing path flows, and the additional path flows, respectively. The impedance parameter θ for trip distribution in Equation (3f) reflects the sensitivity of network users to travel time from an origin to a destination. The existence and uniqueness of the solution in terms of link flows and O–D flows have been proved by Oppenheim [55].

3.3 Capacity reliability measure

Reliability is defined as the probability that the system of interest has the ability to perform an intended function or goal. It can be formulated as the determination of the (supply) capacity of the system to meet certain (demand) requirements [19]. Therefore, the reliability of the system can be defined as an adequacy problem to determine whether the network capacity is sufficient to accommodate the required demand. Capacity reliability in a transportation network explicitly considers the uncertainties associated with the link capacities by treating roadway capacities as continuous quantities subject to routine degradation due to physical and operational factors. Sources of capacity uncertainty include weather conditions, traffic incidents, work zones and construction activities, traffic management and control, etc. (also refer to Chen *et al*. [59] for a detailed description of the sources of uncertainty from both supply and demand). Let NetCap be the network capacity determined by the network capacity problem for a uniform O–D growth (the reserve capacity concept) or a non-uniform growth (the ultimate and practical capacity concepts). Let θ_D be a required demand level. The capacity reliability that can satisfy θ_D is given as

$$R(\theta_D) = \Pr(\text{NetCap} \geq \theta_D) \tag{4}$$

This probability estimates how reliable the existing network is with degradable links in accommodating a required demand level θ_D. The boundary conditions must satisfy the following cases:

(1) $R(\theta_D = 0) = 1.0$
(2) $R(\theta_D = \infty) = 0.0$

The system is 100% reliable when there is no demand and 0% reliable when the demand is infinite.

4. CAPACITY RELIABILITY EVALUATION PROCEDURE

The capacity reliability evaluation procedure adopted in this study is based on the Monte Carlo simulation framework developed by Chen *et al*. [9]. It consists of three main modules: (1) random variate generation (RVG) module, (2) network capacity module, and (3) reliability calculation module. These three modules are synthesized together in a Monte Carlo simulation framework to evaluate the capacity reliability of a transportation network. A flowchart of the procedure is given in Figure 1, and the steps are described as follows.

Step 1: Initialize parameter values: (1) distribution of C_a, (2) initial demand level θ_D and $\Delta\theta_D$, and (3) number of simulations M.
Step 2: Set sample number $m = 1$.
Step 3: Generate values of link capacities according to the distribution properties specified in Step 1: $C^m = \{\ldots, C_a, \ldots\}$.

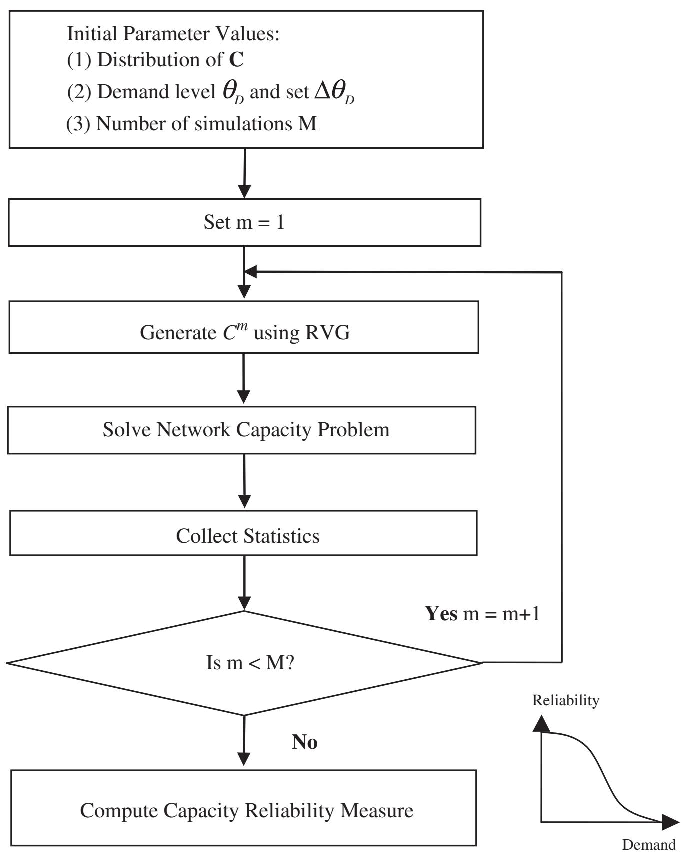


Figure 1. Flowchart of the capacity reliability evaluation procedure.

DOI: 10.1002/atr

Step 4: Solve the network capacity problem for a given realization of link capacities using: (1) the network capacity model based on the reserve capacity concept, (2) the network capacity model based on the ultimate capacity concept, or (3) the network capacity model based on the practical capacity concept.
Step 5: Collect statistics.
Step 6: If sample number m is less than the required sample size M, then increase the sample number by one increment ($m = m + 1$) and go to Step 3. Otherwise, go to Step 7.
Step 7: Compute the capacity reliability measure.

Step 3: Random variate generation procedure. We adopt the RVG procedure developed by Chang *et al.* [61], which is capable of generating multivariate, non-normal, and correlated random variables. Consider that the original random capacities are non-normal and correlated with the vector of mean $\overline{\mathbf{C}}$ and a covariance matrix Cov(C). The following steps are performed to generate multivariate random variates.

(1) Transform the original random capacities to the standard normal condition $C \sim N(\mathbf{p}, \sum)$, where $\mathbf{p}$ and $\sum$ are the mean and correlation matrix of a standard normal.
(2) Decompose the standard normal correlation matrix to its corresponding eigensystem $\sum = \mathbf{V}\Lambda\mathbf{V}^T$, where $\mathbf{V}$ is the eigenvector matrix and $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_a)$ is a diagonal matrix of the eigenvalues.
(3) Generate $\mathbf{Y} = (Y_1, Y_2, \ldots Y_a)^T$ as independent standard normal variates.
(4) Compute the correlated normal random variates given by $\mathbf{C} = \mathbf{p} + \mathbf{V}\Lambda^{1/2}\mathbf{Y}$.
(5) Invert the transformation from standard normal to the original space.

Step 4: Solution procedures for solving the bi-level network capacity models. Bi-level network capacity models are generally difficult to solve because evaluation of the upper-level objective function requires solving the lower-level subprogram. In Step 4, we use an incremental assignment-based procedure to solve the network capacity model based on the reserve capacity concept and a genetic algorithm (GA) combined with a partial linearization algorithm for solving the network capacity models based on the ultimate and practical capacity concepts.

Step 4: (1) Incremental assignment-based procedure. The incremental assignment-based procedure for solving the reserve capacity model can be summarized as follows:

(1) Incremental setting. Select an appropriate $\mu^{(1)}$ applied to a given existing O–D demand matrix so that $\mu^{(1)}\mathbf{q}$ can be allocated to a network without violating the link capacities C_a and set $n = 1$. Then, set an appropriate incremental amount δ.
(2) Initialization. For a given $\mu^{(n)}\mathbf{q}$, perform all-or-nothing assignment based on $t_a = t_a(0)$. This yields $\{v_a^1\}$. Set the counter $k = 1$.
(3) Update. Set $t_a^k = t_a\left(v_a^k\right), \forall a$.
(4) Direction finding. Perform all-or-nothing assignment based on t_a^k. This yields a set of auxiliary flow $\{y_a^k\}$.
(5) Line search. Find α_k that solves

$$\mathrm{Min} \quad Z = \sum_{0 \le \alpha_k \le 1} \int_0^{\substack{ak \\ v}+\alpha_k\left(y_a^k - v_a^k\right)} t_a(w) \ \mathrm{d}w$$

(6) Move. Set $v_a^{k+1} = v_a^k + \alpha_k\left(y_a^k - v_a^k\right)$.
(7) Convergence test. If the convergence criterion is met, stop (the current solution is the set of equilibrium link flows v_a^k) otherwise, set $k = k + 1$ and go to 3.
(8) Constraint check. For each link $a \in A$, if $v_a^k < C_a$, let $\mu^{(n+1)} = \mu^{(n)} + \delta$ and $n = n + 1$, return to 2. Otherwise, let $\mu_{\max} = \mu^{(n)}$ and stop.

DOI: 10.1002/atr

Step 4: (2) Genetic algorithm procedure

In the GA, the best solution is chosen from a number of possible solutions. The approach of the GA implementation for the network capacity problems is that decision variables in the upper-level problem are coded as chromosomes and the chromosomes evolve through successive iterations called generations. During each generation, the chromosomes are evaluated using some measure of fitness, which is calculated by solving the lower-level problem. In the GA evolution process, there are three operators including reproduction, crossover, and mutation operators. Reproduction is the selection process of the chromosomes from the population set for mating purposes. The crossover operator gets genetic material from the previous generation to the subsequent generation. Mutation is a process that introduces a certain amount of randomness to the search and enables the search to find solutions that crossover alone may not encounter. The GA procedure for solving the network capacity models based on the ultimate and practical capacity concept can be summarized as follows:

(1) Select the population size (*NP*) and maximum number of generations (N). Define the crossover rate (p_c) and mutation rate (p_m).
(2) Initialization: Code the decision variables in terms of trip production for all origins $(\ldots,o_i,\ldots)$. Then, set the generation number $n = 1$.
(3) Calculate the fitness for individual chromosomes by solving the lower-level problem using Equations (2f) to (2k) for the ultimate capacity concept or Equations (3f) to (3m) for the practical capacity concept.
(4) Carry out the GA evolution process:
4.1 Reproduce the population according to the fitness function values.
4.2 Conduct the crossover operator through a random choice with p_c.
4.3 Conduct the mutation operator through a random choice with p_m.
This step yields a new population at generation $n + 1$.
(5) If $n = N$, the sample with the highest fitness is adopted as an approximate optimal solution of the problem. Otherwise, set $n = n + 1$ and go to 3.

(a) Partial linearization algorithm for the ultimate capacity model

The solution algorithm is based on the partial linearization method, which is a descent algorithm for continuous optimization problems [62]. Only a part of the objective is linearized in each iteration. A search direction is obtained from the solution of a convex auxiliary problem defined by an approximation of the original objective through the first-order approximation of the additive part of the objective function. A line search is made in the direction obtained with respect to the objective function, and the resulting step size defines a new solution with a reduced objective value. The procedure of solving the combined trip distribution–assignment model is as follows:

(1) Find a set of feasible flows $\left\{v_a^k\right\}, \left\{q_{ij}^k\right\}$, and set $k = 1$.
(2) Calculate link costs $t_a^k = t_a\left(v_a^k\right), \forall a$.
(3) Find the search direction:
3.1 Calculate the minimum travel-time path from each origin i to all destinations based on $\{t_a^k\}$. Let c_{ij}^k denote the minimum travel time from origin i to destination j
3.2 Determine the auxiliary O–D flow by applying a logit distribution model, that is,

$$d_{ij}^k = \frac{o_i e^{-\theta c_{ij}^k}}{\sum_m e^{-\theta c_{im}^k}}, \forall i \in I, j \in J$$

3.3 Assign d_{ij}^k to the minimum travel-time path between origin i and destination j. This also yields a link-flow pattern $\{y_a^k\}$.
(4) Determine the step size. Find α_k that minimizes the function

$$\underset{0\le\alpha_k\le 1}{Min} \quad \sum_{a\in A}\int_0^{v_a^k+\alpha_k\left(y_a^k-v_a^k\right)} t_a(x)\mathrm{d}x+\frac{1}{\theta}\sum_{i\in I}\sum_{j\in J}\left(q_{ij}^k+\alpha_k\left(d_{ij}^k-q_{ij}^k\right)\right)\times\left(\ln\left(q_{ij}^k+\alpha_k\left(d_{ij}^k-q_{ij}^k\right)\right)-1\right)$$

(5) Update O–D and link flows. Set

$$q_{ij}^{k+1}=q_{ij}^k+\alpha_k\left(d_{ij}^k-q_{ij}^k\right),\forall i\in I,j\in J$$
$$v_a^{k+1}=v_a^k+\alpha_k\left(y_a^k-v_a^k\right),\forall a\in A$$

(6) Test the convergence. If the convergence criterion is not achieved, set $k=k+1$ and go to 2. Otherwise, terminate: the solution is $\left\{v_a^{k+1}\right\},\left\{q_{ij}^{k+1}\right\}$.

(b) Partial linearization algorithm for the practical capacity model

The partial linearization algorithm given by Yang *et al*. [54] can be summarized as follows:

(1) Determine an initial value $\left\{v_a^k\right\},\left\{\tilde{q}_{ij}^k\right\}$ and set $k=0$.
(2) Calculate link cost $t_a^k=t_a\left(v_a^k\right)$, the minimum travel time from origin i to destination j, $c_r^{ij(k)}$, and destination cost $c_j^k=c_j\left(\sum_i\left(\tilde{q}_{ij}^k+\bar{q}_{ij}\right)\right)$.
(3) Find the descent direction by obtaining $\tilde{a}_r^{ij(k)}$ that minimizes program P1, which is

$$\text{P1}:\text{Min}\quad Z_1=\sum_{i\in I}\sum_{j\in J}\sum_{r\in R_{ij}}\left[c_r^{ij(k)}+\frac{1}{\theta}\ln\left(\tilde{q}_{ij}^k\right)+c_j\left(\sum_i\left(\tilde{q}_{ij}^k+\bar{q}_{ij}\right)\right)\right]\tilde{a}_r^{ij(k)}$$

subject to

$$\sum_{j\in J}\sum_{r\in R_{ij}}\tilde{a}_r^{ij(k)}=\tilde{o}_i^{(k)},\forall i\in I,\qquad \tilde{a}_r^{ij(k)}\ge 0,\ \forall i\in I,j\in J,r\in R_{ij}$$

and $\bar{a}_r^{ij(k)}$ that minimizes program P2, which is

$$P2:\text{Min}\quad Z_2=\sum_{i\in I}\sum_{j\in J}\sum_{r\in R_{ij}}c_r^{ij(k)}\bar{a}_r^{ij(k)}$$

subject to

$$\sum_{r\in R_{ij}}\bar{a}_r^{ij(k)}=\bar{q}_{ij},\forall i\in I,j\in J,\qquad \bar{a}_r^{ij}\ge 0,\ \forall i\in I,j\in J,r\in R_{ij}$$

Then, set $y_a^k=\sum_{i\in I}\sum_{j\in J}\sum_{r\in R_{ij}}\left(\tilde{a}_r^{ij(k)}+\bar{a}_r^{ij(k)}\right)\delta_{ar}^{ij},\forall a\in A$

$$\tilde{d}_{ij}^k=\sum_{r\in R_{ij}}\tilde{a}_r^{ij(k)},\forall i\in I,j\in J$$

(4) Determine the step size. Find α_k that minimizes the function

$$\underset{0\le\alpha_k\le 1}{\text{Min}}\quad Z(\alpha)=\sum_{a\in A}\int_0^{v_a^k+\alpha_k\left(y_a^k-v_a^k\right)}t_a(x)\mathrm{d}x+\frac{1}{\theta}\sum_{i\in I}\sum_{j\in J}\left(\tilde{q}_{ij}^k+\alpha_k\left(\tilde{q}_{ij}^k-\tilde{d}_{ij}^k\right)\right)\times$$
$$\left(\ln\left(\tilde{q}_{ij}^k+\alpha_k\left(\tilde{q}_{ij}^k-\tilde{d}_{ij}^k\right)\right)-1\right)+\sum_{j\in J}\int_0^{\sum_{i\in I}\left(\tilde{q}_{ij}^k+\bar{q}_{ij}+\alpha_k\left(\tilde{q}_{ij}^k-\tilde{d}_{ij}^k\right)\right)}c_j(y)\mathrm{d}y$$

DOI: 10.1002/atr

(5) Update O–D and link flows. Set

$$\tilde{q}_{ij}^{k+1} = \tilde{q}_{ij}^{k} + \alpha_k\left(\tilde{d}_{ij}^{k} - \tilde{q}_{ij}^{k}\right), \forall i \in I, j \in J$$
$$v_a^{k+1} = v_a^{k} + \alpha_k\left(y_a^{k} - v_a^{k}\right), \forall a \in A$$

(6) Test the convergence test. The condition of convergence is

$$\left|\frac{\tilde{q}_{ij}^{k}}{o_i} - \frac{\exp\left(-\theta\left(c_{ij}^{k} + c_j\left(\sum_{i\in I}\tilde{q}_{ij}^{k}\right)\right)\right)}{\sum_{m\in J}\exp\left(-\theta\left(c_{im}^{k} + c_m\left(\sum_{i\in I}\tilde{q}_{im}^{k}\right)\right)\right)}\right| \leq \varepsilon, \forall i \in I$$

If convergence is not achieved, set $k=k+1$ and go to 2. Otherwise, terminate and use the solution $\left\{v_a^{k+1}\right\}, \left\{\tilde{q}_{ij}^{k+1}\right\}$.

Remark 1

In Step 4, a GA procedure is adopted to solve the ultimate and practical network capacity models. Typical GA operators (i.e., reproduction, crossover, and mutation) are performed to evolve the chromosomes (or trip production from each zone) to obtain better solutions. Detailed descriptions of the GA implementation can be found in Gen and Cheng [63] and Kasikitwiwat's dissertation [64].

Remark 2

The bi-level network capacity models are intrinsically non-convex and hence difficult to solve for a global optimum [65]. Heuristics are typically developed to tackle this class of problems. In Yang *et al.* [54], they developed a successive linear programming (SLP) approach that iteratively solves the lower-level distribution–assignment problem using a partial linearization method and the upper-level trip production problem using a localized linear approximation of the upper-level constraints to solve the resulting linear program. In this study, we developed a GA procedure to solve the ultimate and practical network capacity models. Comparisons using the same test network (see Figure 2 in Section 5) with different initial solutions were conducted in Kasikitwiwat [64] but not reported here for conciseness. The results show that the GA procedure is capable of finding near-optimal solutions (or at least not worse than the SLP approach). For the reserve capacity model, we adopted the incremental assignment-based procedure embedded with a convex combination method (also known as the Frank–Wolfe algorithm) to determine the largest multiplier μ. Care should be exercised in choosing a judicious incremental amount δ. In general, δ should be chosen sufficiently small to increase the estimation accuracy of $\mu_{\max}$.

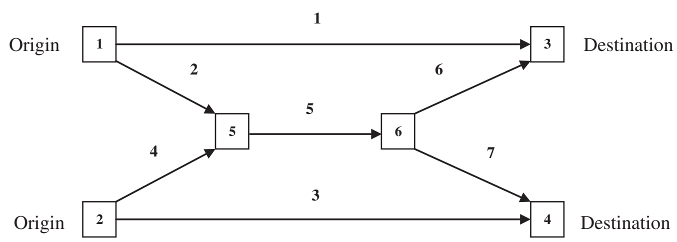


Figure 2. Test network.

DOI: 10.1002/atr

Remark 3

For each set of link capacities generated, a network capacity problem (e.g., reserve capacity, ultimate capacity, and practical capacity) is solved to determine the maximum network capacity. In Step 5, various statistics (e.g., link-flow pattern, O–D travel-time pattern, objective value for each capacity concept, etc.) are collected to compute the capacity reliability measures in Step 7.

5. NUMERICAL RESULTS

In this section, we present some numerical results using the reliability evaluation procedure described earlier. We apply the procedure to a simple network given in Figure 2. The network consists of six nodes, seven links, two origins, two destinations, and four O–D pairs. The link travel-time function used is the standard Bureau of Public Road function:

$$t_a = t_a^f \left(1 + 0.15 \left(\frac{v_a}{C_a} \right)^4 \right)$$

where v_a, t_a^f, and C_a are the flow, free-flow travel time, and random capacity on link a, respectively. Link characteristics are provided in Table II. In the absence of link degradation data, we assume a uniform distribution with an upper bound and a lower bound to generate the random link capacities. Note that the RVG procedure in Step 3 is capable of generating multivariate random capacities that preserves the marginal distribution of random capacities and their correlation structure if link degradation data were available (see Chen *et al.* [9] for results with correlation). As in Table II, the means and standard deviations resulting from 200 Monte Carlo simulations are close to the theoretical values.

5.1. Network capacity analysis

This section compares the network capacity results using three different network capacity models. For the reserve capacity concept, a pre-determined O–D demand pattern is required. This pattern remains fixed while the network capacity model determines the maximum O–D demand multiplier that can be accommodated by assigning the scaled O–D demand matrix to the network using the user-equilibrium assignment method without exceeding the roadway capacity constraints. For the ultimate capacity concept, O–D flows do not need to follow a pre-determined O–D demand pattern because a combined trip distribution–assignment model is adopted in the bi-level program to simultaneously determine both destination choice and route choice of all users. However, physical roadway and zonal capacity constraints are required in the upper-level problem to constrain the maximum total trip production from all origins. Network capacity under this model is treated as the ultimate capacity, which is the upper bound because all network users can choose both destination and route simultaneously. For the practical capacity concept, an existing O–D demand is

Table II. Link free-flow travel times and statistical properties of link capacities.

Link #	Free-flow travel time	Lower bound	Upper bound	Theoretical		Estimated	
				Mean	SD	Mean	SD
1	10.00	75.00	100.00	87.50	7.22	87.33	7.28
2	4.00	60.00	80.00	70.00	5.77	69.91	5.77
3	12.00	60.00	80.00	70.00	5.77	69.98	5.82
4	4.00	37.50	50.00	43.75	3.61	43.52	3.53
5	5.00	90.00	120.00	105.00	8.66	105.39	9.12
6	5.00	37.50	50.00	43.75	3.61	43.80	3.56
7	4.00	37.50	50.00	43.75	3.61	43.71	3.82

DOI: 10.1002/atr

preserved (i.e., only route choice is allowed). Only the additional demands or travelers can choose both destination and route. To choose the destination, network users consider the cost of destination in addition to the cost of traveling to the destination. Therefore, a combined distribution–assignment model with variable destination costs is applied only to the additional demands. Function and data of the destination costs are shown in Table III.

The complete results of these three models are provided in Table IV and graphically presented in Figure 3 to highlight the distinct feature of the three network capacity concepts. For the reserve capacity model, the maximum network capacity is 227.92. The pre-determined O–D demand pattern is preserved with a uniform growth rate (or maximum O–D demand multiplier) of 2.072 for all O–D pairs. For the ultimate network capacity model that allows both destination choice and route choice for all network users, the maximum network capacity with an impedance parameter value of 0.5 is 262.54, an increase of 15.19% compared with the reserve capacity model. For the practical network capacity model that only allows the additional demands to have both destination choice and route choice, the maximum network capacity with an impedance parameter value of 0.5 is 257.58, an increase of 13.08% compared with the reserve capacity model. It is 1.89% less than the network capacity model using the ultimate capacity concept. Moreover, it is observed that the O–D patterns resulting from the three capacity models are quite different. For example, the results indicate that the flows on O–D (1–3), O–D (1–4), and O–D (2–3) from the practical and ultimate capacity models are higher than those in the reserve capacity model, while flows on O–D (2–4) show the opposite. However, the net increase in network capacity is larger in the network capacity models that allow partial and full destination choice. Network capacity estimated with the reserve capacity concept is underestimated because the same increase rate is applied to all O–D pairs. If flows on one O–D pair cause some links in the network to reach the roadway capacities, flows on other O–D pairs will also stop increasing. In this network, link 3, which serves O–D (2–4), is the bottleneck link. The practical and ultimate network capacity models can achieve more network capacity by having more demands in O–D (1–3) (link 1), O–D (1–4) (links 2–5–7), and O–D (2–3) (links 4–5–6) because the capacities of these routes serving these O–D pairs are underutilized in the reserve capacity model. Note that the volume-to-capacity (V/C) ratio of link 4, for the practical network capacity model, is less than the V/C ratio for the reserve capacity model, because link 4 is also used by flows on O–D (2–4) in the reserve capacity model.

In general, using the reserve capacity concept will underestimate the maximum network capacity because of the requirement of preserving the pre-determined O–D pattern, while the network capacity with the ultimate capacity concept will overestimate the maximum network capacity because it does not account for maximum acceptable cost. Therefore, the network capacity with the practical capacity concept, considering background traffic be preserved, may be more practical to achieve.

5.2. Capacity reliability analysis

The capacity reliability measure is calculated via the capacity reliability evaluation procedure. When the capacity of every link is fixed at the upper bound of the uniform distribution (i.e., non-degraded capacity), the maximum network capacity under the non-uniform O–D growth model based on the ultimate capacity concept is 262.54. This value serves as the upper bound for the degradable network; a travel demand level greater than 262.54 cannot be satisfied. The maximum network capacity under the non-uniform O–D growth model based on the practical capacity concept is 257.58. This value serves as the upper bound for the degradable network of the practical capacity concept, and the upper bound of the reserve capacity concept is 227.92.

Table III. Destination cost data $c_j(d_j) = \alpha_j d_j^{\beta_j} - m_j$ used in the practical capacity model.

Destination	m_j	α_j	β_j
3	1.20	0.15	0.25
4	1.50	0.10	0.25

DOI: 10.1002/atr

Table IV. Comparison of three network capacity models.

	System capacity	Trip production		Origin–destination flow			
				O–D (1–3)	O–D (1–4)	O–D (2–3)	O–D (2–4)
Reserve capacity concept		O_1	O_2				
Inputs	—	—	—	40.00	10.00	10.00	50.00
Flow results (multiplier = 2.072)	227.92	103.60	124.32	82.88	20.72	20.72	103.60
Ultimate capacity concept		O_1^{max}	O_2^{max}				
Inputs	—	150	150	—	—	—	—
Flow results	262.54	138.01	124.53	99.97	38.04	43.75	80.78
Practical capacity concept		O_1^{max}	O_2^{max}				
Inputs (existing traffic pattern)	—	150	150	40.00	10.00	10.00	50.00
Flow results	257.58	137.83	119.75	100.80	37.03	31.36	88.39

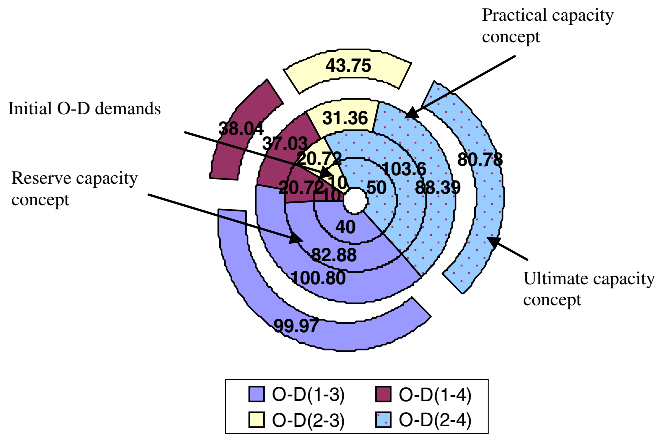


Figure 3. O–D demand patterns and link volume-to-capacity ratios of three network capacity models.

Capacity reliability is dependent on a number of factors. The results are depicted in Figure 4 as a function of demand levels. For demand levels less than 175, the network is 100% reliable for all network capacity models. As the demand levels increase, the capacity reliability starts to deteriorate and completely fails when the demand level is beyond 262.54 (the maximum network capacity under the non-degradable condition). Moreover, capacity reliability is also dependent on the network capacity concept. The capacity reliability calculated using different network capacity concepts deteriorates at different rates. With a demand level of 185 as an example, the capacity reliability using the reserve capacity concept is 85% reliable but 100% reliable for ultimate and practical capacity concepts.

Similarly, using the 85th percentile reliability as a criterion, the capacity reliability using the non-uniform O–D growth model can accommodate a demand level up to 212 for the method considering the practical capacity concept, 217 considering the ultimate capacity concept, and only 185 for the reserve capacity concept. Because the non-uniform O–D growth model corrects the underestimated biased problem in the uniform O–D growth model, the capacity reliability using the non-uniform O–D growth model can accommodate a higher demand level than the capacity reliability using the uniform O–D growth model. As can be seen in Figure 4, the difference of reliability curves between the reserve capacity model and non-uniform models (ultimate and practical capacity concepts) is significant, while the difference of reliability curves between the ultimate capacity concept and the practical concept is not as significant.

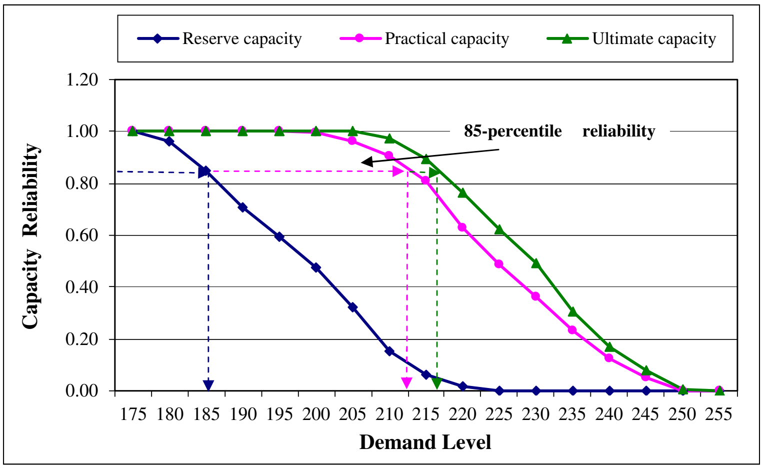


Figure 4. Capacity reliability for different network capacity concepts.

DOI: 10.1002/atr

To obtain more details in the capacity reliability analysis, the statistical values of link flows, O–D costs, and the link-flow distribution are investigated. Tables V and VI present the statistical values of O–D cost and link flows for different network capacity concepts. Statistical values of O–D costs are slightly different with different capacity concepts. For O–D (1–3), O–D (1–4), and O–D (2–3), the mean O–D cost values resulting from the ultimate and practical capacity concepts are higher than O–D costs from the reserve capacity concept, while the mean O–D cost between O–D (2–4) shows the opposite. This is because the O–D pattern for the reserve capacity has the highest demand on O–D (2–4). For the standard deviation of O–D costs, the values are low. The highest standard deviation of O–D cost is the value for O–D (1–3) from the reserve capacity concept. Note that link 1 used in O–D (1–3) has high capacity and free-flow travel time.

For link flows, the statistical values from various concepts of network capacity are different. In most links, the minimum, maximum, mean, and standard deviation of link flows (from 200 realizations), from the ultimate and practical capacity concepts, are higher than those from the reserve capacity concept, for example links 2 and 5. The increase of flows on O–D (1–4) and O–D (2–3) can result in the increase of flows on links 2 and 5 when the network capacity is estimated from the practical

Table V. Statistical values of O–D costs for different network capacity concepts.

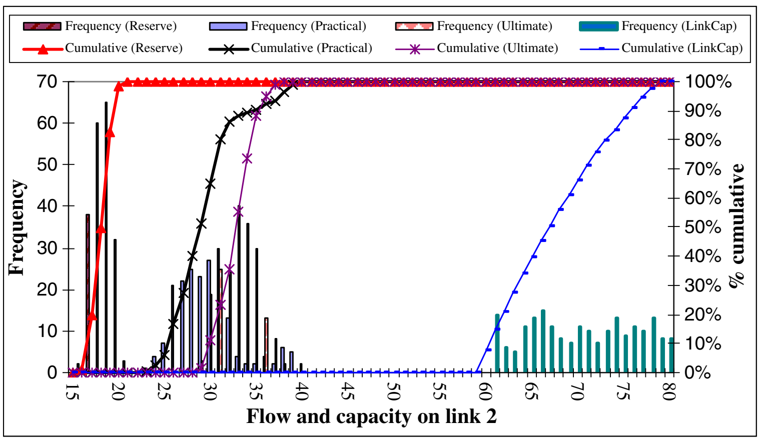
O–D	1–3	1–4	2–3	2–4
Reserve capacity concept				
Min	10.27	13.24	14.23	13.72
Max	11.50	13.61	14.63	13.80
Mean	10.76	13.43	14.46	13.79
Standard deviation	0.30	0.08	0.08	0.01
Practical capacity concept				
Min	10.75	13.28	14.28	13.41
Max	11.56	13.85	14.95	14.04
Mean	11.43	13.64	14.66	13.76
Standard deviation	0.15	0.12	0.14	0.08
Ultimate capacity concept				
Min	11.45	13.25	14.69	13.59
Max	11.50	13.64	15.28	13.80
Mean	11.49	13.44	15.04	13.79
Standard deviation	0.01	0.08	0.12	0.02

Table VI. Statistical values of link flows for different network capacity concepts.

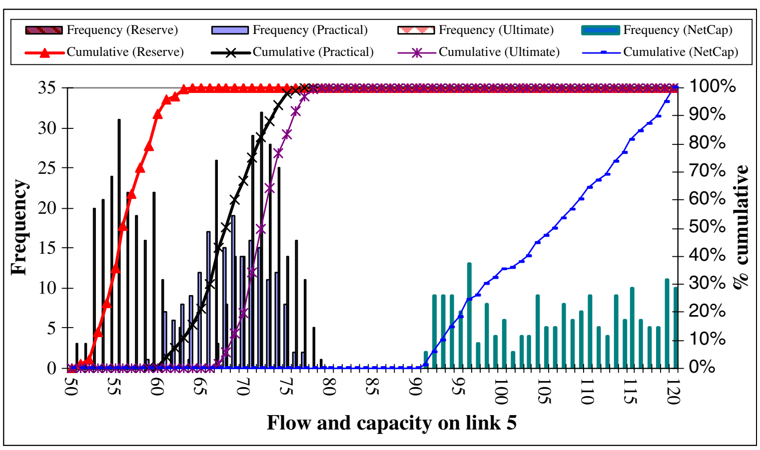
Link #	1	2	3	4	5	6	7
Reserve capacity concept							
Min	63.83	15.96	60.02	33.91	50.17	15.96	33.91
Max	80.68	20.17	79.96	43.50	63.05	20.17	43.50
Mean	72.12	18.03	69.94	38.24	56.28	18.03	38.24
Standard deviation	4.05	1.01	5.79	2.25	2.74	1.01	2.25
Practical capacity concept							
Min	74.93	22.61	59.31	24.23	58.81	21.79	34.32
Max	99.44	39.53	79.49	47.05	76.52	32.23	48.20
Mean	86.12	29.23	69.61	38.80	68.02	27.10	40.92
Standard deviation	6.80	3.46	5.73	3.50	3.89	2.26	2.77
Ultimate capacity concept							
Min	75.00	28.12	59.93	34.63	66.54	33.94	30.09
Max	99.97	37.68	79.63	46.70	78.05	43.36	40.92
Mean	87.22	32.65	69.85	39.51	72.16	37.91	34.26
Standard deviation	7.22	1.99	5.76	2.26	2.55	1.75	2.12

and ultimate capacity concepts. Thus, the network capacity and capacity reliability can be increased. The distributions of link flow and link capacity for links 2 and 5 are graphically shown in Figure 5. The link flow curves resulting from the ultimate and practical capacity models are closer to the link capacity curve than the link flow curve resulting from the reserve capacity model. These distributions indicate that the ultimate and practical capacity concepts utilize the capacities of links 2 and 5 better than the reserve capacity concept.

The capacity reliability assessment results from this specific network indicate that the network capacity models indeed can have a significant impact on the capacity reliability. The general trend is that the reserve capacity model typically underestimates network capacity and hence results in a lower capacity reliability measure for a given demand level. On the other hand, the ultimate capacity model typically overestimates network capacity because it assumes that the O–D demand pattern is completely flexible (i.e., all users are allowed to freely choose both destination and route in the network). This results in overestimating the capacity reliability. The practical capacity model assumes that only the additional demand can deviate from the fixed demand pattern (i.e., only part of the demand can have both destination choice and route choice, while the fixed demand part only has route choice because it must preserve the existing O–D demand pattern). The capacity reliability measure estimated using the practical capacity model is in between those estimated by the reserve capacity model and the ultimate capacity model. It serves as a compromise in determining the capacity reliability of an existing city with a fixed demand pattern, and some spare capacity still exists in the current transportation network.



(a) Flow and capacity on link 2



(b) Flow and capacity on link 5

Figure 5. Distributions of link capacity and link flow for different network capacity concepts.

J. Adv. Transp. 2013; **47**:79–104
DOI: 10.1002/atr

6. CONCLUSIONS

In this paper, alternate capacity reliability measures were proposed for transportation networks. The core of the alternate measures is the network capacity models used to calculate the capacity reliability measures. Three different concepts of capacity (i.e., reserve capacity, ultimate capacity, and practical capacity) were used to determine transportation network capacity. These network capacity models were then integrated into the capacity reliability evaluation procedure developed by Chen *et al*. [9] to assess the capacity reliability measures.

In the previous studies, network capacity determination was based on the premise of preserving a pre-determined (or fixed) O–D pattern, and capacity reliability was evaluated for a prescribed level of service. This reserve capacity model is simple and easy to calculate. It is useful for estimating network capacity when zonal land use information is not available. However, it can only capture the changes in demand volume by uniformly scaling all O–D pairs with the same multiplier (i.e., changes in demand pattern are not considered). Note that the demand response due to changes in demand pattern may require a different time scale as compared to the capacity fluctuations. Nevertheless, the common multiplier assumption was relaxed in this study by adopting two non-uniform network capacity models from two different concepts: ultimate capacity for estimating network capacity of a newly developed city and practical capacity for estimating additional network capacity of an existing city. The practical capacity reliability measure estimates how much more demand volume could be added to a fixed demand pattern by allowing the additional demand to deviate from the fixed demand pattern, while the ultimate capacity reliability measure estimates the maximum network capacity by allowing all users in the network to choose both destination and route. These two alternate capacity models not only allowed a non-uniform O–D growth in the spatial distribution of the O–D demand pattern but also corrected the underestimated biased problem in the reserve capacity model and provided a higher behavioral richness in modeling the interaction between network capacity and level of service of a road network with consideration of zonal development. The practical and ultimate capacity reliability measures can complement the existing reserve capacity reliability measure by allowing the analysis of zonal activity allocations in relation with the physical capacity of zonal land use and network characteristics, which enable evaluating both changes in demand volume and variations in demand pattern. These network capacity models can also be used to assess the capacity flexibility of a road network because of external changes in terms of traffic demand level and traffic demand patterns as demonstrated by Chen and Kasikitwiwat [66]. It is defined as the ability of a road network to accommodate changes in both volume and pattern of traffic demand while maintaining a satisfactory level of performance. Both capacity reliability and capacity flexibility are two useful performance measures that can be used to assess the adequacy of a road network under supply uncertainty and demand changes. It would be interesting as a future research to integrate both capacity reliability and capacity flexibility to assess the system uncertainty with both supply and demand simultaneously.

ACKNOWLEDGEMENTS

The authors are grateful to Prof. William Lam (Editor of Journal of Advanced Transportation) and three referees for providing useful comments and suggestions for improving the quality and clarity of the paper. The work of the first author was supported by a CAREER grant from the National Science Foundation of the United States (CMS-0134161) and an Oriental Scholar Professorship Program sponsored by the Shanghai Ministry of Education in China to Tongji University, the work of the second author was supported by the Royal Thai Government Scholarship and the Center of Advanced Studies in Industrial Technology, Faculty of Engineering, Kasetsart University in Thailand, and the work of the third author was supported by the New Century Excellent Talents Program in University (NCET-08-0406).

REFERENCES

1. Ottomanelli M, Wong CK. Modelling uncertainty in traffic and transportation systems. *Transportmetrica* 2010; **7**(1):1–3.
2. Webster FV, Cobbe BM. Traffic signals. Road Research Technical Paper, No. 56, HMSO, London, 1966.

3. Allsop RE. Estimating the traffic capacity of a signalized road junction. *Transportation Research* 1972; **6**(3):245–255.
4. Wong SC. On the reserve capacity of priority junctions and roundabouts. *Transportation Research Part B* 1996; **30**(6):441–453.
5. Wong SC, Yang H. Reserve capacity of a signal-controlled road network. *Transportation Research Part B* 1997; **31**(5):397–402.
6. Ceylan H, Bell MGH. Reserve capacity for a road network under optimized fixed time traffic signal control. *Journal of Intelligent Transportation Systems* 2004; **8**(2):87–99.
7. Sumalee A, Luathep P, Lam WHK, Connors RD. Transport network capacity evaluation and design under demand uncertainty. *Transportation Research Record* 2009; **2090**:93–101.
8. Chen A, Yang H, Lo HK, Tang W. A capacity related reliability for transportation networks. *Journal of Advanced Transportation* 1999; **33**(2):183–200.
9. Chen A, Yang H, Lo HK, Tang W. Capacity reliability of a road network: an assessment methodology and numerical result. *Transportation Research Part B* 2002b; **36**(3):225–252.
10. Chen A, Tatineni M, Lee D-H, Yang H. Effect of route choice models on estimating network capacity reliability. *Transportation Research Record* 2000; **1733**:63–70.
11. Sumalee A, Kurauchi F. Network capacity reliability analysis considering traffic regulation after a major disaster. *Network and Spatial Economics* 2006b; **6**(3):205–219.
12. Lo HK, Tung YK. Network with degradable links: capacity analysis and design. *Transportation Research Part B* 2003; **37**(4):345–363.
13. Gao Z, Song Y. A reserve capacity model of optimal signal control with user equilibrium route choice. *Transportation Research Part B* 2002; **36**(4):313–323.
14. Chootinan P, Wong SC, Chen A. A reliability-based network design problem. *Journal of Advanced Transportation* 2005; **39**(3):247–270.
15. Chen A, Chootinan P, Wong SC. New reserve capacity model of a signal-controlled road network. *Transportation Research Record* 2006; **1964**:35–41.
16. Yim K, Wong SC, Chen A, Wong CK, Lam WHK. A reliability-based land use and transportation optimization model. *Transportation Research Part C* 2011; **19**(2):351–362.
17. Chen A, Zhou Z, Chootinan P, Ryu S Yang C, Wong SC. Transport network design problem under uncertainty: a review and new developments. *Transport Reviews* 2011a; **31**(6):743–768.
18. Chen A, Kim J, Lee S, Kim Y. Stochastic multi-objective models for network design problem under demand uncertainty. *Expert Systems with Applications* 2010a; **37**(2):1608–1619.
19. Ang AHS, Tang WH. *Probability Concepts in Engineering Planning and Design, Vol. I Basic Principle; Vol. II Decision, Risk and Reliability*. Wiley: New York, 1990.
20. Bell MGH, Iida Y. *Transportation Network Analysis*. John Wiley and Sons, Inc: New York, 1997.
21. Nicholson A, Du ZP. Degradable transportation systems: an integrated equilibrium model. *Transportation Research Part B* 1997; **31**(3):209–224.
22. Lam WHK. Special issue: network reliability and transport modeling. *Journal of Advanced Transportation* 1999; **33**(2):121–122.
23. Bell MGH, Cassir C. *Reliability of Transport Networks*. Research Studies Press Ltd.: Baldock, England, 2000.
24. Bell MGH, Iida Y. *The Network Reliability of Transport*. Elsevier Science: Oxford, U.K., 2003.
25. Nicholson A, Dantas A. Proceedings of the 2nd International Symposium on Transportation Network Reliability, Christchurch, New Zealand, 2004.
26. Sumalee A, Kurauchi F. Guest editorial: reliability and emergence issues in transportation network analysis. *Network and Spatial Economics* 2006a; **6**(3):168–172.
27. Murray AT, Grubesic TH. *Critical Infrastructure: reliability and Vulnerability*. Springer-Verlag: Berlin, Heidelberg, 2007.
28. van Zuylen H. Proceedings of the 3rd International Symposium on Transportation Network Reliability, Hague, The Netherlands, 2007.
29. Kurauchi F, Sumalee A. Editorial: evaluation and enhancement of transport network reliability: network equilibrium based approach. *Journal of Advanced Transportation* 2008; **42**(3):209–212.
30. Schmocker J-D, Lo HK. Editorial: special issue on improving network reliability through ITS technologies. *Journal of Intelligent Transportation Systems* 2009; **13**(1):1.
31. Levinson D, Liu H, Bell MGH. Proceedings of the 4th International Symposium on Transportation Network Reliability, Minnesota, USA, 2010.
32. Iida Y, Wakabayashi H. An approximation method of terminal reliability of a road network using partial minimal path and cut set. In *Proceeding of the Fifth WCTR*, Yokohama, Japan, 1989.
33. Asakura Y, Kashiwadanu M. Road network reliability caused by daily fluctuation of traffic flow. In *Proceeding of the 19th PTRC Summer Annual Meeting*, Brighton, 1991; 73–84.
34. Bell MGH, Cassir C, Iida Y, Lam WHK. A sensitivity-based approach to network reliability assessment. In *Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, Jerusalem, 1999; 283–300.
35. Asakura Y. Reliability measures of an origin and destination pair in a deteriorated road network with variable flows. In *Transportation Networks: Recent Methodological Advances*, Bell MGH (ed.). Pergamon Press: Oxford, 1998; 273–287.

36. Chen A, Ji Z, Recker W. Travel time reliability with risk sensitive travelers. *Transportation Research Record* 2002a; **1783**:27–33.
37. Chen A, Ji Z, Recker W. Effect of route choice models on estimation of travel time reliability under demand and supply variations. In *The Network Reliability of Transport*, Bell MGH, Iida Y (eds). Elsevier Science: Pergamon, 2003; 93–118.
38. Lo HK, Luo XW, Siu BWY. Degradable transport network: travel time budget of travelers with heterogeneous risk aversion. *Transportation Research Part B* 2006; **40**(9):792–806.
39. Chen A, Ji Z. Path finding under uncertainty. *Journal of Advanced Transportation* 2005; **39**(1):19–37.
40. Shao H, Lam WHK, Tam ML. A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand. *Network and Spatial Economics* 2006; **6**(3):173–204.
41. Szego G. Measures of risk. *European Journal of Operational Research* 2005; **163**(1):5–19.
42. Heydecker B, Lam WHK, Zhang N. Use of travel demand satisfaction to assess road network reliability. *Transportmetrica* 2007; **3**(2):139–171.
43. Bell MGH, Schmocker JD. *Public transport network reliability: topological effects and the importance of information*. In *Proceedings of the Third International Conference on Transportation and Traffic Studies*, Guilin, China, 2002; 453–460.
44. Yang H, Bell MGH. Traffic restraint, road pricing and network equilibrium. *Transportation Research Part B* 1997; **31**(4):303–314.
45. Ferrari P. Road pricing and network equilibrium. *Transportation Research Part B* 1995; **29**(5):357–372.
46. Yang H, Bell MGH. Models and algorithms for road network design: a review and some new developments. *Transport Reviews* 1998a; **18**(3):257–278.
47. Yang H, Bell MGH. A capacity paradox in network design and how to avoid it. *Transportation Research Part A* 1998b; **32**(7):539–545.
48. Morlok EK, Riddle SM. Estimating the capacity of freight transportation systems: a model and its application in transport planning and logistics. *Transportation Research Record* 1999; **1653**:1–8.
49. Aggarwal KK. Integration of reliability and capacity in performance measure of a telecommunication network. *IEEE Transactions on Reliability* 1985; **34**(2):184–186.
50. Chan Y, Yim E, Marsh A. Exact and approximate improvement to the throughput of a stochastic network. *IEEE Transactions on Reliability* 1997; **46**(4):473–486.
51. Li D, Dolezal T, Haimes Y. Capacity reliability of water distribution networks. *Reliability Engineering and System Safety* 1993; **42**(1):29–38.
52. Billington R, Li W. *Reliability assessment of electric power systems using Monte Carlo methods*. Plenum Press: New York, 1994.
53. Nijkamp P, Vleugel AM, Kreutzberger E. Assessment of capacity in infrastructure networks: a multidimensional view. *Transportation Planning and Technology* 1993; **17**(4):301–310.
54. Yang H, Bell MGH, Meng Q. Modeling the capacity and level of service of urban transportation networks. *Transportation Research Part B* 2000; **34**(4):255–275.
55. Oppenheim N. Equilibrium trip distribution/assignment with variable destination costs. *Transportation Research Part B* 1993; **27**(3):207–217.
56. Sheffi Y. *Urban transportation networks: equilibrium analysis with mathematical programming methods*. Prentice Hall: England Cliffs, NJ, 1985.
57. Zhou Z, Chen A, Behkor S. C-logit stochastic user equilibrium model: formulations and solution algorithm. *Transportmetrica* 2012; **8**(1):17–41.
58. Chen BY, Lam WHK, Sumalee A, Shao H. An efficient solution algorithm for solving multi-class reliability-based traffic assignment problem. *Mathematical and Computer Modelling* 2011b; **54**(5–6):1428–1439.
59. Chen A, Zhou Z. The α-reliable mean-excess traffic equilibrium model with stochastic travel times. *Transportation Research Part B* 2010; **44**(4):493–513.
60. Zhou Z, Chen A. Comparative analysis of three user equilibrium models under stochastic demand. *Journal of Advanced Transportation* 2008; **42**(3):239–263.
61. Chang C-H, Tung Y-K, Yang J-C. Monte Carlo simulation for correlated variables with marginal distributions. *Journal of Hydraulic Engineering* 1994; **120**(3):313–331.
62. Patriksson M. *The traffic assignment problem: models and methods*. VNU Science Press: Utrecht, The Netherlands, 1994.
63. Gen M, Cheng R. *Genetic Algorithms and Engineering Optimization*. John Wiley and Sons, Inc.: New York, 2000.
64. Kasikitwiwat P. Capacity reliability and capacity flexibility of a transportation network. Ph.D. dissertation. Utah State University, Logan, Utah, 2005.
65. Friesz TL, Cho HJ, Mehta NJ, Tobin RL, Anandalingam G. A simulated annealing approach to the network design problem with variational inequality constraints. *Transportation Science* 1992; **26**(1):18–26.
66. Chen A, Kasikitwiwat P. Modeling network capacity flexibility of transportation networks. *Transportation Research Part A* 2011; **45**(2):105–117.
67. Chen A, Zhou Z, Lam WHK. Modeling stochastic perception error in the mean-excess traffic equilibrium model with stochastic travel times. *Transportation Research Part B* 2011b; **45**(10):1619–1640.

DOI: 10.1002/atr