Scenario-Based Approach to Analysis of Travel Time Reliability with Traffic Simulation Models

Jiwon Kim, Hani S. Mahmassani, Peter Vovsha,
Yannis Stogios, and Jing Dong

This study established a conceptual framework for capturing the probabilistic nature of travel times with the use of existing traffic simulation models. The framework features three components: scenario manager, traffic simulation models, and trajectory processor. The scenario manager captures exogenous sources of variation in travel times through external scenarios consistent with real-world roadway disruptions. The traffic simulation models then produce individual vehicle trajectories for input scenarios while further introducing randomness that stems from endogenous sources of variation. Finally, the trajectory processor constructs distributions of travel time either for each scenario or for multiple scenarios to allow users to investigate scenario-specific impact on variability in travel times and overall system reliability. Within this framework, the paper discusses methodologies for performing scenario-based reliability analysis that focuses on (*a*) approaches to obtaining distributions of travel times from scenario-specific outputs and (*b*) issues and practices associated with designing and generating input scenarios. The proposed scenario-based approach was applied to a real-world network to show detailed procedures, analysis results, and their implications.

With growing concern over unreliable travel times in urban networks and the associated costs of unexpected delays and frustration, travel time reliability has become an increasingly important issue in the arena of transportation network planning and traffic operations. Transportation policy makers and professionals are placing greater emphasis on improvement in reliability—consistency or dependability in travel times—along with improvement in average travel time. This emphasis calls for incorporating the reliability aspects into planning, operations, and economic evaluation models so that outputs of these models can adequately support transportation experts in developing a more reliable transportation system.

In the context of travel time reliability, significant progress has been made in measuring reliability, which entails developing and recommending various reliability indicators for practical use (*1–4*) and valuing reliability, which ranges from assessing the perceived value of reliability to incorporating reliability measures into travel demand modeling and network equilibrium frameworks (*5–8*). Another important yet less investigated aspect is modeling reliability, which involves identifying and capturing various sources of travel time unreliability in simulation or analytical models to reproduce realistic distributions of travel time or reliability measures. While efforts have been made analytically (*9*) and empirically (*3*) to predict variability in travel times in the presence of demand and capacity variations, little attention has been devoted to the use of existing traffic simulation models to produce reliability measures that predict and evaluate reliability levels of urban networks. By recognizing the important role of simulation-based dynamic traffic assignment models in the field of transportation planning and operations, this study attempts to establish a systematic and practical framework for producing reliability measures as output of simulation tools.

One way to capture the probabilistic nature of travel times by using simulation models is to conduct multiple simulation runs with different scenarios (e.g., different combinations of demand, capacity, and external events), possibly with different weights or occurrence probabilities, and to construct the resulting distribution of travel times to characterize overall system reliability. In this approach, primary emphasis is placed on designing and generating input scenarios to investigate the realistic travel time variability. This approach thus forms the basis for the scenario-based travel time reliability analysis that is the main focus of this paper. The paper is structured as follows. A conceptual framework for modeling and evaluating travel time reliability using simulation models is presented first. Within this framework is further discussion of scenario-based methodologies for constructing distributions of travel times, assessing reliability measures, and understanding impacts of scenarios on variability in travel times. Next, a real-world application is provided to show detailed procedures and analysis results. Finally, a summary and concluding remarks are provided.

METHODOLOGY

Reliability Modeling Framework That Uses Traffic Simulation Models

Before a methodological framework is built, one must understand the sources of uncertainty that affect the travel time reliability in the roadway environment. A previous study defined seven major root causes of travel time variability: (*a*) traffic incidents, (*b*) work zones, (*c*) weather, (*d*) special events, (*e*) traffic control devices,

J. Kim and H. S. Mahmassani, Transportation Center, Northwestern University, 215 Chambers Hall, 600 Foster Street, Evanston, IL 60208. P. Vovsha, Parsons Brinckerhoff, 1 Penn Plaza, Second Floor, New York, NY 10119. Y. Stogios, Delcan Corporation, 625 Cochrane Drive, Suite 500, Markham, Ontario L3R 9R9, Canada. J. Dong, Department of Civil, Construction, and Environmental Engineering, Iowa State University, 394 Town Engineering, Ames, IA 50111. Corresponding author: H. S. Mahmassani, masmah@northwestern.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2391, Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 56–68.
DOI: 10.3141/2391-06

(*f*) fluctuations in demand, and (*g*) inadequate base capacity (*10*). Many existing simulation tools view and model these factors as exogenous events by means of user-specified scenarios (*11*). Distinct from these exogenous factors are also endogenous sources of variation that are inherently reproduced, to varying degrees, by given traffic simulation models. Many studies have proposed ways to capture random variation in various traffic phenomena within particular micro- and mesosimulation models. Examples include flow breakdown (*12*), incidents attributable to drivers' risk-taking behaviors (*13*), and heterogeneity in driving behaviors (*14*).

On the basis of this identification, this study establishes a conceptual framework for modeling and estimating travel time reliability by using simulation models. As Figure 1 shows, the framework features three components: scenario manager, traffic simulation model, and trajectory processor. The primary role of the scenario manager is to prepare input scenarios for the traffic simulation models; these scenarios are a core part of this framework, as they directly affect the final distributions of travel times. Once the scenario manager generates a set of input scenarios, which represent any mutually consistent combinations of demand- and supply-side random factors, these scenarios are simulated in a selected traffic simulation model in conjunction with average demand obtained at a demand–supply equilibrium point under normal conditions encompassing any systematic variations. While exogenous sources of variation are captured through scenarios by the scenario manager, endogenous variation sources are captured in the traffic simulation model, which depends on the modeling capability of the selected tool.

In this framework, the traffic simulation models refer to particle-based models, namely micro- and mesoscopic simulation models (*15, 16*) that produce individual vehicle (or particle) trajectories. Regardless of the specific reliability measures of interest, to the extent that they can be derived from the distribution of travel times, the availability of particle trajectories in the output of a simulation model enables construction of any level of distributions of travel times of interest [e.g., networkwide, origin–destination (O-D), path, and link]. Then, the key building block for producing measures of reliability in this framework consists of particle trajectories and the associated experienced traversal times through the entirety or part of the travel path. Tasks such as converting simulated trajectories to various reliability measures are performed by the trajectory processor. This processor obtains the scenario-specific distribution of travel times from each simulation run and constructs the overall distribution of travel times aggregated over multiple scenarios.

While chaining of these three modules completes the necessary procedures for performing a scenario-based reliability analysis, worthy of mention are two feedback loops to further incorporate behavioral aspects of travelers into the reliability modeling framework. The inner loop in Figure 1 suggests that information from scenario-specific travel times might be used to make scenario-conditional demand adjustment (e.g., a change in departure time under severe weather conditions). The outer loop indicates that the overall uncertainty in the system might affect the average demand by shifting the equilibrium point (i.e., reliability-sensitive network equilibrium) on the basis of travel demand forecasting models that

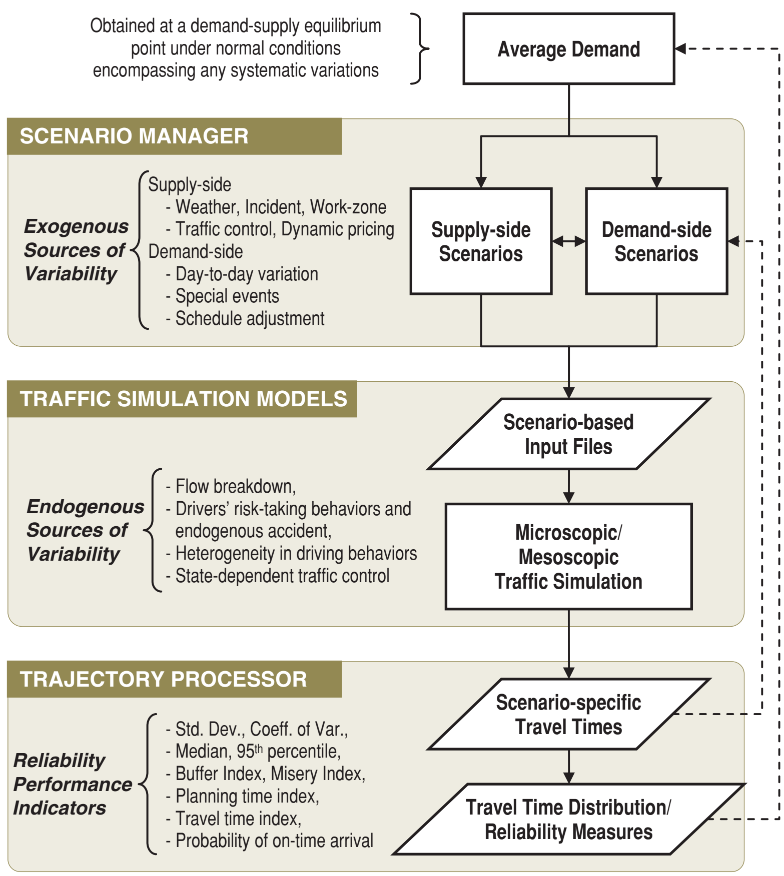

FIGURE 1 Core elements of reliability modeling framework (std. dev. = standard deviation; coeff. of var. = coefficient of variation).

predict the impact of reliability measures on travel patterns [e.g., Zhou et al. (*7*) and Jiang et al. (*8*)].

Scenario-Based Reliability Analysis

This section elaborates on the basic idea of the scenario-based reliability analysis within the previously described framework. Conceptually, traffic simulation models can be viewed as input–output functions in which inputs are scenarios that represent exogenous sources of roadway disruptions and outputs are distributions of travel times experienced by travelers under such disruptions. The objective of the scenario-based reliability analysis is to investigate variability in the output distribution of travel times by controlling the input scenario (i.e., input scenarios can be generated completely at random or in a more directed manner on the basis of a particular experimental design). Endogenous sources of random variations are not part of the control variables, as those are considered part of the logic of the traffic simulation model.

Let S denote an input scenario for the traffic simulation and X a scenario component that represents a supply- or demand-side random factor such as weather, incident, or day-to-day demand variation. Scenario S is defined by a set of selected scenario components [i.e., $S = \{X_1, X_2, \ldots, X_J\}$, where each component X_j, $j = 1, 2, \ldots, J$ is a vector of numerous attributes describing temporal (e.g., start time and duration), spatial (e.g., event location) and state (i.e., intensity or condition) properties of the instance of a given factor]. Suppose that N input scenarios S_i, $i = 1, 2, \ldots, N$ have been generated. Then the output distribution of travel times for each scenario is obtained by

$$T_i = g(S_i) \qquad i = 1, 2, \ldots, N \tag{1}$$

where T_i represents a collection of travel time t for a given O-D–path–link of interest under the ith scenario S_i, and $g(\cdot)$ represents a black-box representation of the given traffic simulation model. Let $f_i(t)$ denote the probability density function of scenario-specific travel times $t \in T_i$ under S_i. Then the overall probability density function of the distribution of travel times aggregated over N scenarios, $f(t)$, is calculated by the weighted sum of $f_i(t)$ as follows:

$$f(t) = \sum_{i=1}^{N} w_i f_i(t) \tag{2}$$

where w_i denotes the weight of the ith scenario with $\sum_{i=1}^{N} w_i = 1$, which is typically obtained from the scenario probability $w_i = P(S_i)$. Figure 2 presents a schematic diagram to illustrate the procedure of constructing the overall distribution of travel times on the basis of this concept.

Approaches to Assessment of Reliability

Travel time reliability is a relative concept in that it depends on the temporal and spatial boundaries for which travel times are observed. For example, the travel time reliability for weekdays is different from that for weekends on the same road network. Therefore, defining of time and space domains needs to precede assessment of reliability. In general, the time domain is specified by a date range of the overall period (e.g., June 1 to August 31, 2012), day of week (e.g., Monday to Friday), and time of day (6 to 10 a.m.), or it could be a specific season or day of each year (e.g., Thanksgiving Day). The space domain defines the level at which travel times are collected and the reliability measures are calculated (e.g., network level, O-D level, path level, or link level).

Two approaches are explored to assess the travel time reliability for given time and space domains: (*a*) the Monte Carlo approach and (*b*) the mix-and-match approach. The former tries to generate all possible scenarios that could occur within the given temporal and spatial boundaries to introduce realistic variations in the resulting distribution of travel times, while the latter constructs scenarios by

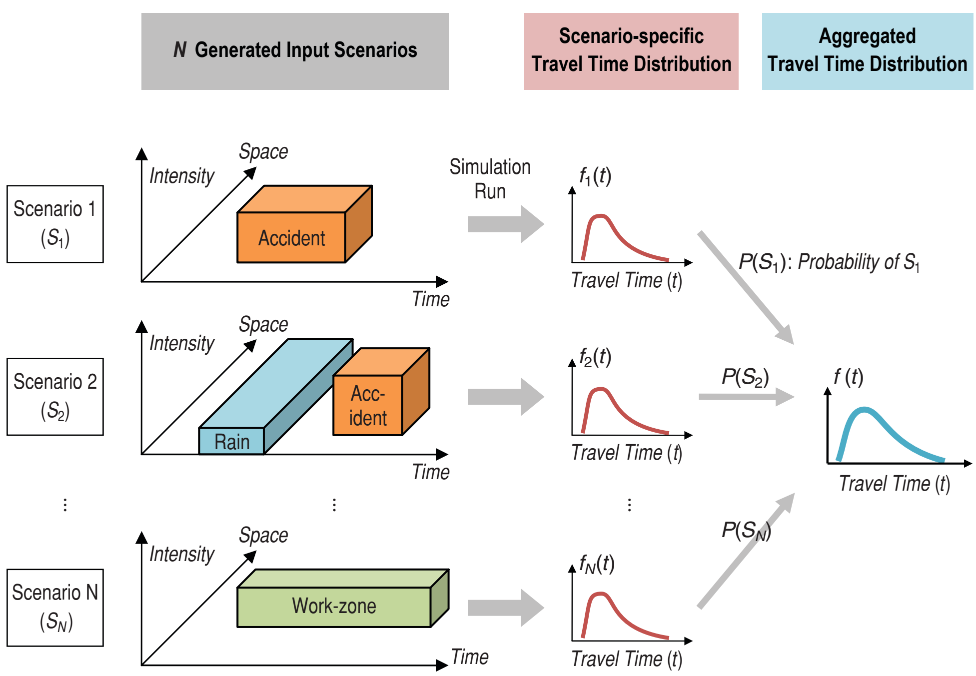

FIGURE 2 Construction of distribution of travel times from scenario-specific simulation outputs.

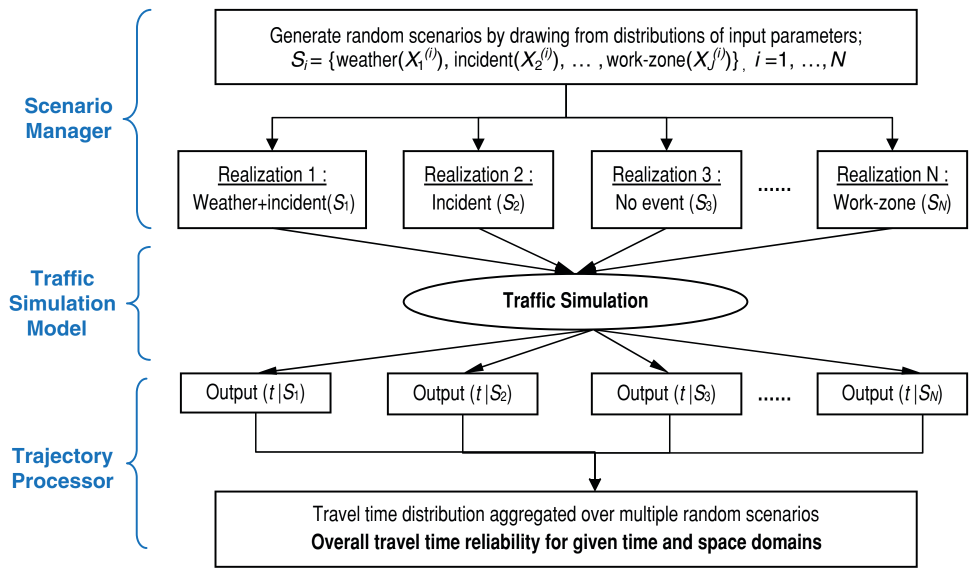

FIGURE 3 Monte Carlo approach.

manually choosing various combinations of scenario components. These approaches are discussed in more detail next.

Monte Carlo Approach

Monte Carlo simulation is used to prepare input scenarios aimed at propagating uncertainties in selected scenario components X into uncertainties in the generated scenarios S_i, $i = 1, 2, \dots, N$, which can, in turn, be translated into the resulting distribution of travel times. As Figure 3 shows, the scenario manager performs Monte Carlo simulation to generate hundreds or thousands of input scenarios by sampling from the joint probability distribution of scenario components. Because each scenario is equally likely, the trajectory processor is simply allowed to aggregate distributions of travel times from a large number of simulation runs to obtain the most likely (probable) outcome of a set of reliability performance indicators for the given time and space domains.

Mix-and-Match Approach

Instead of generating scenarios randomly given the underlying stochastic processes, one could explicitly specify scenarios with particular historical significance or policy interest. The mix-and-match approach aims to construct input scenarios in a more directed manner by enumerating all possible combinations of specific input factors or by directly using known historical events or specific instances (e.g., holiday, ball game, etc.). The schematic diagram in Figure 4 illustrates this approach with a simple example. Two

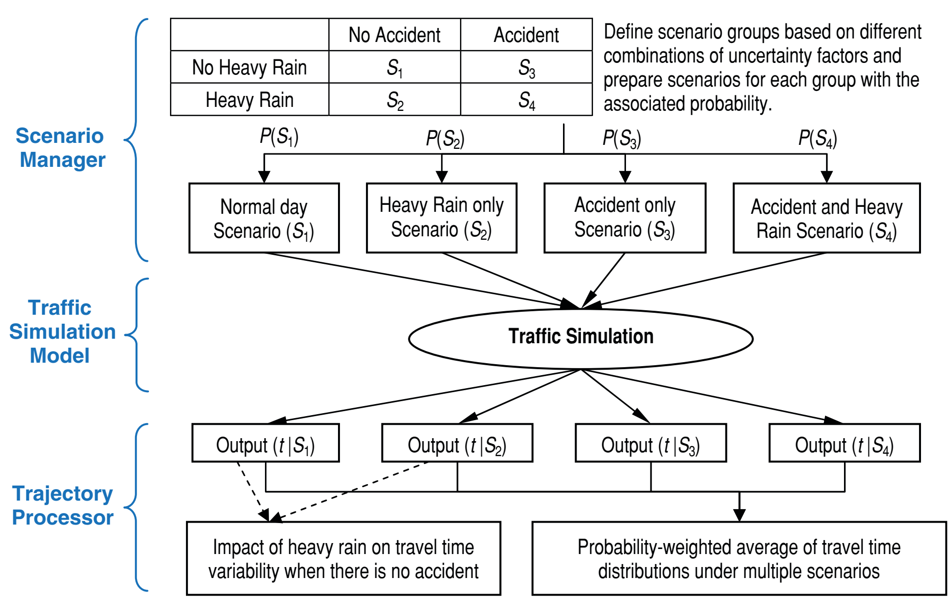

FIGURE 4 Mix-and-match approach.

scenario components, accident and heavy rain, are considered, and each component has two discrete states: occur and not occur. From the Cartesian product of the states of two components, four possible scenario groups are defined, as Figure 4 shows. One may assume that a representative scenario exists for each group, with the scenario probability assigned on the basis of the joint probability of accident and heavy-rain events. Then a probability-weighted average of distributions of travel times under all four scenarios can be used as the expected distribution of travel times to approximate the overall reliability measures. A more informative use of this approach is to understand the impact of a particular scenario component on travel time variability by investigating gaps between different combinations of output results.

Combined Approach

Unlike the simple example above, however, it is often necessary to allow randomness in scenarios within each group, especially when no representative scenario has been predefined. It is also possible to have no probability value for each scenario group known to users. In both of those cases, the Monte Carlo approach can be used in conjunction with the mix-and-match approach [i.e., sampling of random scenarios from their conditional distributions given each group (for the former)] and generating a large number of scenarios for the entire scenario space and categorizing them into the associated groups to obtain the group probabilities (for the latter).

Generating Scenarios by Considering Dependencies

One of the practical issues in generating scenarios is considering the dependencies in various random factors. As represented by the dotted arrows in Figure 5, certain scenario components are dependent on other components. Incident occurrence, in which event properties (e.g., frequency, duration, and severity) tend to be affected by weather and other external events, is the most prominent example. The authors investigated weather-conditional incident rates (incidents per hour per lane mile) by measuring the number of incidents during the total time of exposure to different weather conditions by using historical incident data from 2007 to 2010 in Chicago. As Figure 6 shows, incident rates tended to increase as the severity of rain or snow events increased. In addition to incidents, dependencies were also observed on the traffic management side: weather-responsive traffic management strategies were deployed on the basis of the types and severities of weather events (*17*), and traffic incident management was triggered by incident events. In the scenario manager, such dependencies are taken into account during the generation process. Once the scenario components of interest are defined, the scenario manager identifies dependency relationships between components and derives a generation order such that components that affect others are generated before their dependent ones. Following the generation order, the scenario manager generates each component sequentially (e.g., weather → incident → incident management) so that each component is sampled from its distribution, conditioned on all previously sampled components.

APPLICATION

In this section, the presented framework is applied to a real-world network to show detailed procedures and analysis results by using a mesoscopic traffic simulation tool, DYNASMART-P (*16, 18*).

Time and Space Domains and Data Collection

Suppose that one was interested in evaluating the reliability of travel times in a Long Island, New York, network during a weekday (Monday to Friday) morning peak (6 to 10 a.m.) during the winter (November 2010 to February 2011). For this, an O-D pair between Washington

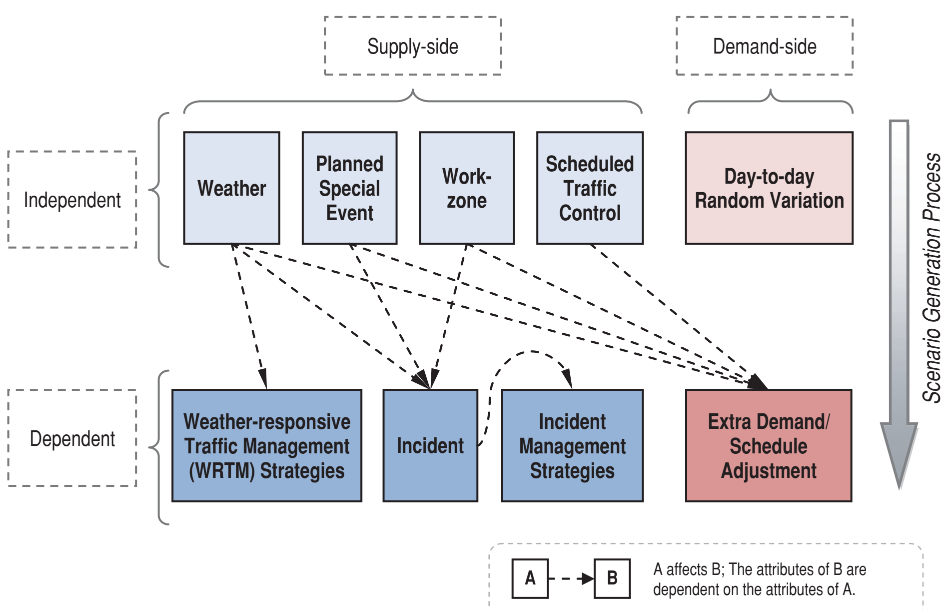

FIGURE 5 Various scenario components and dependency relationships.

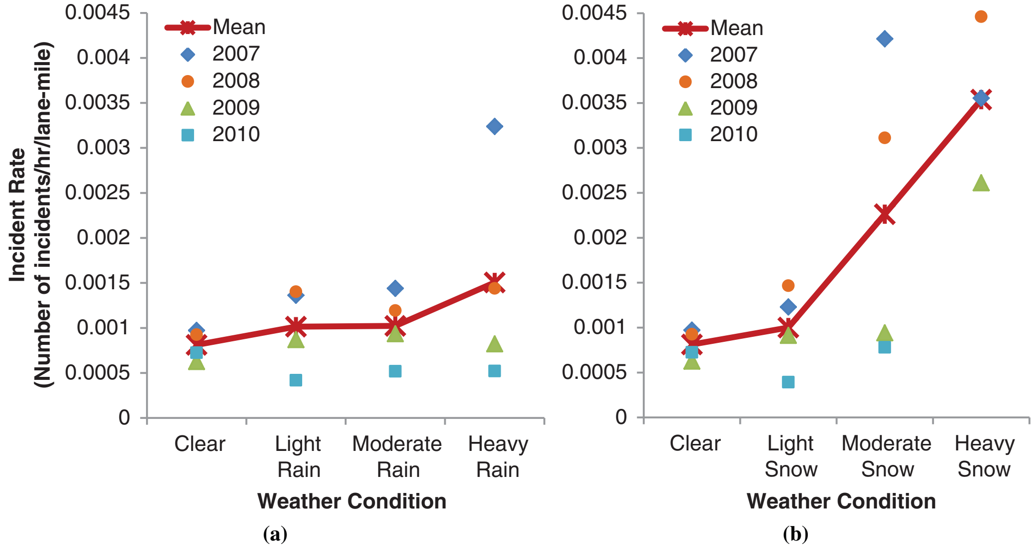

FIGURE 6 Weather-conditional incident rates (Chicago incident data from 2007 to 2010): (*a*) rain and (*b*) snow.

Avenue and Cross Island Parkway was selected, as shown in Figure 7, where the trip distance along the Long Island Expressway (I-495) is 27.5 mi. Two uncertainty factors were considered as scenario components: weather and incident. To estimate parameters for specifying weather and incident characteristics, historical data for the given time and space domains were collected and analyzed. Weather data were obtained from the nearest automated surface observing system (ASOS) station at Republic Airport, Farmingdale, New York, where the percentage of hours of each weather condition was as follows:

- Clear: 92.05%,
- Rain: 4.91% (light: 84.86%; moderate: 12.97%; heavy: 2.18%), and
- Snow: 3.05% (light: 84.85%; moderate: 8.76%; heavy: 6.39%).

Incident data were collected from the INFORM system and provided by the New York State Department of Transportation (*19*). The incident data contained information on event locations (red triangles in Figure 7) and severities in relation to the number of closed lanes, which were distributed as follows:

- No lane closed: 35.34%,
- One lane closed: 50.32%,
- Two lanes closed: 11.17%, and
- Three or more lanes closed: 3.17%.

The overall incident rate (i.e., the number of incidents per observation hour per total lane mile) is measured as 0.002 incident per hour per lane mile.

FIGURE 7 Study network and selected O-D pair (Long Island).

TABLE 1 Input Parameters and Sampling Methods

Scenario Component	Properties Required for Sampling Event Instances			
	Frequency	Duration	Intensity	Location
Weather				
Input parameter	Always[a]	Time period for each weather condition	Discrete states: {CL, LR, MR, HR, LS, MS, HS}	Networkwide
Sampling method	Nonparametric[b]	Nonparametric[b]	Nonparametric[b]	Apply to entire network
Incident				
Input parameter	Mean incident rate λ(incidents/h/lane mile)	Two parameters in fitted model	Percentage of capacity loss (number of lanes closed)	Section specific
Sampling method	Parameteric–Poisson distribution[c] $\lambda_{CL} = 0.0019$ $\lambda_{LR} = 0.0024$ $\lambda_{MR} = 0.0047$ $\lambda_{HR} = 0.0071$ $\lambda_{LS} = 0.0043$ $\lambda_{MS} = 0.0095$ $\lambda_{HS} = 0.0189$	Parametric–gamma distribution Shape = 1.210 Scale = 31.553	Nonparametric–empirical PMF[d] $P(0.15) = .35$ $P(0.3) = .5$ $P(0.6) = .11$ $P(0.9) = .04$	Parametric–(homogeneous): Poisson point process in space

NOTE: CL = clear; LR = light rain; MR = moderate rain; HR = heavy rain; LS = light snow; MS = moderate snow; HS = heavy snow; PMF = probability mass function.
[a]In this experiment, weather events are viewed as always present with one of the seven states: CL, LR, MR, HR, LS, MS, and HS.
[b]Use the actual measured values; randomly draw from historical time series of weather data.
[c]λ_x = mean incident rate under weather condition x.
[d]$P(x)$ = probability that the fraction of link capacity lost because of the instance becomes x (remaining capacity becomes $1 - x$).

Input Parameters and Sampling Methods

Each scenario component is characterized by four major event properties: frequency, duration, intensity, and location, where each property is specified either parametrically or nonparametrically. Table 1 presents input parameters and sampling methods for each property of weather and incident components.

Modeling of weather events in a fully parametric manner is not a trivial task and requires a complex stochastic model that captures a temporal clustering of the event points. As development of such a model is beyond the scope of this paper, a nonparametric sampling approach was used; in it, the empirical data were directly used for generating weather scenarios. The scenario manager was populated with 5-min ASOS weather observations for the analysis period, and it randomly sampled the entire time series of 4-h weather scenarios from the data. This approach is especially useful when dependence exists in the data structure, as it preserves existing dependencies. Weather data exhibit dependence not only in sequence (a time series) but across different parameters (e.g., precipitation intensity, visibility, duration, etc.). On the basis of the categorization used in ASOS data, seven mutually exclusive and exhaustive states were defined: clear, light rain, moderate rain, heavy rain, light snow, moderate snow, and heavy snow; any time point during the scenario horizon was assigned one of these states, as illustrated in Figure 8*a*.

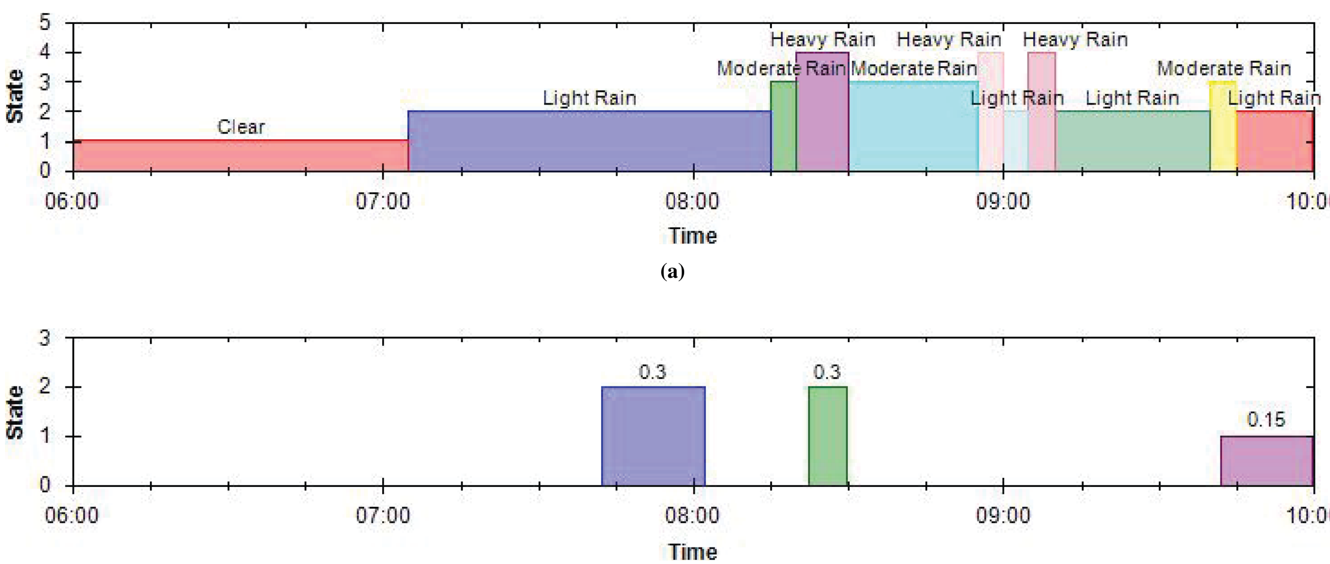

FIGURE 8 Temporal profiles represented by rectangular pulse with duration (width) and intensity (height) for one instance of a scenario consisting of (*a*) weather and (*b*) incident events.

In contrast, many random properties of incident events can be modeled by using known parametric probability distributions. For frequency, incidents were assumed to follow a Poisson process with the mean incident rate. As noted earlier, however, the rate is highly dependent on the prevailing weather conditions, and therefore the weather-conditional mean incident rates for seven weather conditions were estimated on the basis of the historical data, as presented in Table 1. To reproduce incident instances following this state-contingent incident rate, a discrete-event approach to simulation was applied to identify discrete time points in which the weather state changed on the basis of a given (sampled) weather time series, and the incident occurrence pattern at each variable time interval was determined by applying the associated mean incident rates. To validate this approach, 1,000 scenarios with and without consideration of dependencies between weather and incident were tested and simulated incident rates were compared with actual observed ones, as shown in Figure 9. The results show that the scenarios from the weather-dependent incident sampling reproduced the real-world incident frequency successfully, while the scenarios generated in the weather-independent manner significantly underestimated the likelihood of incident occurrence under severe weather conditions. For incident duration, the gamma distribution was selected on the basis of the model-fitting results, and two input parameters were estimated, as follows: shape = 1.210 and scale = 31.553. Incident intensity was expressed as the percentage capacity loss (the fraction of link capacity lost because of the instance), and the empirical probability mass function was constructed on the basis of the observed pattern for the number of lanes closed, as presented in Table 1.

Scenario Specification and Generation

This application used the combined approach, in which a discrete set of scenario groups were defined, as in the mix-and-match approach, but random scenarios for each group and the group probability were obtained from Monte Carlo sampling. Six scenario groups were defined on the basis of the Cartesian product of three weather states (clear, rain, and snow) and two incident states (incident and no incident). To calculate scenario group probabilities, 10,000 scenarios were generated and classified into one of those six scenario groups. Each scenario represented a single day (6 to 10 a.m.) with the combination of weather and incident events (e.g., Figure 8). The probability of each group occurring is presented in Table 2. Scenarios with clear weather and incidents accounted for 61% of the total trials as the most likely scenario, and scenarios with snow and no incident accounted for 0.4% as the least likely scenario.

In sampling random scenarios for each group, the number of scenarios required to estimate the mean travel time with no worse than a 10% error and at least 90% confidence were initially identified (*20*). The calculation result showed that the sample size of 20 scenarios would meet the criteria. However, when the current authors considered that variability measures such as the standard deviation or other reliability metrics tend to require a larger sample size, 40 was selected as the final sample size for this experiment. Therefore, 40 scenarios were randomly selected for each group and simulated by using DYNASMART-P to obtain scenario-specific (or scenario group–specific) distribution of travel times. For the clear–no incident group, however, only one scenario was simulated, as it did not involve any randomness.

Analysis Results

After completion of traffic simulation for the selected scenarios, distributions of travel times were obtained, as presented in Figure 10, for which the *y*-axis represents the probability mass function and the *x*-axis represents the O-D travel time in minutes. Figure 10*a* shows the combined (probability-weighted) distribution of travel times obtained by using the method in Equation 2, and Figure 10, *b* to *g,* shows the scenario-specific distributions of travel times. From the scenario-specific probability mass functions in Figure 10, one

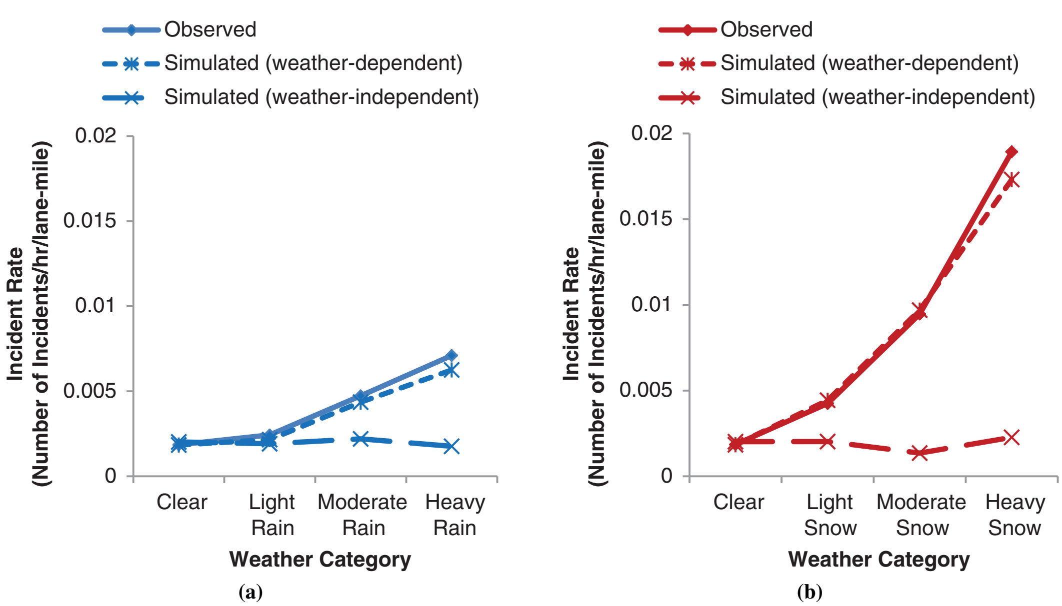

FIGURE 9 Weather-conditional incident rates: observed versus simulated for (*a*) rain and (*b*) snow (Long Island incident data).

TABLE 2 Traffic Simulation Results and Estimated Reliability Measures

Scenario Group	CL_nINC	CL_INC	RA_nINC	RA_INC	SN_nINC	SN_INC	Total
Probability of Occurrence	.242	.610	.023	.072	.004	.049	1
Descriptive Statistics							
Number of scenarios	1	40	40	40	40	40	201
Number of observations	1,431	57,640	57,690	56,310	57,640	56,676	285,956
Mean travel time (min)	27.65	27.52	28.69	28.74	29.24	31.64	27.87
Median travel time (min)	25	26	27	27	28	29	26
Reliability Measures							
Standard deviation (min)	7.41	6.48	6.26	6.13	6.15	8.29	6.86
Coefficient of variation	0.27	0.24	0.22	0.21	0.21	0.26	0.25
80th percentile (min)	27	27	28	29	29	33	28
95th percentile (min)	41	41	41	41	42	46	42
Buffer index (%)[a]	48.27	48.98	42.9	42.67	43.64	45.37	50.7
Buffer time (min)[b]	13.35	13.48	12.31	12.26	12.76	14.36	14.13
Percentage on time (%)[c]	91.4	91.22	90.72	90.39	91.3	85.65	90.61
Planning time index[d]	1.64	1.64	1.64	1.64	1.68	1.84	1.68
Misery index[e]	2.27	2.09	2.09	2.05	2.09	2.35	2.15

NOTE: CL_nINC = clear, no incident; CL_INC = clear, incidents; RA_nINC = rain, no incident; RA_INC = rain, incidents; SN_nINC = snow, no incident; SN_INC = snow, incidents.
[a]The difference between the 95th percentile travel time and the average travel time, normalized by the average travel time (*3*).
[b]The difference between the 95th percentile travel time and the average travel time (*3*).
[c]The percentage of trips with travel times < (1.25 ∗ median travel time) (*3*).
[d]The 95th percentile travel time divided by free-flow travel time (*3*).
[e]The average of the highest 5% of travel times divided by the free-flow travel time (*3*).

can see that travel times become more dispersed as the weather state changes from clear to snow and the incident state changes from no incident to incident. Significantly high dispersions are observed in distributions of travel times under snow conditions, but impact of those dispersions on the combined distribution appears to be small because of the low probabilities.

Various statistics and reliability performance measures were extracted from each distribution of travel times and are presented in Figure 11 and Table 2. For the individual scenario group, the mean and median travel times tend to grow from left to right, while the standard deviation is higher on the sides than in the middle. Having such a high standard deviation in the snow–incidents case appears reasonable, as the distribution of travel times is highly dispersed, as shown in Figure 10*g*. But the relatively high standard deviation for the clear–no incident case seems to require a different explanation, with one of the reasons possibly being that the standard deviation is quite sensitive to the tails of a distribution and slight changes in the tails could lead to substantially different standard deviations (*21*). Although the distributions of travel times for clear–no incident and clear–incidents had little visible difference and the maximum travel time for clear–no incidents was smaller than that for clear–incidents, the relative impact of the tail of clear–no incident on the standard deviation appeared to be greater than that of clear–incidents. This result could be partly because of a much smaller sample size for clear–no incident. This tendency was also apparent for the misery index measure (Figure 11*j*), for which clear–no incident showed a higher value than other groups did (except snow–incidents), an indication that the average of the highest five percentages of travel times were higher in this group than in others.

For the 95th-percentile travel time (Figure 11*f*), all the scenario groups had similar values and only snow–incidents showed a noticeable difference. This finding was also true for the planning time index (Figure 11*i*), which is the 95th-percentile travel time divided by the free-flow travel time. This result suggests that the 95th percentile may be too extreme to reflect different characteristics under different scenarios. A previous study also noted this issue and recommended the use of the 80th percentile instead (*3*). As Figure 11*e* shows, the 80th-percentile travel time appears to capture the effects of different weather and incident conditions better than the 95th-percentile does.

Another important observation concerns the buffer index (Figure 11*g*), which measures the relative distance between the central (mean) and extreme (95th-percentile) values and represents the extra buffer time (i.e., the percentage of the mean travel time that travelers should add to the mean to ensure on-time arrival 95% of the time). From the scenario-specific distributions of travel times, buffer index values for clear–no incident and clear–incidents were estimated to be higher than those of snow–incidents. However, the actual buffer time, calculated as the difference between the mean and 95th-percentile travel times, was higher under snow–incidents. In general, caution is required when reliability measures across groups are being compared, as some measures are normalized by scenario-dependent reference values (e.g., mean and median), and such relative distances should be interpreted differently from measures of absolute distance to a global reference point (e.g., free-flow travel time).

CONCLUSION

While simulation-based traffic prediction models have been widely used for operational and planning purposes for decades, there has been no systematic development of approaches to modeling travel

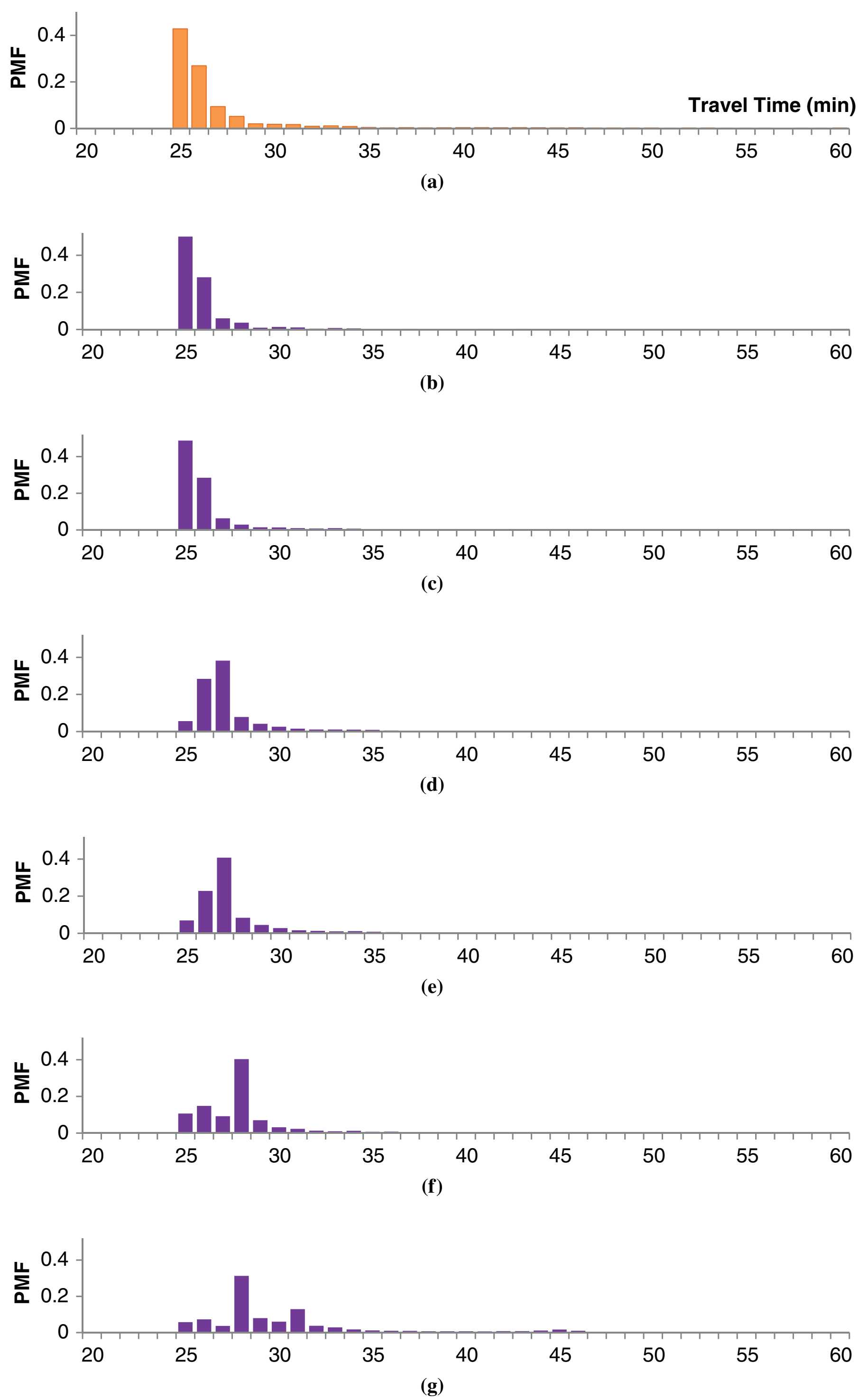

FIGURE 10 Overall and scenario-specific distributions of travel times (right truncated): (*a*) overall, (*b*) clear–no incident (P = .242), (*c*) clear–incidents (P = .610), (*d*) rain–no incident (P = .023), (*e*) rain–incidents (P = .072), (*f*) snow–no incident (P = .004), and (*g*) snow–incidents (P = .049).

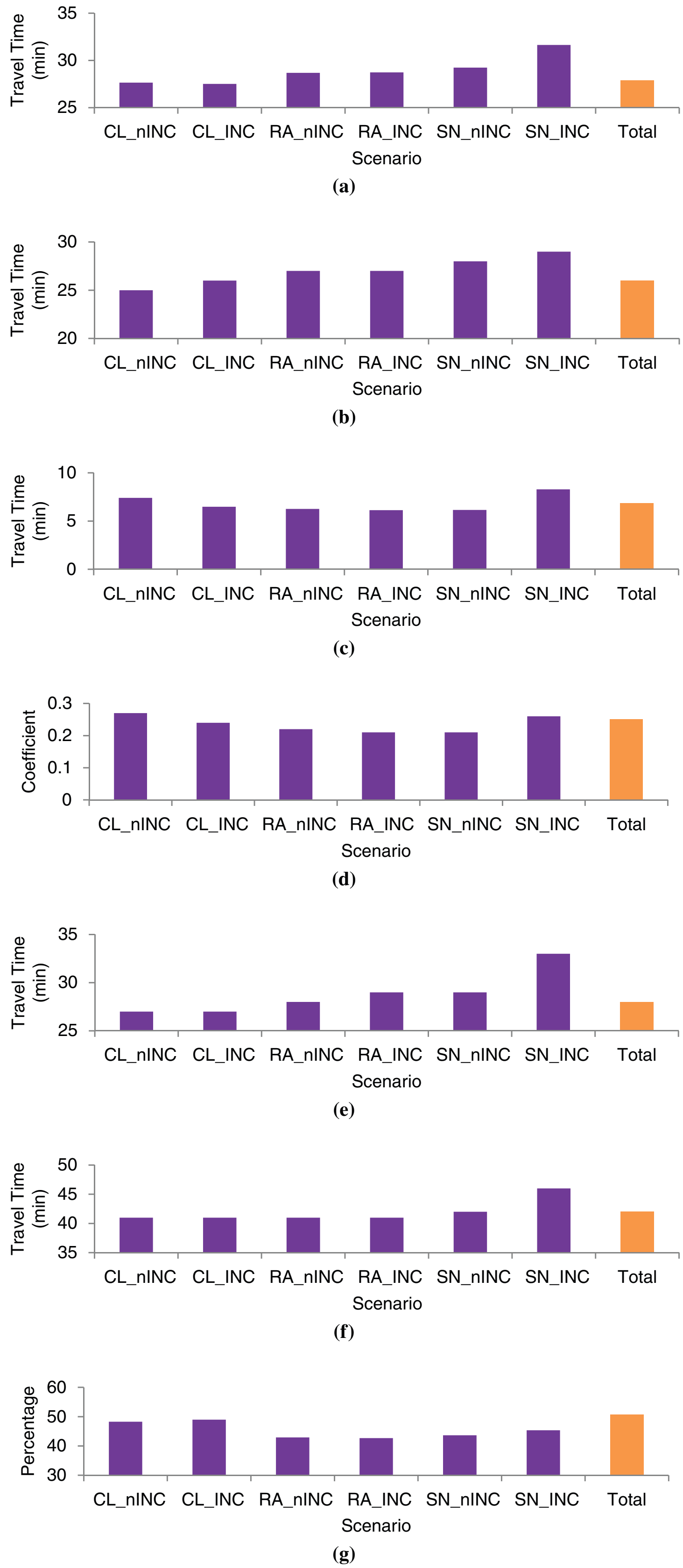

FIGURE 11 Travel time reliability measures by scenario group: (*a*) mean, (*b*) median, (*c*) standard deviation, (*d*) coefficient of variation, (*e*) 80th percentile, (*f*) 95th percentile, and (*g*) buffer index.

(*continued*)

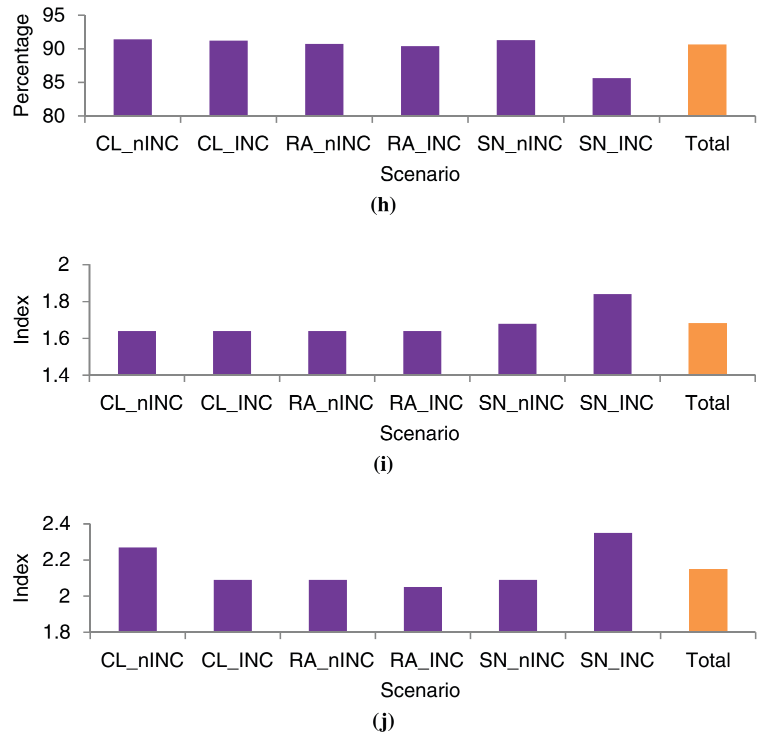

FIGURE 11 *(continued)* **Travel time reliability measures by scenario group: (*h*) percentage on time, (*i*) planning time index, and (*j*) misery index.**

time reliability within the framework of traffic simulation models. This paper established a conceptual framework for capturing the probabilistic nature of travel times by using existing traffic simulation models. The framework features three components: scenario manager, traffic simulation models, and trajectory processor. The scenario manager captures exogenous sources of travel time variation through external scenarios consistent with real-world roadway disruptions. The traffic simulation models then produce individual vehicle trajectories for input scenarios while further introducing randomness that stems from endogenous sources of variations. Finally, the trajectory processor constructs distributions of travel times for either each scenario or multiple scenarios on the basis of simulated trajectories to allow users to investigate scenario-specific impact on travel time variability as well as on overall system performance. Within this framework, this paper discussed methodologies for performing scenario-based reliability analysis that focuses on (*a*) approaches to obtaining overall distributions of travel times from scenario-specific outputs and (*b*) issues and practices associated with designing and generating input scenarios. The proposed scenario-based approach was applied to a real-world network to show detailed procedures, analysis results, and their implications.

This paper demonstrated the use of traffic simulation models in generating distributions of travel times that reflect various demand- and supply-side uncertainty factors. Although endogenous variations from scenario components were excluded at this point, the scenario-based approach is not limited to modeling only external events. Rather it expands the view of what can be specified as scenarios. Any phenomena that are characterized by certain event properties (e.g., frequency, duration, and intensity) can be generated and provided as inputs to traffic simulation models. For instance, flow breakdown can also be specified as an external event by identifying triggering mechanisms and dependencies with other external factors such as weather, as discovered by Kim et al. (*22*). Therefore, many extensions and developments of this framework are possible.

ACKNOWLEDGMENTS

This paper is based on research conducted under SHRP 2 Project L04, Incorporating Reliability Performance Measures in Operations and Planning Modeling Tools, and a National Science Foundation grant, Toward More Reliable Mobility: Improved Decision Support Tools for Transportation Systems. The authors acknowledge the valuable comments of the SHRP 2 project Technical Expert Task Group, as well as those of various collaborators at Northwestern University (especially Alireza Talebpour, Lan Jiang, Ali Zockaie, Tian Hou, and Meead Saberi), at PB Americas (Bob Donnelly), and at Delcan (Ken Curry and Gelu Marcu). The authors are especially grateful to William Hyman and Stephen Andrle of SHRP 2 for their continued support and encouragement throughout this effort. The paper has benefited considerably from the comments of anonymous referees.

REFERENCES

1. Lomax, T., D. Schrank, S. Turner, and R. Margiotta. *Selecting Travel Time Reliability Measures.* Texas Transportation Institute, Texas A&M University System, College Station, 2003.
2. Cambridge Systematics, Inc., Dowling Associates, Inc., System Metrics Group, Inc., and Texas Transportation Institute. *NCHRP Report 618: Cost-Effective Performance Measures for Travel Time Delay, Variation,*

and Reliability. Transportation Research Board of the National Academies, Washington, D.C., 2008.

3. Margiotta, R., T. Lomax, M. Hallenbeck, R. Dowling, A. Skabardonis, and S. Turner. *SHRP 2 Report S2-L03-RR: Analytical Procedures for Determining the Impacts of Reliability Mitigation Strategies.* Transportation Research Board of the National Academies, Washington, D.C., 2010.
4. Mahmassani, H. S., T. Hou, and J. Dong. Characterizing Travel Time Variability in Vehicular Traffic Networks: Deriving a Robust Relation for Reliability Analysis. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2315,* Transportation Research Board of the National Academies, Washington, D.C., 2012, pp. 141–152.
5. Fosgerau, M., and A. Karlström. The Value of Reliability. *Transportation Research Part B: Methodological,* Vol. 44, No. 1, 2010, pp. 38–49.
6. Carrion, C., and D. Levinson. Value of Travel Time Reliability: A Review of Current Evidence. *Transportation Research Part A: Policy and Practice,* Vol. 46, No. 4, 2012, pp. 720–741.
7. Zhou, X., H. S. Mahmassani, and K. Zhang. Dynamic Micro-Assignment Modeling Approach for Integrated Multimodal Urban Corridor Management. *Transportation Research Part C: Emerging Technologies,* Vol. 16, No. 2, 2008, pp. 167–186.
8. Jiang, L., H. S. Mahmassani, and K. Zhang. Congestion Pricing, Heterogeneous Users, and Travel Time Reliability: Multicriterion Dynamic User Equilibrium Model and Efficient Implementation for Large-Scale Networks. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2254,* Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 58–67.
9. Clark, S., and D. Watling. Modelling Network Travel Time Reliability Under Stochastic Demand. *Transportation Research Part B: Methodological,* Vol. 39, No. 2, 2005, pp. 119–140.
10. Cambridge Systematics, Inc.; Texas Transportation Institute; University of Washington; and Dowling Associates. F-SHRP *Web Document 3: Providing a Highway System with Reliable Travel Times: Study 3—Reliability.* Transportation Research Board of the National Academies, Washington, D.C., 2003. http://onlinepubs.trb.org/onlinepubs/f-shrp/f-shrp_webdoc_3.pdf.
11. Mahmassani, H. S., J. Dong, J. Kim, R. B. Chen, and B. Park. *Incorporating Weather Impacts in Traffic Estimation and Prediction Systems.* FHWA-JPO-09-065. FHWA, U.S. Department of Transportation, 2009.
12. Dong, J., and H. S. Mahmassani. Flow Breakdown and Travel Time Reliability. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2124,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 203–212.
13. Hamdar, S. H., and H. S. Mahmassani. From Existing Accident-Free Car-Following Models to Colliding Vehicles: Exploration and Assessment. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2088,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 45–56.
14. Kim, J., and H. S. Mahmassani. Correlated Parameters in Driving Behavior Models: Car-Following Example and Implications for Traffic Microsimulation. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2249,* Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 62–77.
15. Chang, G.-L., H. S. Mahmassani, and R. Herman. Macroparticle Traffic Simulation Model to Investigate Peak-Period Commuter Decision Dynamics. In *Transportation Research Record 1005,* TRB, National Research Council, Washington, D.C., 1985, pp. 107–121.
16. Mahmassani, H. S. Dynamic Network Traffic Assignment and Simulation Methodology for Advanced System Management Applications. *Networks and Spatial Economics,* Vol. 1, No. 3–4, 2001, pp. 267–292.
17. Mahmassani, H. S., J. Kim, T. Hou, A. Zockaie, M. Saberi, L. Jiang, O. Verbas, S. Cheng, Y. Chen, and R. Haas. *Implementation and Evaluation of Weather Responsive Traffic Estimation and Prediction System.* Final report. FHWA-JPO-12-055. FHWA, U.S. Department of Transportation, 2012.
18. Mahmassani, H. S., and H. Sbayti. *DYNASMART-P User's Guide,* Version 1.6, Northwestern University, Evanston, Ill., 2009.
19. INFORM homepage. New York State Department of Transportation, Albany, http://www.informny.com/. Accessed Jan 9, 2012.
20. Law, A., and W. D. Kelton. Output Data Analysis for a Single System. *Simulation Modeling and Analysis,* 3rd ed., McGraw-Hill, New York, 1999, pp. 496–552.
21. Wilcox, R. R., and H. J. Keselman. Modern Robust Data Analysis Methods: Measures of Central Tendency. *Psychological Methods,* Vol. 8, No. 3, 2003, pp. 254–274.
22. Kim, J., H. S. Mahmassani, and J. Dong. Likelihood and Duration of Flow Breakdown: Modeling the Effect of Weather. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2188,* Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 19–28.

The authors are responsible for the content of this paper.

The Traffic Flow Theory and Characteristics Committee peer-reviewed this paper.