1. Report No. SWUTC/10/167275-1	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Quantifying Travel Time Variability in Transportation Networks		5. Report Date March 2010
		6. Performing Organization Code
7. Author(s) Stephen D. Boyles, Avinash Voruganti and S. Travis Waller		8. Performing Organization Report No. 167275-1
9. Performing Organization Name and Address Center for Transportation Research University of Texas at Austin 3208 Red River, Suite 200 Austin, Texas 78705-2650		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. 10727
12. Sponsoring Agency Name and Address Southwest Region University Transportation Center Texas Transportation Institute Texas A&M University System College Station, Texas 77843-3135		13. Type of Report and Period Covered
		14. Sponsoring Agency Code
15. Supplementary Notes Supported by general revenues from the State of Texas		

16. Abstract

Nonrecurring congestion creates significant delay on freeways in urban areas, lending importance to the study of facility reliability. In locations where traffic detectors record and archive data, approximate probability distributions for travel speed or other quantities of interest can be determined from historical data; however, the coverage of detectors is not always complete, and many regions have not deployed such infrastructure. This report describes procedures for estimating such distributions in the absence of this data, considering both supply-side factors (reductions in capacity due to events such as incidents or poor weather) and demand-side factors (such as daily variation in travel activity). Two demonstrations are provided: using data from the Dallas metropolitan area, probability distributions fitting observed speed data are identified, and regression models developed for estimating their parameters. Using data from the Seattle metropolitan area, the appropriate capacity reduction applied to planning delay functions in the case of an incident is identified.

17. Key Words Transportation Planning, Regression Models, Supply-Side Analysis, Demand-Side Analysis, Travel Demand, Travel Speed Distribution		18. Distribution Statement No restrictions. This document is available to the public through NTIS: National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161	
19. Security Classif.(of this report) Unclassified	20. Security Classif.(of this page) Unclassified	21. No. of Pages 39	22. Price

Form DOT F 1700.7 (8-72)   **Reproduction of completed page authorized**

Quantifying Travel Time Variability in Transportation Networks

by

Stephen D. Boyles
Avinash Voruganti
S. Travis Waller

Research Report SWUTC/10/167275-1

March 2010

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

ABSTRACT

Nonrecurring congestion creates significant delay on freeways in urban areas, lending importance to the study of facility reliability. In locations where traffic detectors record and archive data, approximate probability distributions for travel speed or other quantities of interest can be determined from historical data; however, the coverage of detectors is not always complete, and many regions have not deployed such infrastructure. This report describes procedures for estimating such distributions in the absence of this data, considering both supply-side factors (reductions in capacity due to events such as incidents or poor weather) and demand-side factors (such as daily variation in travel activity). Two demonstrations are provided: using data from the Dallas metropolitan area, probability distributions fitting observed speed data are identified, and regression models developed for estimating their parameters. Using data from the Seattle metropolitan area, the appropriate capacity reduction applied to planning delay functions in the case of an incident is identified.

ACKNOWLEDGEMENTS

EXECUTIVE SUMMARY

Uncertainty characterizes transportation systems. Whether in long-range transportation planning, operational analyses, impact studies, or in investment banks considering the risk of investing in a toll road, the catalog of unknown factors is lengthy. In the short-term, daily demand fluctuations, incidents, and poor weather contribute to unreliability of travel times, and variability in travel speeds, link flows, and other "operational metrics." Researchers have developed algorithms which can account for these in planning and operations, but two major challenges prevent immediate adoption of these methods:

1. These models typically require explicit probability distributions on travel speeds (or other metrics) for all roadway segments.
2. Intelligent transportation systems (ITS) often provide data on particular freeways, allowing empirical distributions to be derived, but coverage is often sparse, detector malfunctions all too common, and such data may not include all of the metrics of interest.

This report details a method for estimating the distribution of any operational metric (such as speed, volume, or planning capacity) using ITS data. Broadly speaking, the method involves (1) identifying a corpus of ITS data, and other relevant data sets (on weather, incidents, and so forth); (2) identifying the best-fitting distribution to describe the distribution of the operational metric in question; (3) performing regression analysis to link the distribution parameters to roadway geometry and other factors; and (4) applying these regression equations to locations *without* data to estimate distributions at these locations.

The method is divided into a supply-side analysis, and a demand-side analysis. In the supply-side analysis, a more accurate prediction is provided by partitioning the data set according to the "state" of the freeway. States include the presence of an incident, poor weather, and so forth, depending on the available data sets. A best-fitting distribution is identified for each state, and the law of total probability applied to create an unconditional distribution. A further refinement is possible through a demand-side analysis, which incorporates fluctuations in daily travel demand; a demonstration using Dallas-Ft. Worth data shows a sixfold improvement in the accuracy of an estimate of travel speed standard deviation.

A total of two demonstrations are provided. The aforementioned experiment in the Dallas-Ft. Worth region involves predictions of travel speed distributions, while a second experiment using data from Seattle involves creation of a sketch planning network accounting for incident conditions (perhaps to evaluate different information-provision strategies to travelers). These demonstrations show the flexibility of the method to handle different problems in operations and planning, and to represent a diversity of operational metrics.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

The extent and effects of congestion in urban areas have been known and documented for decades (TTI mobility report), and is only likely to grow more severe in coming years. Broadly speaking, congestion can be divided into recurring and nonrecurring components, the former occurring due to systematic capacity shortages, such as during peak periods, and the latter occurring due to less predictable causes, such as incidents or inclement weather. It is well-known that nonrecurring congestion causes a substantial amount of total delay, with estimates ranging from 13–30% (Skabardonis et al., 2003) to over 50% (Lindley, 1987); furthermore, its unpredictability imposes additional burden on travelers because travel times become unreliable. As a result of uncertain travel times, freight shipments may be late, bus transit services become less attractive, and commuters may leave earlier than is often necessary, losing time that could be more productively spent in order to hedge against late arrival at work.

At the same time, researchers and practitioners are aware of the costs of uncertain travel, and have begun to adapt their methodologies accordingly. For example, in logistics, adaptive and stochastic shortest path algorithms (Polychronopoulos and Tsitsiklis, 1996; Miller-Hooks, 2001; Waller and Ziliaskopoulos, 2002; Gao, 2006) allow vehicle routes to be updated in response to travel information. In transportation planning, incorporating the value of travel reliability has been found to significantly enhance mode choice models (Small et al., 2005; Pinjari and Bhat, 2006; Liu et al., 2007). From the perspective of adaptive congestion pricing, properly accounting for uncertain conditions is needed to ensure optimum conditions obtain (Kobayashi and Do, 1996; Lindsey, 2008; Boyles et al., 2009). When considering the network-level impacts of information-providing devices (Boyles, 2009; Unnikrishnan and Waller, 2009; Boyles and Waller, 2009), distributions on travel time or other measures of delay are also crucial.

Although macroscopic in scope, all of these models rely on facility-level descriptions of uncertain travel, often requiring an explicit probability distribution for travel time or speed as an input. In some locations, traffic detectors record and archive data which may be used to generate these distributions directly. However, coverage is often sparse, commonly existing only on major freeways in metropolitan areas (Lomax et al., 2003). As a result, while such data is invaluable for beginning to study facility reliability, additional modeling is needed to estimate distributions for *all* freeways in a region.

In general, one is concerned with the distribution of one or more *operational metrics*, defined as any physical quantity describing traffic flow on a single roadway segment. This includes quantities such as capacity, free-flow speed, average speed, travel time, queuing delay, average capacity reduction due to an incident, and heavy-vehicle proportion, while excluding measures such as total system travel time (which describes an entire network, not a single segment), roadway geometry (which describes the facility itself, not traffic flow), and the average annual number of icy ideas (which describes the region). While the latter two have an impact on operational measures, and are included as part of the analysis, they are not the principal quantities of interest. Perhaps the distribution of speed (or equivalently, travel time) is the most common operational metric of interest. For this reason, speed is commonly used in this report whenever a specific operational metric would prove useful as an illustration, although the techniques are more generally applicable.

The main contribution of this report is the development of a statistical procedure to generate probability distributions for operational metrics, even where no archived data is available, based on facility-specific attributes. Broadly speaking, the causes of this uncertainty can be divided into those related to roadway capacity (the *supply-side*: incidents, weather, etc.) and those related to travel behavior (the *demand-side*: daily demand fluctuations, special events, etc.) This division is not absolute: some events, such as inclement weather, both reduce roadway capacity and demand for travel. From the perspective of analysis, facility-level field data is much more readily available for the supply-side than for the demand-side, so a two-phase procedure is adopted in this research:

1. A model for generating distributions to represent supply-side sources of uncertainty is developed and calibrated using field data; the output of this procedure may be used directly, or refined using a second-phase demand-side analysis.
2. If data on travel demand variability is available, an analytical, macroscopic demand-side model is provided, which builds on the above supply-side analysis to allow this source of uncertainty to be more accurately considered.

These results can be applied in several ways: first, they can be used to generate more complete statistics of travel reliability in a region; second, they can be used to identify "sensitive facilities" which are especially subject to high variability due to capacity disruptions or demand fluctuations, which is useful when locating variable message signs or other infrastructure; and finally, they can be used as input to the logistics, planning, and pricing models mentioned above.

The remainder of this report is organized as follows: first, a review of relevant research is described. Following this, the general methods of the supply-side and demand-side models are presented. The next sections describe the data sources and specific procedures used, as well as the modeling results themselves. The paper concludes with a summary of the key results, and directions for future research.

BACKGROUND

The research literature contains relatively little on constructing complete probability distributions, often emphasizing estimation of statistics such as confidence intervals or the proportion of late trips (Lomax et al., 2003) or quantifying the proportion of delay attributed to nonrecurring causes (Skabardonis et al., 2003; Lindley, 1987). Still, considerable research has been conducted on several related problems.

Predicting the impact of incidents requires two distinct efforts: estimating the effects of incidents that occur, and estimating the likelihood of incidents in the first place. Regarding the former, researchers have employed analytical approaches based in traffic flow theory (Wirasinghe, 1978; Morales, 1986; Boyles and Waller, 2007), as well as statistical approaches based on field data (Golob et al., 1987; Garib et al., 1997). These are often coupled with models predicting the duration of incidents, for which a number of statistical techniques have been applied, using linear regression (Garib et al., 1997), Poisson regression (Jones et al., 1991), nonparametric regression (Smith and Smith, 2002), hazard-based models (Nam and Mannering, 2000), decision trees (Ozbay and Kachroo, 1999), and Bayesian methods (Ozbay and Noyan, 2006; Boyles and Waller, 2007). Multiple techniques also exist for estimating incident frequency, often as functions of roadway geometry, weather, and flow (Karlaftis and Golias, 2002; Golob and Recker, 2003).

The Highway Capacity Manual (Transportation Research Board, 2000) provides some guidance on the impact of poor weather, suggesting reductions in both capacity and free flow speed as a result of rain, snow, or fog, based on previous research into these factors (Lamm et al., 1990; Ibrahim and Hall, 1994; Hogema et al., 1994; Aron et al., 1994). This information can be combined with regional historical weather data to estimate both the frequency of these events, as well as their impact on the transportation system.

Researchers have also considered the impact of demand fluctuations; almost by necessity, these approaches are macroscopic in nature, rather than facility-specific. The effect of day-to-day demand variations has been studied using simulation techniques (Asakura and Kashiwadani, 1991), equilibrium sensitivity analysis (Bell et al., 1999), and statistical techniques (Clark and Watling, 2005), providing some initial insight on how to model this phenomenon.

GENERAL METHODS

This section presents a general method for estimating distributions of operational metrics on freeways, first considering the impact of supply-side uncertainty (e.g., due to incidents) and then considering demand-side uncertainty (e.g., due to special events). A large difference exists in the availability and accuracy of facility-level data (which is routinely collected in the field through permanent detectors) and demand data (which is inherently macroscopic and difficult to directly measure in terms of mapping observable data to an origin/destination trip table). For this reason, the supply-side model is founded in the statistical analysis of facility-level data, while the demand-side model is treated as a refinement of the supply-side model when relevant data is available.

Supply-Side Analysis
1. Identify set of freeway states.
2. For each detector location and state, identify the best-fitting distribution and the likelihood-maximizing parameters.
3. Use rank sums to identify the best-fitting distribution for each state over all locations.
4. Regress best-fitting distribution parameters against roadway characteristics at detector locations.
5. For target locations, identify distributions for each state based on regression results.
6. Obtain unconditional distribution from law of total probability.

Demand-Side Analysis
1. Estimate OD-link proportions.
2. Identify intrinsic and extrinsic sensitivities.
3. Use tangent plane approximation to refine formulas for moments of the unconditional distribution.
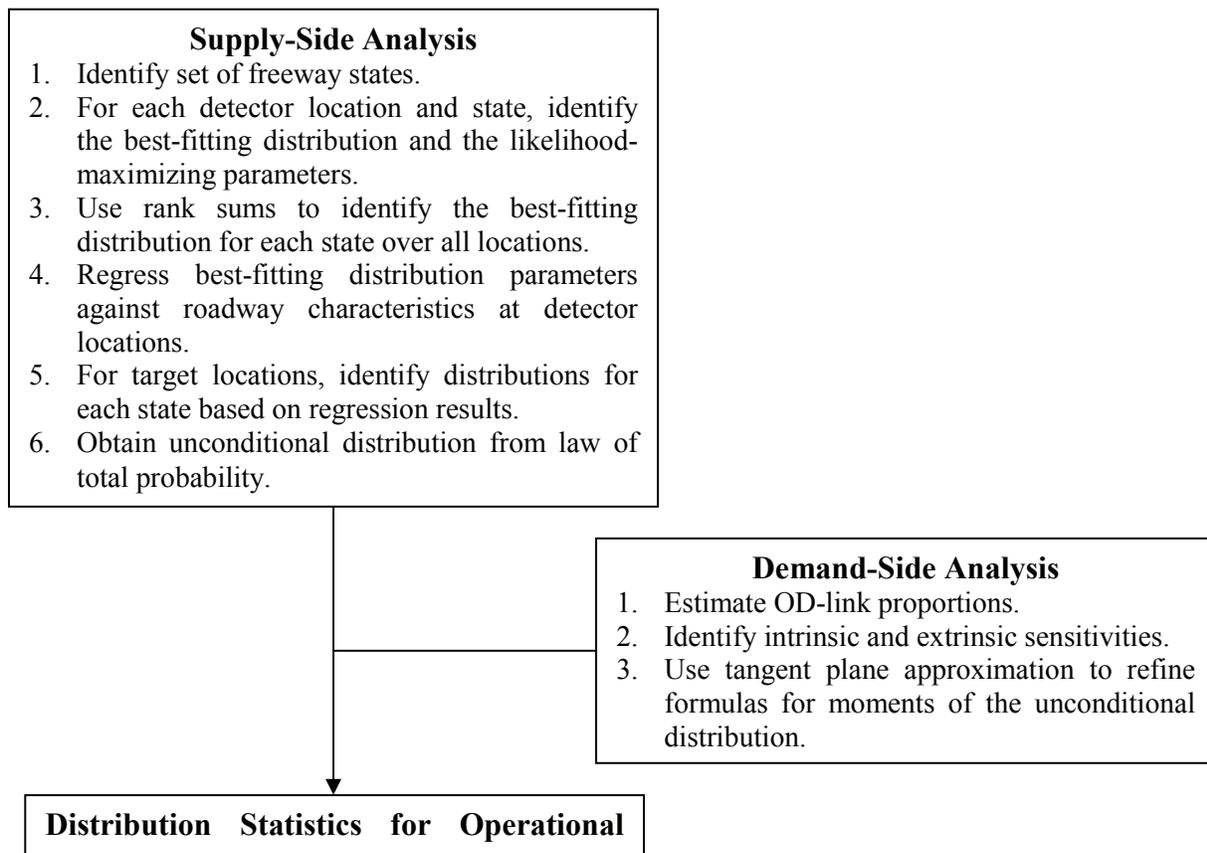
Distribution Statistics for Operational

Figure 1. Overview of analysis procedure

Supply-Side Uncertainty

Many urban areas routinely collect detailed operational traffic data on freeways, using induction loop detectors, side-fire radar, and other technologies. Waller et al. (2008) provide an extensive review of data collection and archival methods.   For arcs where this data is available, an empirical distribution of travel time, speed, volume, and other measures is directly available; the importance of the methods in this section is estimating distributions of operational metrics when direct observations are not available. Because transferability is a central requirement of this method, the procedure distinguishes between different freeway "states" (such as the presence of an incident or poor weather), disentangling the likelihood of an incident from the magnitude of its effect on operational metrics and allowing heterogeneity in both quantities.

Briefly, for a given operational metric, the procedure uses the available data to identify the family of probability distribution (normal, gamma, etc.) which best describes the observed variation in that metric within each state, and uses a regression model to relate these distributions' location and shape parameters to roadway characteristics, such as geometry and position within the network.   These regression models can then be used to estimate the parameters for these distributions on any freeway arc. Figure 1 summarizes the steps in the process, explained below in detail.

The first step is to identify the relevant states S that freeway arcs can exist in. The appropriate state definition depends primarily on data availability (one must know the prevailing state for each travel data observation), but also on modeling scope, geographic resolution, and other factors. For concreteness, four possible states might be ``no incident, good weather'' (NIGW), ``no incident, poor weather'' (NIPW), ``incident present, good weather'' (IPGW) and ``incident present, poor weather'' (IPPW) – although the procedure is certainly not restricted to this configuration.

Concurrently, one should specify the operational metric of choice. This metric should be readily calculable from the available data. If the metric is directly observed, as is often the case with traffic volume or speed, each measurement can be treated as a separate observation. Other metrics may require additional filtering of the data set: if one is interested in free-flow speed, the data set should be restricted to observations with very low volume. If one is interested in the dependence of roadway capacity on freeway state, one option is to consider the highest recorded volume during each state, at each detector location, or a high percentile if outliers are a concern.[1]

At this time, it is also appropriate to consider the roadway characteristics which will be used as explanatory variables for the regression. Potential factors to include are lane width, shoulder width, number of lanes, interchange spacing, lane position of the detector, distance from the city center, proximity to weaving sections or other bottlenecks, speed limits, roadway curvature, roadway grade, peak hour factors, the proportion of heavy vehicle traffic, and any other roadway characteristic that could influence its reliability. Clearly, the exact set of factors which will be used depend on data availability, and the factors deemed most significant in a particular region.

---

[1] More sophisticated procedures are clearly possible. One alternative, based on the fundamental relation $v = uk$ between volume v, space-mean speed u, and density k, is to consider clusters of neighboring observations and seek points where $dv/dk = k(du/dk) + u$ nearly vanishes, i.e., where $du/dk \approx -u/k$.

Next, the necessary data sets must be assembled, including the travel data and the operational metrics, which must be partitioned according to the freeway state at the time of measurement. With the state definitions used in this example, incident logs and historical weather records must be consulted in order to classify observations of the operational metric into these categories.

The best-fitting distributions are then identified using the following procedure: let $\mathbb{P}$ be the set of candidate probability distribution families $f(\phi; \mathbf{x})$ (with $\mathbf{x}$ the parameters for each distribution), and A_D the set of arcs with detector data. Then, for each arc $a \in A_D$, for each state $s \in S$, and each distribution family $\mathcal{P} \in \mathbb{P}$, identify the distribution parameters maximizing the likelihood of the observed sample. A chi-squared statistic $(\chi^2)_{as}^{\mathcal{P}}$ can then be calculated, representing the goodness-of-fit for this distribution.

These are used to generate a numerical ranking $R_{as}^{\mathcal{P}}$ of the distributions, where the lowest rank is associated with the best-fitting distribution, that is, the lowest $(\chi^2)_{as}^{\mathcal{P}}$. A rank-sum $R_s^{\mathcal{P}} = \sum_{a \in A_D} R_{as}^{\mathcal{P}}$ is then calculated for each distribution and each state; the best-fitting distribution for each state is the one with the lowest rank-sum. Using a rank-sum, rather than the chi-squared statistics themselves, provides a measure of robustness against outlying results.

Equipped with the best-fitting distributions $f_s^*(\phi; \mathbf{x})$ for each freeway state $s \in S$, the likelihood-maximizing parameters $\mathbf{x}_s^*$ from each detector location are then retrieved – for example, this would include the mean and the variance for the normal distribution, and two shape parameters for the gamma distribution – and regressed against the roadway characteristics to identify the relationship between these and the travel time distributions. These regression models can then be applied to all freeway arcs (not just those in A_D), using their physical characteristics to estimate the distributions within each state.

Last but not least, the probability that each state $s \in S$ occurs must also be calculated. In the case of weather, simple consultation of historical observations should suffice, with the probabilities more or less constant across the network. For incidents, it may be desirable to apply one of the models developed in the previous literature which relate incident frequency to roadway geometry and other factors, and incident probabilities calculated individually on different arcs. In general, depending on the state definition, an additional regression model may be needed to relate the state probabilities to roadway characteristics. Putting this together, an unconditional density function for ϕ is given by

$$f(\phi) = \sum_{s \in S} f_s^*(\phi; \mathbf{x}_s^*) \Pr(s = S) \tag{1}$$

from the law of total probability.

Demand-Side Uncertainty

Variability in travel times is caused not only by fluctuations in roadway capacity; variation in demand for travel also plays a significant role, especially when identifying "sensitive facilities" which are especially susceptible to unreliable conditions. This section provides methods to refine the supply-side analysis developed in the previous section, in order to account for this

phenomenon. It is more difficult to account for demand uncertainty, since travel demand is harder to observe, and since it occurs at the macroscopic level, rather than at the level of individual corridors or facilities. Thus, these methods are necessarily more analytical and approximate in nature, and additional assumptions must be made. As mentioned previously, the supply-side and demand-side analyses can function either independently or together, depending on available data and scope of the application.

Demand is inherently macroscopic in nature, since it is rooted in individuals' desire to travel from one location to another, possibly distant, location. Further, these individuals choose routes which may span multiple regions and corridors of the transportation system. Thus, it is impossible to rigorously account for demand uncertainty without taking a correspondingly macroscopic modeling perspective, in which travel delay is represented by cost functions mapping demand for travel on a roadway segment to the corresponding travel time. Let the transportation network be described by a graph $G = (N, A)$ consisting of a set N of nodes and a set A of directed arcs. For each arc $(i, j) \in A$, let x_{ij} represent the demand for travel on this arc. Let travel demand from node r to node s be given by the random variable $\tilde{D}_{rs}$. Let $\phi_{ij}(x_{ij})$ represent the operational metric of interest on arc (i, j) as a function of x_{ij}.

The degree to which links are affected by demand uncertainty depends on two primary factors. First, the "typical" operating condition must be considered to determine the local sensitivity of ϕ_{ij} to changes in x_{ij}, an effect we term *intrinsic sensitivity*. Second, the network structure must be considered: where alternate routes exist, links are less sensitive to demand fluctuations than where there is no viable alternative; this effect we term *extrinsic sensitivity*. Figure 2 shows these two effects for the case of average travel speed. Each of these is discussed in turn, and then combined into a single measure. The resulting formulas will allow the mean and variance of ϕ_{ij} to be estimated, incorporating demand uncertainty.

The intrinsic sensitivity is defined as $d\phi/dx$, evaluated at the typical demand value x_0. Quantifying extrinsic sensitivity involves relating uncertainty in macroscopic demand to the uncertainty in demand for an individual arc. Following Clark and Watling (2005) and Unnikrishnan (2008), let q_{rs}^{ij} denote the proportion of travelers departing origin node r, and using arc (i, j) *en route* to destination node s. Thus, the actual demand for individual arcs is itself random, and given by
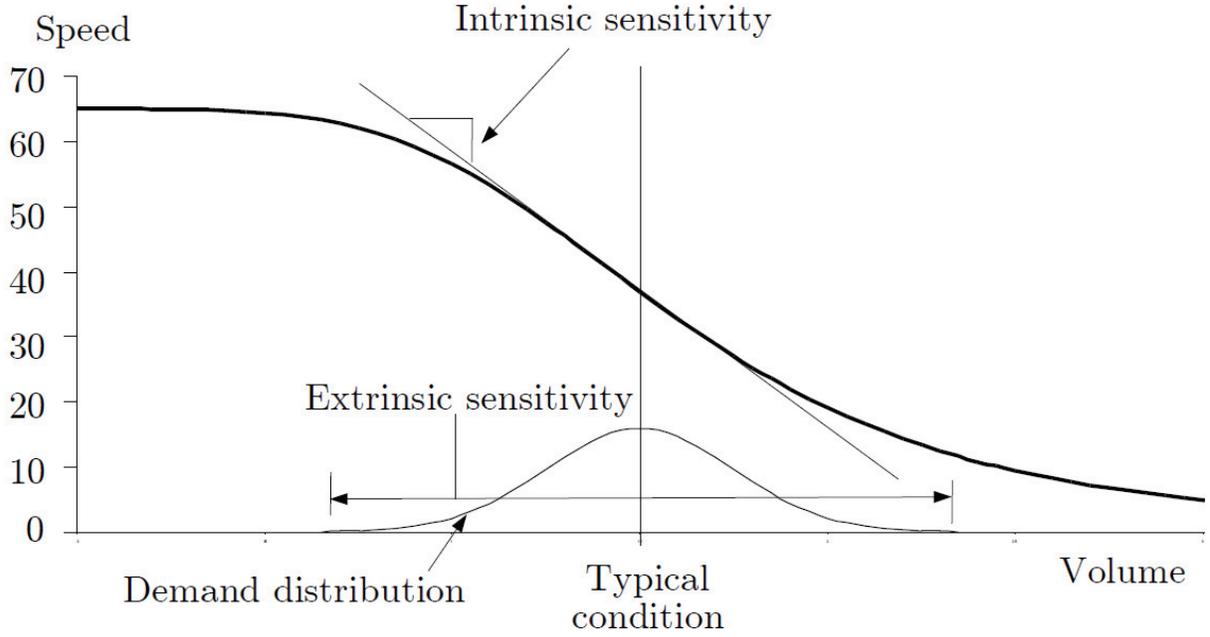
Figure 2. Comparison of intrinsic and extrinsic sensitivity for travel speed.

$$\tilde{x}_{ij} = \sum_{(r,s)\in N\times N} \tilde{D}_{rs} q_{rs}^{ij} \tag{2}$$

Thus, the mean and variance of arc flows are given by

$$\mu_{ij} = E\left[\tilde{x}_{ij}\right] = \sum_{(r,s)\in N\times N} q_{rs}^{ij} E\left[\tilde{D}_{rs}\right] \tag{3}$$

and

$$\sigma_{ij}^2 = Var\left[\tilde{x}_{ij}\right] = \sum_{(r,s)\in N\times N} \sum_{(t,u)\in N\times N} q_{rs}^{ij} q_{tu}^{ij} Cov\left[\tilde{D}_{rs}, \tilde{D}_{tu}\right] \tag{4}$$

which simplifies to

$$\sigma_{ij}^2 = \sum_{(r,s)\in N\times N} \left(q_{rs}^{ij}\right)^2 Var\left[\tilde{D}_{rs}\right] \tag{5}$$

if origin-destination (OD) demands are statistically independent. In general, it is difficult to derive the exact probability density function for arc flows, as it requires a large multiple integral ($O(n^2)$ integrations per arc), however, several special cases are worth noting:

- If OD demands are independent and normally distributed, is also normally distributed with the mean and variance as given above.

9

- If OD demands are independent and given by a Poisson distribution, is also given by a Poisson distribution with rate parameter where λrs is the rate parameter for OD pair (r, s).

Applying relation (5) and the chain rule, extrinsic and intrinsic sensitivity can be combined, deriving the sensitivity of ϕ_{ij} to a change in demand from an arbitrary OD pair (r, s) to be

$$\frac{d\phi_{ij}}{dD_{rs}} = \frac{d\phi_{ij}}{dx_{ij}}\frac{dx_{ij}}{dD_{rs}} = \frac{d\phi_{ij}}{dx_{ij}}q_{rs}^{ij} \tag{6}$$

Taking a tangent plane approximation to s at the point $\mathbf{D}$, the value of ϕ_{ij} resulting from a change in demand $\mathbf{\Delta D}$ can be approximated by

$$\phi_{ij}(x_{ij}^0) + \frac{d\phi_{ij}}{dx_{ij}}\sum_{(r,s)\in N\times N} q_{rs}^{ij}\Delta D_{rs} \tag{7}$$

Given some density function $g_{ij}(x)$ for demand on link (i, j), and taking the linear approximation at the point $x_0 = E[x]$, and fixing a value ϕ_0, the mean and variance of travel speed can be estimated as

$$E[\phi|\phi_0] \approx \int \left(\phi_0 + \phi'(x_0)(x - x_0)\right)g_{ij}(x)\,dx = \phi_0 + \phi'(x_0)(E[x] - x_0) = \phi_0 \tag{8}$$

$$
\begin{aligned}
Var[\phi|\phi_0] &\approx \int \left(\phi_0 + \phi'(x_0)(x - x_0)\right)^2 g_{ij}(x)\,dx - \phi_0^2 \\
&= (\phi_0 - \phi'(x_0)x_0)^2 + 2(\phi_0 - \phi'(x_0)x_0)E[x] + [\phi'(x_0)]^2 E[x^2] - \phi_0^2 \\
&= [\phi'(x_0)]^2 Var[x]
\end{aligned} \tag{9}
$$

This can now be combined with the supply-side analysis from the previous section, which derived a density function $f_{ij}(\phi)$ for the operational metric on each link (i,j). Fixing x_0, the derivative $\phi'(x_0)$ still depends on s, implying that the above formulas condition on a given "base" condition. Using the law of total variance, we can write *un*conditional expressions for these quantities:

$$
\begin{aligned}
E[\phi] &= \int E[\phi|\phi_0]f(\phi_0)\,d\phi_0 = \int \phi_0 f(\phi_0)\,d\phi_0 \\
&= \mu_\phi
\end{aligned} \tag{10}
$$

$$Var[\phi] = Var\big[E[\phi|\phi_0]\big] + E\big[Var[\phi|\phi_0]\big] = Var[\phi_0] + E\left[\big(\phi'(\phi_0)\big)^2 Var[x]\right]$$

$$= \sigma_\phi^2 + E\left[\big(\phi'(x_0)\big)^2\right]\sum_{(r,s)\in N\times N}\big(q_{rs}^{ij}\big)^2 Var[\bar{D}_{rs}] \tag{11}$$

To give a concrete example, consider the average travel speed u. In a macroscopic sense, travel time functions are often assumed as a function of x. Thus, we can apply the transformation $u =$

L/t, with u the travel speed, L the arc length, and t the traversal time. The intrinsic sensitivity of speed to a change in demand can be represented by the derivative

$$\frac{du}{dx} = -\frac{L}{t^2}\frac{dt}{dx} = -\frac{u^2}{L}\frac{dt}{dx} \tag{12}$$

For instance, a commonly-used travel time function, developed by the Bureau of Public Roads, is

$$t = t_0\left(1 + \alpha\left(\frac{x}{c}\right)^\beta\right) \tag{13}$$

with t_0 the free-flow speed, c the capacity, and α and β calibrated parameters. Using this function produces

$$\frac{du}{dx} = -\frac{t_0\alpha\beta}{Lc^\beta}x^{\beta-1}u^2 \tag{14}$$

indicating that this link is "robust" to demand uncertainty if either the demand d or travel speed s is low – that is, changes in demand have lesser effect if few people are using the link (close to free-flow), or if the link is already highly congested (speed will not degrade much further) – and is most "sensitive" when the quantity $x^{\beta-1}u^2$ is maximized. This can be generalized; almost all commonly-used delay functions are increasing and convex; that is, dt/dx is positive and increasing in x, so du/dx as defined by (1) is small when x (and thus dt/dx) is small, or when s is small.

Substituting into (6), the sensitivity of travel speed on an arc (i, j) to a change in demand from an arbitrary OD pair (r, s) is

$$\frac{du_{ij}}{dD_{rs}} = -\frac{u_{ij}^2}{L_{ij}}\frac{dt_{ij}}{dx_{ij}}\frac{dx_{ij}}{dD_{rs}} = \frac{du_{ij}}{dx_{ij}}q_{rs}^{ij} \tag{15}$$

For the BPR relation, the following tangent plane approximation is

$$u - \frac{t_0\alpha\beta}{Lc^\beta}x^{\beta-1}u^2 \sum_{(r,s)\in N\times N} q_{rs}^{ij}\Delta D_{rs} \tag{16}$$

Thus

$$E[u] = \int E[u|u_0]f(u_0)du_0 = \int u_0 f(u_0)du_0 = \mu_{ij} \tag{17}$$

$$Var[u] = \sigma_{ij}^2 + \left(\frac{t_{ij}^0\alpha\beta}{L_{ij}c_{ij}^\beta}x^{\beta-1}\right)^2 E[u^4]\sum_{(r,s)\in N\times N}\left(q_{rs}^{ij}\right)^2 Var[\tilde{D}_{rs}] \tag{18}$$

where $E[u^4]$ is the fourth raw moment of the speed distribution found from the supply-side analysis (1), using equations (8) and (9). Certainly, other travel time functions can be used as well, such as those given in the Highway Capacity Manual (Transportation Research Board,

2000); the latter may be more appropriate when only link volumes are available, rather than link demands from a planning model.

DEMONSTRATIONS

The remainder of the paper provides two demonstrations of the above methods in different applications. The first application involves estimation of freeway speed distributions, using data from the Dallas-Ft. Worth metropolitan area in Texas. The second involves identifying the appropriate capacity reduction to apply for planning models which can account for incidents (as in Unnikrishnan and Waller, 2009), using data from Seattle, Washington. For reasons of brevity, the demand-side procedure is only presented for the first application.

Dallas-Ft. Worth

Three main sources of data were obtained: a set of archived loop detector observations providing speed data; a set of logs detailing incident locations, times, and durations; and a set of weather data providing information on temperature and precipitation, for approximately nine months between January 1 and September 10, 2007. The Dallas Traffic Management System, operated by the Texas Department of Transportation, provides archived five-minute loop detector data (speed, occupancy, volume, long vehicle volume) on a publicly available website.[2] Data from this archive are stored in a separate file for each day; these were converted into detector-specific files for use in this project.

Incident logs for this region were also obtained from the Texas Department of Transportation. The incident logs contain information regarding the location, time and duration of the incident. The location of the incident is described using the name of the road, direction of traffic flow and the approaching intersection. Based on this information the corresponding detector is identified. Once the corresponding detector has been identified, the incident information is merged with the loop detector data by matching the date and time of the incident. Based on the coverage of this data set, and of the loop detector data, a set of seventy-two detectors at nineteen locations was identified, and extracted for further analysis. Finally, daily weather observations were obtained from the National Weather Forecast Office.[3] Precipitation information was extracted from this data set, and merged with the loop detector data.

These data sets were merged into a common file, and divided into three segments: **no incident, good weather** (NIGW), **poor weather** (PW), and **incident present** (IP). Approximately 88.5% of observations correspond to the NIGW state, 9.7% to the PW state, and 1.7% to the IP state. As relatively few (less than 0.2%) observations existed for cases where *both* an incident and poor weather conditions were present, the choice was made to omit this latter category, and classify any such observations under both the PW and IP categories. Furthermore, this data set only contained weather data at the resolution of one day, so "poor weather" was applied to all observations on days with a half inch or more of precipitation. This suffices for a demonstration; for field application, a more disaggregate data source would be highly valuable.

For each detector and each category (NIGW, PW, IP), fifteen different probability distributions comprised the set of candidates $\mathbb{P}$: the normal, lognormal, beta, chi-squared, Erlang, exponential, fatigue life, Frechet, gamma, generalized extreme value, Gumbel, logistic, log-logistic, Rayleigh,

---

[2] http://ttidallas.tamu.edu/detectordataarchive/DalTrans/Default.htm. Accessed March 20, 2008.
[3] http://www.srh.noaa.gov/fwd/f6.htm. Accessed April 20, 2008.

and Weibull distributions. As described in the previous section, these rankings are aggregated across detectors within each state. Based on this process, the normal distribution is seen to best describe speeds for the NIGW category, while the beta distribution best describes speeds for the PW and IP categories. Table 1 shows an example of the rankings for one particular detector in the IP state; Figure 3 shows the beta distribution fitted to that detector's data along with a histogram of observed speeds.
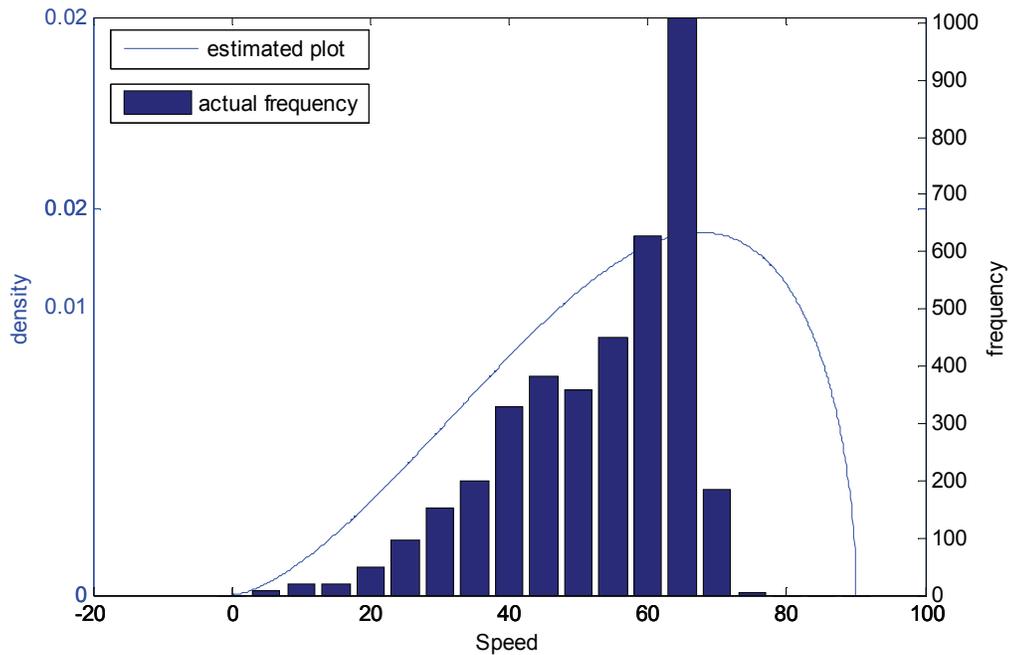

Figure 3. Fitted beta distribution versus actual frequency distribution

Table 1. Sample ranking of distributions for a particular location and state.

Distribution	Chi Squared Statistic	Rank
Beta	9.826	1
Log-Logistic	10.297	2
Weibull	14.62	3
Fatigue Life	14.802	4
Gamma	14.966	5
Lognormal	15.516	6
Erlang	19.212	7
Gumbel Max	22.481	8
Gen. Extreme Value	25.341	9
Normal	25.541	10
Frechet	31.559	11
Logistic	34.933	12
Rayleigh	42.646	13
Chi-Squared	68.756	14
Exponential	119.66	15

These distributions are specified by two parameters each: the mean μ and standard deviation σ for the normal distribution, which has density

$$f(u; \mu, \sigma) \frac{1}{\sigma\sqrt{2\pi}} e^{-(u-\mu)^2/(2\sigma^2)} \tag{19}$$

and two shape parameters a and b for the beta distribution, which has density

$$f(u; a, b) = \frac{u^{a-1}(1-u)^{b-1}}{B(a, b)} \qquad a \leq x \leq b; p, q \geq 0 \tag{20}$$

where B is the beta function. Additionally, a linear transformation is applied to the beta distribution to scale its support to [0, 90], with speeds measured in miles per hour.

In order to estimate speed distributions at locations where no data is present, we attempt to link these distribution parameters to roadway geometry and segment characteristics using linear regression. Drawing inspiration from the Highway Capacity Manual procedure for calculating free-flow speed on freeway segments (Transportation Research Board, 2000), the independent variables chosen for regression were the lane width, shoulder width, number of lanes, interchange spacing, and lane position (that is, whether the detector is located on an inner or outer lane), while the dependent variables were the best-fitting distribution parameters for each detector.

Three such regressions were performed, for the NIGW, PW, and IP scenarios; regression results are shown in Tables 2 and 3, with *t*-statistics given in parentheses. One notable result is that the

R^2 values are very low for the NIGW and PW scenarios, indicating that these geometric factors do not play a significant role in determining speed distributions when no incident is present. The past literature clearly shows that geometry affects the probability of an incident occurring, and this data reveals that geometry also affects the severity of an incident (as measured by the resulting speed distributions); but this initial investigation suggests that the impact is limited beyond this. That is to say, the influence of geometry is much more pronounced in the state probabilities $\Pr(s = S)$ than in the state-specific distributions $f_s(u)$,

Using this procedure, three probability distributions $f_{NIGW}(u)$, $f_{IP}(u)$, and $f_{PW}(u)$, can be constructed for any freeway segment, using its geometric properties to choose the distribution parameters. These can then be combined into an unconditional speed distribution using the law of total probability:

$$f(u) = f_{NIGW}(u)\Pr(NIGW) + f_{IP}(u)\Pr(IP) + f_{PW}(u)\Pr(PW)$$

(21)

The probability of an incident Pr(*IP*) as a function of roadway characteristics can be calculated using any of the models described in the literature review; the probability of poor weather Pr(*PW*) can be approximated using past records; and the remaining probability Pr(*NIGW*) can be calculated as *Pr(NIGW)* = 1 – Pr(*IP*) – Pr(*PW*). (Note that by neglecting the case with both an incident and poor weather, these equations are only approximate.) Figure 4 illustrates how these conditional density functions can be combined into an overall, unconditional probability density function for eastbound I-635 at Josey Lane. Note that the unconditional density most closely resembles the NIGW density, which is the most probable state. In particular, using the geometric characteristics of this site, the density functions are

$$f_{NIGW}(u) = 0.03507 e^{-(s-46.36)^2/258.8}$$

(22)

$$f_{IP}(u) = \frac{u^{1.815}(90-u)^{3.020}}{90^{5.835} B(2.815,4.02)}$$

(23)

$$f_{PW}(u) = \frac{u^{2.784}(90-u)^{2.375}}{90^{6.159} B(3.874,3.375)}$$

(24)

Table 2. Output from the linear regression models for the NIGW scenario

Coefficient\Parameter	mean	t-statistic	std. dev.	t-statistic
Constant	41.23	10.12	15.51	0.45
Outer lane dummy	-	-	–40.11	–1.2
Number of lanes	-	-	47.08	1.32
Shoulder width (m)	5.13	3.1	-	-
R-squared	0.2		0.06	

Table 3. Output from the linear regression models for the IP and PW scenarios

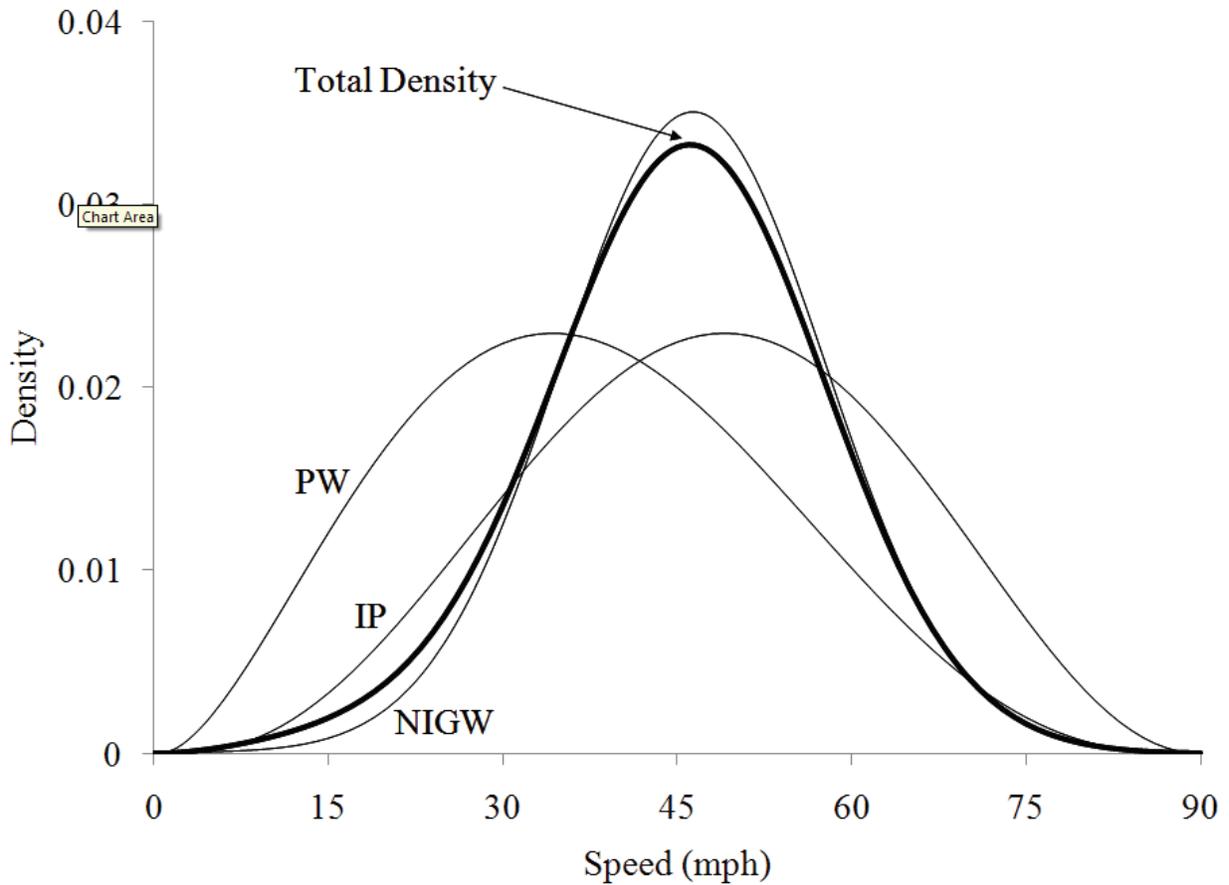| Coefficient\Parameter	IP case				PW case			
	a	*t*-stat	*b*	*t*-stat	*a*	*t*-stat	*b*	*t*-stat
Constant	33.07	4.23	25.73	3.93	6.925	2.98	5.19	3.51
Outer lane dummy	-	-	1.69	1.52	-	-	0.9	2
Number of lanes	−6.09	−3.9	−5.85	−4.03	−0.698	−1.2	−0.51	−1.1
Shoulder width (m)	2.12	5.3	1.79	1.7	-	-	-	-
Interchange spacing (mi)	−14.18	1.83	−3.78	1.5	-	-	-	-
R-squared	0.49		0.46		0.03		0.1	

Figure 4. Conditional and unconditional probability density functions

with the total density given by equation (21). From this, the predicted mean and standard deviation of travel speed at this segment are 45.4 mph and 14.8 mph, respectively, compared to an observed mean and standard deviation of 40.0 mph and 15.3 mph (Table 4). Thus, the regression models overestimated the travel speed by approximately 13.5%, and underestimated the standard deviation by 3.10%.

A more refined estimate is possible using the demand-side procedure. In the absence of a calibrated demand model, we use the observed volumes as a proxy for $\sum_{(r,s) \in N \times N} \tilde{D}_{rs} q_{rs}^{ij}$, and thus $\sum_{(r,s) \in N \times N} \left(q_{rs}^{ij} \right)^2 Var\left[\tilde{D}_{rs} \right]$ is simply the variance in observed volumes (17.8×10^6 veh^2/hr^2). Because we are using volume, rather than demand, we use the speed formula from the Highway Capacity Manual (Transportation Research Board, 2000):

$$u = FFS - \left[\frac{1}{9}(7FFS - 340)\left(\frac{v_p + 30FFS - 3400}{40FFS - 1700} \right)^{2.6} \right] \qquad (25)$$

based on the observed free-flow speed $FFS = 68$ mph obtained from the detector data and the effective volume v_p (almost exactly the per-lane volume, owing to the low heavy vehicle count reported at this location). Substituting the derivative of (25) into (11), the new estimate of the standard deviation of travel speed is 15.2 mph, which matches the observed standard deviation of 15.3 mph to 0.5%, roughly six times more accurate than the estimate given by the supply-side analysis alone

.

Table 4. Numerical summary for Dallas-Ft. Worth, I-635 EB at Josey Lane.

Statistic	Observed	Base estimate (supply-side only)	Refined estimate (supply + demand)
Mean	40.0 mph	46.4 mph	46.4 mph
Standard deviation	15.3 mph	14.8 mph	15.2 mph

Seattle

A second demonstration of these procedures can be applied using data from the Seattle, Washington metropolitan area. In this case, the goal is development of a planning network which represents incidents in addition to typical operating conditions, perhaps to be used to identify the locations where variable message signs would be most beneficial, as in Boyles and Waller (2009). To do this, we need to estimate the planning capacity on each freeway in the presence and absence of incidents, in addition to the probability of incidents occurring. Unlike the previous example, here we are interested in a single point estimate of these values, rather than the entire distribution. To keep the example comprehensible, only a sketch network is constructed with a small number of nodes and arcs; for field applications, additional *vraisemblance* is necessary and a more disaggregate analysis is needed. Figure 5 and Table 5 show the network structure along with information on the freeway segments each link represents.

Data sets including volume, speed, and density information on selected freeways in the network, as well as the presence of any incidents, were obtained from TransNOW. In this application, each arc can exist in one of two states – no incident (NI) or incident present (IP). A delay function must be estimated for each arc in each state, which are assumed to be of the BPR form of equation (13). Thus, the free-flow travel time and capacity must be specified for each arc in the network, including those not covered by the data set. For both the NI and IP states, the free-flow travel time is given by the segment length divided by the speed limit, converted to appropriate units.[4]

Regarding capacity, recall that the "capacity" parameter c in the BPR equation does *not* actually represent the true roadway capacity, but merely a shape parameter for converting travel demand on an arc to the experienced travel delay. For the typical BPR parameters $\alpha = 0.15$ and $\beta = 4$, c is often taken to be the "practical capacity" of the roadway, roughly corresponding to level of service E. Kockelman (2003) estimates this to be roughly 80% of the true capacity.

If c_{ij}^{NI} represents the practical roadway capacity of arc (i, j) in the NI state, we express the practical capacity in the IP state as $n_{ij} c_{ij}^{NI}$ where n_{ij} represents the capacity decrease due to the incident's presence. Thus c_{ij}^{NI} and n_{ij} form the operational metrics of interest, along with the probability of an incident occurring $\Pr_{ij}(IP)$. For planning models, a single value is needed, rather than a distribution. Thus, the only candidate distribution in $\mathbb{P}$ is the degenerate distribution, located at the most likely value of these parameters, and the ranking process is not needed.

For each arc, c_{ij}^{NI} is estimated from the data set by finding the maximum observed volume during any 5-minute time interval, scaled to appropriate units, and multiplied by

---

[4] The speed limit for all of the freeways in this network is 60 mph; although average travel speed is often higher than this at near free-flow conditions, the data set truncated speed measurements at the speed limit.
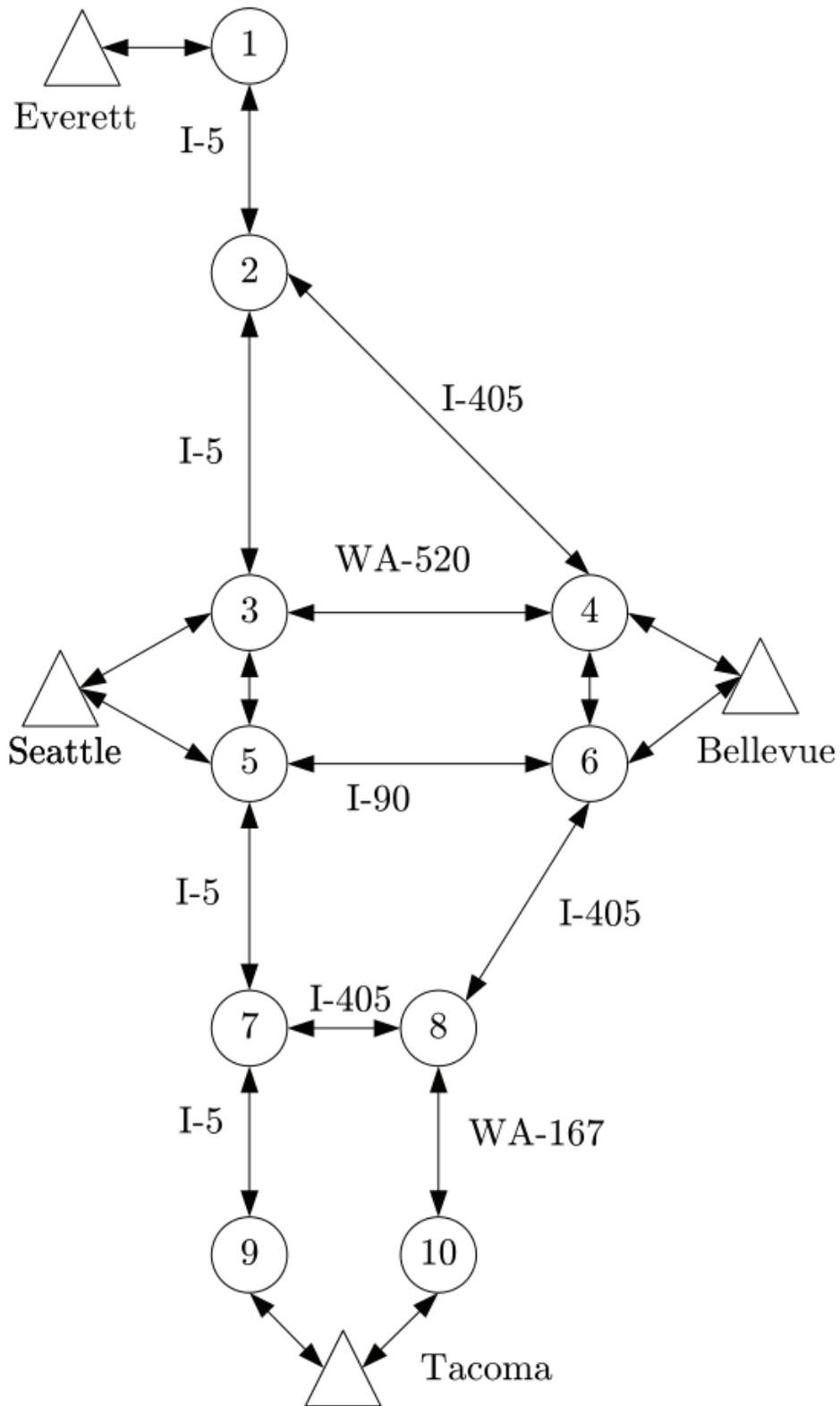
Figure 5. Sketch Seattle network used for demonstration.

Table 5. Seattle data highway segments and network arcs.

Highway segment	Length (mi)	Network arcs
I-5 NB Seatac to Seattle	13	(7,5)
I-5 NB Seattle to WA-526	23	(5,3), (3,2), (2,1)
I-5 SB WA-526 to Seattle	23	(1,2), (2,3), (3,5)
I-5 SB Seattle to Seatac	13	(5,7)
WA-167 NB Auburn to Renton	10	(10,8)
WA-167 SB Renton to Auburn	10	(8,10)
I-405 NB Tukwila to Bellevue	14	(7,8), (8,6)
I-405 NB Bellevue to WA-524	16	(6,4), (4,2)
I-405 SB WA-524 to Bellevue	16	(2,4), (4,6)
I-405 SB Bellevue to Tukwila	14	(6,8), (8,7)

80%. Estimation of η_{ij} is more involved. Simply reducing c_{ij}^{NI} to account for a lane blockage is insufficient for several reasons: the capacity at the bottleneck is not the true quantity of interest, but rather the travel delay on the entire roadway segment. Upstream of the incident, additional delay is incurred; however, downstream of the incident, some travel savings may be obtained due to a metering effect which reduces congestion below its normal level. Instead, the following procedure is used to choose η_{ij} to match the observed decrease in travel speeds during incident conditions.

Suppressing the arc subscript for brevity, let u^{NI} and u^{IP} represent the average travel speeds during these states, and let u_f denote the free-flow travel speed. First, we express the travel demand x producing travel speed u^{NI} by inverting (13) and expressing quantities in terms of speed:

$$x = c^{NI}\left(\frac{u_f/u^{NI} - 1}{\alpha}\right)^{1/\beta} \tag{26}$$

Now, we have

$$\frac{u^{IP}}{u^{NI}} = \frac{1 + \alpha\left(x/c^{NI}\right)^{\beta}}{1 + \alpha\left(x/\eta c^{NI}\right)^{\beta}} \tag{27}$$

and solving for η yields

$$\bar{n} = \frac{x}{c^{NI}} \left[\frac{1}{\alpha} \left(\frac{1 + \alpha (x/c^{NI})^\beta}{u^{IP}/u^{NI}} - 1 \right) \right]^{-1/\beta} \tag{28}$$

Although it may not be obvious from the above procedure, $\bar{n}$ is in fact independent of c^{NI}, and is solely determined by u_f and the ratio u^{NI}/u^{IP}.

At this point, regression is applied to estimate the values of these parameters as functions of roadway characteristics. Given the relative paucity of locations due to aggregation, only three explanatory variables were considered: the number of lanes n_L, the interchange density I_D expressed in interchanges per mile, and the distance from the Seattle city center d_{CBD}, measured in miles. After pruning insignificant variables, the resulting regression equations are

$$c \approx 1670 n_L - 66 \tag{29}$$

$$\bar{n} \approx 0.736 + 0.142 I_D - 0.015 n_L \tag{30}$$

$$\Pr(IP) \approx 0.02 d_{CBD} + 0.11 n_L - 0.54 \tag{31}$$

Further details on the regression can be seen in Table 6. Applying these equations, delay functions can be created for each arc in the Seattle network; all of the relevant parameters are shown in Table 7.

Table 6. Regression results for Seattle data.

Coefficient\Parameter	c^{NI}	t-statistic	$\bar{n}$	t-statistic	$\Pr(IP)$	t-statistic
Constant	-66	-0.13	0.73	13.2	-0.54	-5.08
Distance to CBD (mi)	-	-	-	-	0.02	6.02
Interchange density (mi^{-1})	-	-	0.10	5.93	-	-
Number of lanes	1670	12.70	-0.01	-1.37	0.11	5.93
R-squared	0.95		0.46		0.84	

Table 7. Capacities and link incident capacity reductions for Seattle

Arc	Free-flow time	Base capacity c_{ij}^{NI}	Reduction factor η	Incident capacity c_{ij}^{IP}
(1,2)	7	6864	0.762	5229
(2,1)	7	6864	0.762	5229
(2,3)	14	6864	0.762	5229
(2,4)	16	5895	0.712	4200
(3,2)	14	6864	0.762	5229
(3,4)	10	3825	0.759	2903
(3,5)	4	6864	0.762	5229
(4,2)	16	5722	0.778	4451
(4,3)	10	3825	0.759	2903
(4,6)	3	5895	0.712	4200
(5,3)	4	6864	0.762	5229
(5,6)	10	6614	0.735	4858
(5,7)	13	8762	0.781	6847
(6,4)	3	5722	0.778	4451
(6,5)	10	4944	0.749	3704
(6,8)	9	5609	0.800	4487
(7,5)	13	7577	0.793	6009
(7,8)	2	5994	0.809	4848
(7,9)	12	7449	0.741	5523
(8,6)	9	5994	0.809	4848
(8,7)	2	5609	0.800	4487
(8,10)	10	4079	0.797	3251
(9,7)	12	7449	0.741	5523
(10,8)	10	3950	0.836	3304

CONCLUSION

This paper developed a two-step procedure for estimating distributions of operational metrics on freeways, even at locations where no observed data are present. The main step focuses on factors affecting roadway supply, such as incidents or weather. For these scenarios (as well as a base case), a method for identifying the most appropriate probability distributions was described, with their parameters estimated using linear regression models. Combining the distributions developed for each scenario into one unified, unconditional probability distribution is accomplished using the law of total probability.

If data is also available on variation in macroscopic travel demand, this analysis can be refined using an analytical demand-side procedure, which combines the previously-calculated speed distribution with roadway sensitivity to demand fluctuations, to produce refined estimates of the operational metric distributions, and their means and variances. Two examples were provided; in the Dallas-Ft. Worth example, the data indicated that roadway geometry influences the state probabilities (e.g., of incidents) more so than the speed distributions within each state. For this particular case, the demand-side refinement improved the accuracy of the estimated standard deviation in speed sixfold. In the Seattle example, a sketch planning network was created to reflect the possibility of incidents. Developing a procedure to estimate the operational metric based on data required more effort for this case.

As is, this procedure can be applied to generate input data for transportation and logistics models requiring a probability distribution on speeds, travel times (as in vehicle routing problems), or other operational metrics; or to assist with identifying facilities which are particularly vulnerable to changes in roadway supply and demand, and designing mitigation strategies to address these problems. Still, many directions for further research remain. As the demonstrative examples suggest, identifying the geometric and roadway factors most influencing a particular operational metric is of central importance, and it would be helpful to determine the most important factors for a variety of metrics. Further, the supply-side and demand-side models are based on different first principles (owing to the macroscopic nature of demand modeling, and the microscopic nature of facility-level data), and developing a framework based on a common set of assumptions would be fruitful. Dynamic traffic assignment models have potential to be applied here, but tractable, large-scale analytical formulations of such models remain elusive. Finally, obtaining accurate data on demand variability can be difficult, and developing rigorous procedures in the presence of more easily collected proxies for such data (such as cordon counts) would be highly valuable.

REFERENCES

Aron, M., M. Ellenberg, P. Fabre, and P. Veyre. (1994). "Weather related traffic management." In Towards an Intelligent Transport System: Proceedings of the First World Congress on Applications of Transport Telematics and Intelligent Vehicle-Highway Systems, Vol. 3. Ertico, Brussels, Belgium, 1089-1096.

Asakura, Y. and M. Kashiwadani. (1991). "Road network reliability caused by daily fluctuations of traffic flow." Proc., the 19th PTRC Summer Annual Meeting, Brighton, 1991, pp. 73-84.

Bell, M. G. H., C. Cassir, Y. Iida, and W. H. K. Lam. (1999). "A sensitivity-based approach to network reliability assessment." Proc., the 14th International Symposium on Transportation and Traffic Theory, 1999, Jerusalem, Israel, 283-300.

Boyles, S. and S. T. Waller. (2007). "A stochastic delay prediction model for real-time incident management." ITE Journal 77, 18-24.

Boyles, S. and S. T. Waller. (2009). "Optimal information location for adaptive routing." Accepted for publication in Networks and Spatial Economics.

Boyles, S., D. Fajardo, and S. T. Waller. (2007). "A naive Bayesian classifier for incident duration." Proc., the 86th Annual Meeting of the Transportation Research Board, January 2007, Washington, D.C.

Clark, S. and D. Watling. (2005). "Modelling network travel time reliability under stochastic demand." Transportation Research Part B. 39, 119-140.

Gao, S. (2005). Optimal Adaptive Routing and Traffic Assignment in Stochastic Time-Dependent Networks. Ph.D. thesis, Massachusetts Institute of Technology.

Garib, A., A. E. Radwan, and H. Al-Deek. (1997). "Estimating magnitude and duration of incident delays." Journal of Transportation Engineering. 123, 459-486.

Golob, T. F. and W. W. Recker. (2003). "Relationships among urban freeway accidents, traffic flow, weather, and lighting conditions." Journal of Transportation Engineering. 129, 343-353.

Golob, T., W. Recker, and J. Leonard. (1987). "An analysis of the severity and incident duration of truck-involved freeway accidents." Accident Analysis and Prevention. 19, 375-395.

Hogema, J. H., A. R. A. Vanderhorst, and P. J. Bakker. (1994). Evaluation of the 16 Fog-Signaling System with Respect to Driving Behaviour (Evaluatie Van Het a 16 Mistsignaleringssysteem in Termen Van Het Rijgedrag.) Publication TNO-TM 1994 C-48, TNO Technische Menskunde, Soesterberg, The Netherlands.

Ibrahim, A. T. and F. L. Hall. (1994). "Effect of adverse weather conditions on speed-flow-occupancy relationships." Transportation Research Record. *1457*, 184-191.

Jones, B., L. Janssen, and F. Mannering. (1991). "Analysis of the frequency and duration of freeway accidents in Seattle." Accident Analysis and Prevention. 23, 239-255.

Karlaftis, M. G. and I. Golias. (2002). "Effects of road geometry and traffic volumes on rural roadway accident rates." Accident Analysis and Prevention. 34, 357-365.

Kobayashi, K. and M. Do. (1996). "The informational impacts of congestion tolls upon route traffic demands." Transportation Research Part A. 26, 505-529.

Kockelman, K. M. (2004). "Traffic Congestion." In Handbook of Transportation Engineering, Myer Kutz, ed., McGraw-Hill Professional, New York, USA, Chapter 12.

Lamm, R., E. M. Choueiri, and T. Mailaender. (1990). "Comparison of operating speeds on dry and wet pavements of two-lane rural highways." Transportation Research Record. 1280, 199-207.

Lindley, J. (1987). "Urban freeway congestion: quantification of the problem and effectiveness of potential solutions." ITE Journal. 57, 27-32.

Lindsey, R. (2009). "Cost recovery from congestion tolls with stochastic capacity and demand." Journal of Urban Economics 66, 16-24.

Liu, H. X., X. He, and W. Recker. (2007). "Estimation of the time-dependency of values of travel time and its reliability from loop detector data." Transportation Research Part B. 41, 448-461.

Lomax, T., D. Schrank, S. Turner, and R. Margiotta. (2003). Selecting Travel Reliability Measures. Texas Transportation Institute and Cambridge Systematics.

Miller-Hooks, E. (2001). "Adaptive least-expected time paths in stochastic, time-varying transportation and data networks." Networks. 37, 35-52.

Morales, J. (1986). "Analytical procedures for estimating freeway traffic congestion." Public Roads. 50, 55-61.

Nam, D. and F. Mannering. (2000). "An exploratory hazard-based analysis of highway incident duration." Transportation Research Part A. 35, 85-102.

Ozbay, K. and N. Noyan. (2006). "Estimation of incident clearance times using Bayesian networks approach." Accident Analysis and Prevention. 38, 542-555.

Ozbay, K. and P. Kachroo. (1999). Incident Management in Intelligent Transportation Systems. Artech House, Boston, USA.

Pinjari, A. and C. Bhat. (2006). "On the nonlinearity of response to level of service variables in travel mode choice models." Transportation Research Record. 1977, 67-74.

Polychronopoulos, G. and J. Tsitsiklis. (1996). "Stochastic shortest path problems with recourse." Networks. 27, 133-143.

Skabardonis, A., P. Varaiya, and K. F. Petty. (2003). "Measuring recurrent and nonrecurrent traffic congestion." Transportation Research Record. 1856, 118-124.

Small, K. A., C. Winston, and J. Yan. (2005). "Uncovering the distribution of motorists' preferences for travel time and reliability." Econometrica. 73, 1367-1382.

Smith, K. and B. Smith. (2002). Forecasting the Clearance Time of Freeway Accidents. Publication STL-2001-012, Center for Transportation Studies, University of Virginia, USA.

Transportation Research Board. (2000). Highway Capacity Manual, National Research Council, Washington, DC, USA.

Unnikrishnan, A. and S. T. Waller. (2009). "User equilibrium with recourse." Networks and Spatial Economics. 9(4), 575-593.

Waller, S. T. and A. Ziliaskopoulos. (2002). "On the online shortest path problem with limited arc cost dependencies." Networks. 40, 216-227.

Wirasinghe, S. C. (1978). "Determination of traffic delays from shock-wave analysis." Transportation Research. 12, 343-348.