Probabilistic Graphical Models of Fundamental Diagram Parameters for Simulations of Freeway Traffic

Ajith Muralidharan, Gunes Dervisoglu, and Roberto Horowitz

Freeway traffic simulations must account for the probabilistic nature of model parameters to capture observed variations in traffic behavior. Fundamental diagrams specify freeway section parameters describing the flow–density relationship in macroscopic simulation models. A triangular fundamental diagram—specified with the free-flow speed, congestion wave speed, and capacity—is commonly adopted in first-order cell transmission models. Capacity (defined as the maximum flow observed in a given freeway section over a particular day) exhibits significant day-to-day variation, and capacity variations across different sections of the freeway are significantly correlated. Free-flow speeds do not exhibit significant variation, but congestion wave speeds exhibit variation uncorrelated with section capacities or parameters from other sections. A probabilistic graphical approach is presented to model the probabilistic distribution of fundamental diagram parameters of an entire freeway section chosen for simulation. More than 1 year of data from dozens of loop detectors along a 25-mi section of the I-210 freeway westbound in Los Angeles, California, are used for demonstration. The parameters of the distribution are estimated with the expectation–maximization algorithm to account for missing observations. Model selection from among plausible models indicates that a first-order spatial Markov model is appropriate to capture the capacity distribution, which is the joint probability distribution of freeway section capacities. Stochastic simulations with sampled parameters demonstrate that capacity variations can lead to significant variations in congestion patterns and freeway performance.

Traffic flow simulations offer a cost-effective solution for studying the effects of operational strategies such as ramp metering and the management of traffic demand and incidents (e.g., accidents and lane closures). Macroscopic models simulate aggregate traffic behavior and offer reliable, fast simulations suitable for integration into a real-time management system. Models are calibrated with measured data to specify the parameters and inputs that will replicate observed flow and speed contours as well as performance measures such as vehicle miles traveled (VMT) and vehicle hours traveled (VHT). These models typically are calibrated to replicate 1 specific day of traffic flow. This deterministic method of modeling fails to capture the rich day-to-day variations in traffic flow behavior that can be attributed to the stochasticity of calibration parameters and inputs. Probabilistic models of these parameters used in stochastic simulations provide an enhanced framework for the evaluation of proposed improvement strategies.

The cell transmission model (CTM) is a first-order macroscopic simulation model suitable for simulating freeway traffic (*1*). Construction of the freeway model and its calibration are specified by well-defined steps (*2*). The freeway is modeled as a set of sequential segments, and on- and off-ramps are represented by input and output links. A fundamental diagram (the empirical curve relating observed densities to observed flows) is used to specify the parameters of each segment. Density and flow measurements obtained from loop detectors along the freeways are used to calibrate these empirical curves. Input flows into the on-ramps and split ratios (routing parameters) for the off-ramps complete the model specification. Ramp flows, if missing, can be obtained by using a model-based imputation technique. The Performance Measurement System (PeMS), an online repository of traffic data that also contains links to related publications, provides a rich archive of loop detector data for freeways in California (*3*).

Most stochastic simulations that model traffic flow are based on microsimulation models, and their stochastic nature is based on the randomness of demand and driver behavior. In comparison, stochasticity must be introduced in the demands and model parameters (e.g., fundamental diagram parameters) in macroscopic models. The study presented here focuses on the stochastic nature of fundamental diagram parameters with an emphasis on the capacity parameter of freeway sections, which are essential inputs to the CTM and must be calibrated on the basis of available observations. As defined in the *Highway Capacity Manual*, the theoretical capacity (design capacity) of a freeway section is the maximum flow that can possibly traverse the cross-section of that section in a certain period (*4*). However, this theoretical capacity can be reached in practice only when the section is in bottleneck flow conditions (i.e., no active downstream effects on the analyzed section). In contrast, the operational capacity of a section, defined as maximum flow across the section observed over a chosen period, can be estimated for all sections (*5*). This capacity value can be influenced by both queues forming upstream of the section (*6*) and downstream congestion spilling into the section (*7–9*). Moreover, operational capacity can be affected by external conditions such as weather, driver behavior, and so on. All of these factors lead to a high variability in operational capacity over different days (*10*).

In some studies, capacity has been modeled as a random variable; capacity is replaced by the notion of breakdown flow, which is defined as the flow across a section in free flow just before it switches to congested-flow conditions. Several studies model the capacity as a continuous random variable (*11–14*). Of these, Ozbay and Ozguven (*14*) perform a Bayesian nonparametric estimation, Brilon et al. (*12*) suggest a Weibull distribution, and Polus and Pollatschek (*13*)

A. Muralidharan and G. Dervisoglu, 2167C Etcheverry Hall, and R. Horowitz, 5138 Etcheverry Hall, Department of Mechanical Engineering, University of California, Berkeley, CA 94720. Corresponding author: A. Muralidharan, ajith@berkeley.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2249, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 78–85.
DOI: 10.3141/2249-10

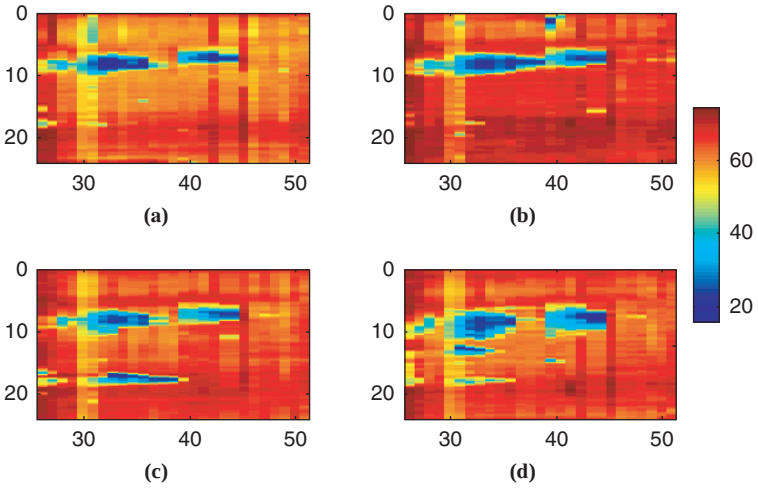

FIGURE 1 Speed contours (mph) observed on (*a*) May 11, (*b*) May 12, (*c*) May 27, and (*d*) July 23, 2009 [*x*-axis = milepost; *y*-axis = time (h)].

use a shifted-gamma distribution to fit capacity data gathered on various days and locations. Wu et al. model the capacity as a normally distributed random variable and use this probabilistic capacity to formulate a chance-constrained ramp metering strategy (*15*).

However, the notion of capacity defined in literature cannot be adapted for use in macroscopic simulation models such as the CTM. Capacity is one of the fundamental diagram parameters calibrated from loop detector measurements over a chosen time horizon (*10*). In this paper, the stochastic variations of capacities (calibrated using a fundamental diagram) are modeled across days observed in almost 1 year of data. The joint probability model (modeled with a probabilistic graphical model approach) estimated from the observed data will be directly applicable for stochastic simulations with the CTM. To the authors' best knowledge, probabilistic models of this notion of capacity are not available in literature.

The paper is organized as follows. First, the diverse nature of congestion patterns and freeway performance observed in practice is summarized, highlighting the observed variations in fundamental diagram parameters calibrated from data and establishing the need to model joint capacity distributions. Next, a probabilistic graphical model approach to represent the joint probability distributions is described. Then, the effect of parameter variations is demonstrated by detailed studies with stochastic simulations. Finally, conclusions of the study are presented.

STOCHASTIC NATURE OF TRAFFIC

Traffic exhibits day-to-day variations that result in different observed congestion patterns, performance measures (VHT, VMT, and so on), and observed flow–density relationships. These variations can be attributed to the stochastic nature of demand as well as other parameters such as section capacity. In this paper, a 25-mi section (Milepost 51 to Milepost 26) of the I-210 freeway westbound in Los Angeles, California, is used for demonstration. Figure 1 illustrates the speed contours (with traffic flowing from right to left) observed for 4 typical days; the extent as well as the time of congestion varies. Figure 2 is a scatter plot of VHT versus VMT that highlights significant variations in freeway traffic performance. In particular, VHT (which measures amount of time spent by all vehicles on the freeway) varies significantly for similar levels of VMT (which measures the usage level of the freeway facility).

Fundamental diagrams, which capture the flow–density relationship, also exhibit significant variations. The commonly used triangular fundamental diagram (Figure 3) is determined by a bimodal constrained least-squares algorithm that separates the data into congested and free-flow modes by speed (<60 mph is deemed congested), sets the maximum observed flow value as the apex of the diagram, and performs constrained least-squares fits on the free-flow and congested parts of the data to estimate free-flow speed and congestion wave speed. Dervisoglu et al. present a similar calibration scheme based on data from multiple days, applying an approximate quantile regression scheme for the estimation of congestion wave speed (rather than the direct least-squares fit used in this paper) (*10*).

Figure 4 is a box plot of capacities observed during 1 year, from April 2009 to April 2010. The vehicle detector stations corresponding to some of the bottleneck locations are marked; locations c and

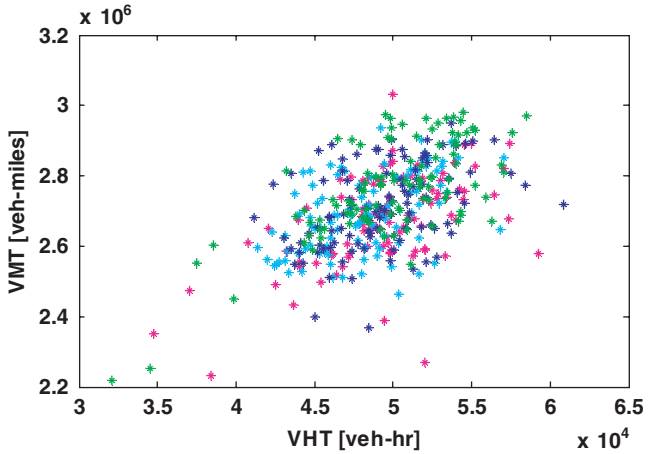

FIGURE 2 VMT versus VHT for I-210W freeway from April 2009 to April 2010 (colors represent different days of week).

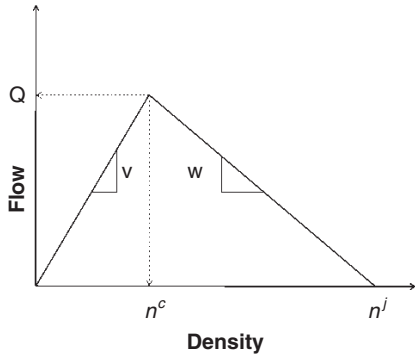

FIGURE 3 Triangular fundamental diagram characterized by free-flow speed (v), congestion wave speed (w), and capacity (Q) (n^c and n^j are critical density and jam density, respectively).

d are frequent bottlenecks, whereas others are less frequent. To exclude the effect of downstream congestion spillback and lack of demand in the capacity determination, data points corresponding to these cases were excluded from the plots. The daily maximum flow varies substantially, and the variation depends on location.

Fundamental diagrams during 40 days of data are plotted in Figure 5*a;* the free-flow speed shows negligible variations. A scatter plot of capacities in adjacent sections indicates a strong correlation between capacities in adjacent sections, and this feature is typical to other section pairs (not necessarily adjacent) on the freeway (Figure 5*b*). Figure 5*c* is a scatter plot of congestion wave speed w and capacity of a freeway section Q. Congestion wave speed (w), restricted to the range of 5 to 20 mph in the calibration procedure, exhibits weak correlation with the capacity of the section as well as parameters from the adjacent section. These observations indicate that the capacity distribution of the entire freeway must be modeled as a multidimensional joint distribution, while independent distributions can be used to model the congestion wave speed distribution.

PROBABILISTIC GRAPHICAL MODEL

A probabilistic graphical model is a graph-based representation of the conditional independence properties in the probability distribution. In this paper, undirected graphical models (also known as Markov random fields) represent the joint distribution of capacities along the freeway. Koller and Friedman provide a detailed view of the theory of probabilistic graphical models (*16*). The candidate graphical models for the joint distribution of section capacities are presented in Figure 6. Each node in the graphical model of freeway capacity distribution corresponds to section capacity. Undirected edges between nodes highlight direct dependence between variables. Variables in the graphs are conditionally independent if no path exists between them after the observed variables and their corresponding edges are deleted.

Models A, B, and C (Figure 6) represent first-, second-, and third-order (spatial) Markov distributions, respectively. Individual section capacities are separated into five equally sized bins, and the resulting discrete probability distribution is modeled. Discretization circumvents the problem of identifying good candidate distributions of continuous capacity for the section as well as joint probabilities, because they do not correspond to well-known candidate distributions. However, discretization results in a loss of accuracy in modeling the distribution.

The data set used for parameter estimation consisted of 216 N-dimensional points, where N is equal to 27 sections. These data points correspond to weekday capacities and exclude days when data were not available. A section capacity is considered unobserved or missing during a particular day when the detector malfunctions or the section does not get congested. It also is classified as missing if maximum section flow is observed when the downstream is congested in an effort to exclude the effect of downstream spillback on the capacity estimation. However, missing capacity is restricted to be more

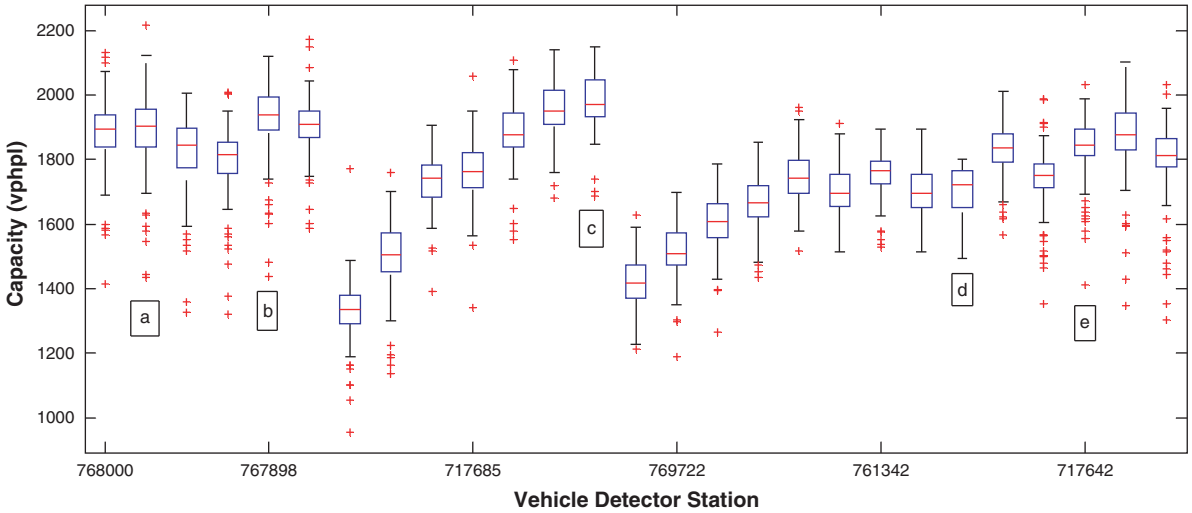

FIGURE 4 Original box plot showing capacity [in vehicles/hour/lane (vphpl)] variation in successive freeway sections, with labels indicating vehicle detector stations (VDSs) corresponding to bottleneck locations: (*a*) 767986 (Milepost 50.29), (*b*) 767898 (Milepost 47.79), (*c*) 717675 (Milepost 38.21), (*d*) 717663 (Milepost 30.78), and (*e*) 717642 (Milepost 28.03).

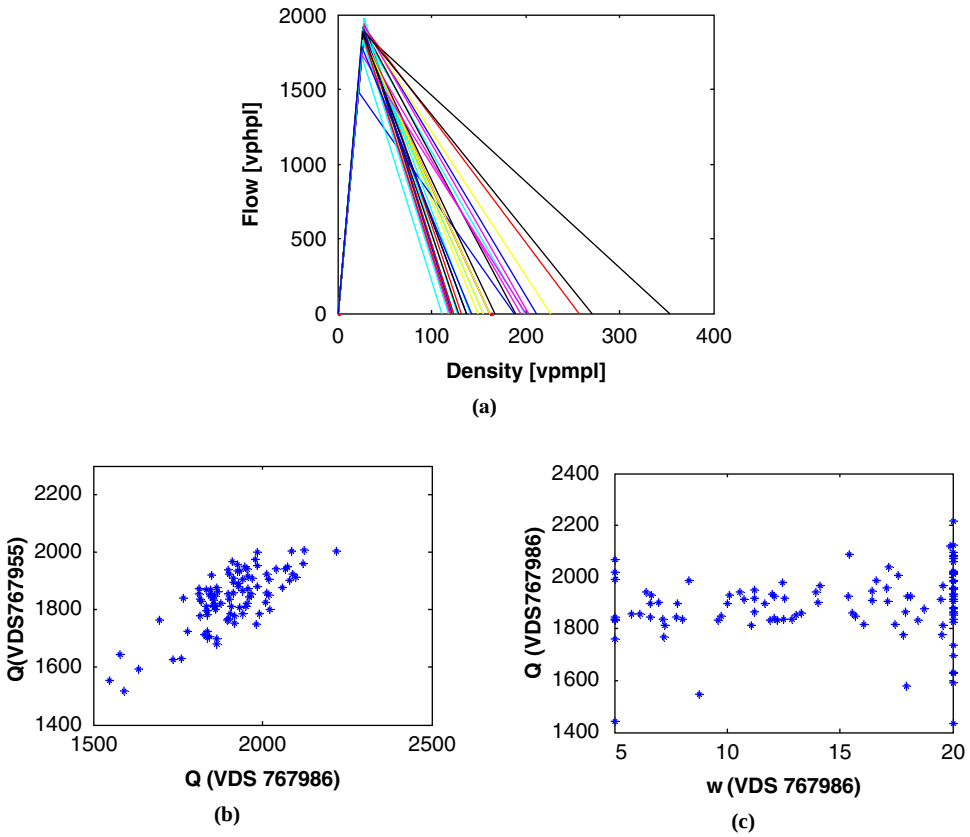

FIGURE 5 (*a*) Calibrated fundamental diagrams across different days at VDS 717637 [where density is measured in vehicles/mile/lane (vpmpl)], (*b*) scatter plot of capacities in adjacent sections, and (*c*) scatter plot of capacity and congestion wave speed in a freeway section.

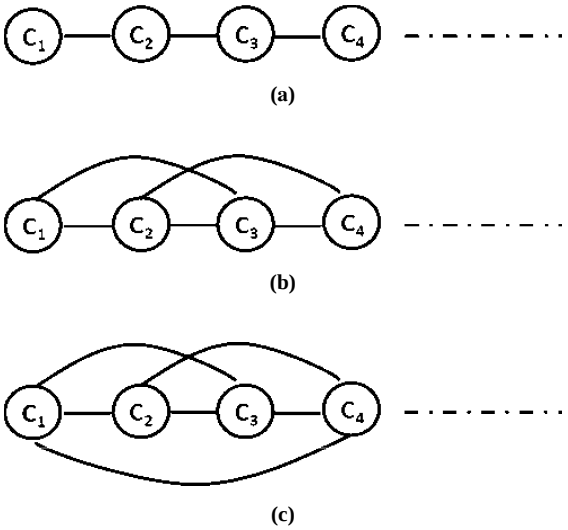

FIGURE 6 Candidate graphical models for joint probability distribution: (*a*) Model A, (*b*) Model B, and (*c*) Model C.

than the maximum observed flow for the day. Most data points contained multiple missing observations for section capacity. To eliminate outliers (due to accidents or special unknown events), capacities outside the range of median plus 1.5 times the interquartile range were denoted as missing. The interquartile range corresponds to the distance between the 25th and 75th percentiles of the data.

Let $C = (C_1, C_2, \ldots, C_N)$ represent the random variable corresponding to section capacities, where $C_i \in (1, \ldots, n_b)$ represents the set of plausible values and the number of bins n_b is equal to 5. The probability distribution is factorized as

$$p(C) = \frac{1}{Z} \prod_{(i,j)\in E} \psi_{i,j}(C_i, C_j)$$

where

E = set of edges in the graph,
Z = normalization constant, and
$\psi_{i,j}$ = potential functions.

The general factorization places potential functions on all cliques (maximally connected subgraphs) in the graph, but given the data size and the high rate of missing data, potential functions are placed at the edges to reduce overfitting. For the proposed discrete factorization, each potential function is an $n_b \times n_b$ table.

With complete observations, parameter learning can be carried out with the iterative proportional fitting (IPF) algorithm (*16, 17*);

with missing observations, the expectation–maximization algorithm must be used (*17*). Missing data are assumed to be missing at random. The expectation–maximization algorithm is a two-step hill-climbing algorithm. The E Step can be interpreted as filling the probability of occurrence of different plausible missing values in the observations with the model, and the M Step uses the output of the E Step to learn the new parameters with the IPF algorithm. The IPF algorithm requires empirical marginals to learn the parameters of the distribution, which are provided in the E Step.

Let $m_{i,j}$ be the dimensional table corresponding to the empirical marginals. The learning algorithm is given as follows.

Given data $\{c^p: p \in \{1, \ldots, N_P\}\}$
Initialize all elements of $\psi_{r,s}$ to nonzero initial values.
Repeat.

E STEP
Set all elements of $m_{i,j}$ to 0
for $p = 1$ to N_P

$$m_{i,j}(i_1, j_1) = m_{i,j}(i_1, j_1) + P(C_i = i_1, C_j = j_1 \mid C = c^p, \psi_{r,s})$$
$$\forall (i, j) \in E \text{ and } i_1, j_1 \in (1, \ldots, n_b)$$

$$m_{i,j}(i_1, j_1) = \frac{m_{i,j}(i_1, j_1)}{N_P} \qquad \forall (i, j) \in E \text{ and } i_1, j_1 \in (1, \ldots, n_b)$$

M STEP (IPF)
Initialize $\psi^0_{r,s}$ to $\psi_{r,s}$
Repeat $t = 1, 2, \ldots,$ until sufficient convergence
Choose an edge $(r, s) \in E$
Calculate $m^{t-1}_{r,s}(p, q) = P(C_r = p, C_s = q \mid \psi^{t-1}_{a,b})$
$\forall p, q \in (1, \ldots, n_b)$
if $m_{r,s}(p, q) \neq 0$

$$\psi^t_{r,s}(p, q) = \psi^{t-1}_{r,s}(p, q) \times \frac{m_{r,s}(p, q)}{m^{t-1}_{r,s}(p, q)}$$

otherwise

$$\psi^t_{r,s}(p, q) = 0$$

LOG LIKELIHOOD CALCULATION

Calculate $L = \sum_p \log(P(c^p \mid \psi_{r,s})$
Until L converges.

The algorithm (E Step and M Step) is iterated until L converges. Indices r, s, i, and j indicate the sections considered. The term $P(C_i = i_1, C_j = j_1 \mid C = c^p, \psi_{r,s})$ involves calculation of probabilities given the model parameters from the previous step. In the case that both section capacities are observed as $(C_i = \bar{i}, C_j = \bar{j})$, then the above term is equal to 1 if $i_1 = \bar{i}, j_1 = \bar{j}$ and 0 otherwise. In the case that the section capacities indicated by the indices (one or both) are missing, then the model from the previous M Step is used to calculate these probabilities, given other observed section capacities. When the capacity of any particular section corresponds to downstream congestion spillback or lack of demand, then the missing data are allowed to take on only values above these recorded maximums.

TABLE 1 Predictive Log Likelihood for Model Selection

Model	Median Predictive Log Likelihood
A	–24.12
B	–39.23
C	–60.91
Independent distributions	–24.97

The term $m^{t-1}_{r,s}(p, q)$ corresponds to model marginals. The probabilities mentioned above are easily calculated with the elimination algorithm, which is one of the inference algorithms commonly used in practice.

The expectation–maximization algorithm presented above can be used to learn the parameters for Models A, B, and C and the independent distribution model. Model selection can be performed by using a 10-fold cross-validation. Data are randomly partitioned into 10 subpartitions of equal size, and the model is learned from nine partitions in rotation and evaluated on the other partition. The log likelihood function was chosen for evaluation; results are summarized in Table 1 (higher values indicate better fit). The median predictive log likelihood values for Model A and the model corresponding to independent distributions are similar, but evidence of high correlation between capacities (e.g., Figure 5*b*) suggests that Model A best represents the joint probability distribution of the capacity function. In addition, sampling from independent distributions for performing simulations might trigger "false" bottlenecks not observed in practice.

STOCHASTIC SIMULATIONS

The first-order spatial Markov capacity distribution is learned from the entire capacity data set using the algorithm presented in the previous section. The free-flow speed v is modeled as a deterministic variable because it shows little day-to-day variation. The congestion wave speed w of each section is modeled as an independent variable. The ramp demands and split ratios were obtained for four days: May 11, 12, 26, and 27, 2009. The imputation algorithm described by Muralidharan and Horowitz was used to obtain ramp demands and split ratios for ramps for which data were incorrect or missing (*18*).

Capacity and congestion wave speed were sampled from their respective distributions and used for stochastic simulations. The sample from the capacity distribution corresponds to the bins used. Actual capacities can be obtained by sampling an (assumed) uniform distribution of capacities within the bin capacity range. Because a stochastic model for demands and split ratios was not available, three stochastic simulations were performed with each of the four demand profiles. Each simulation corresponded to a single sample from the capacity and congestion wave speed distributions.

The velocity contours of 12 stochastic simulations are shown in Figure 7 (where rows from top to bottom correspond to demand–split ratio data for May 11, 12, 26, and 27, 2009). These contour plots highlight the fact that capacity variations can lead to appreciable differences in congestion patterns observed on the freeways. The velocity

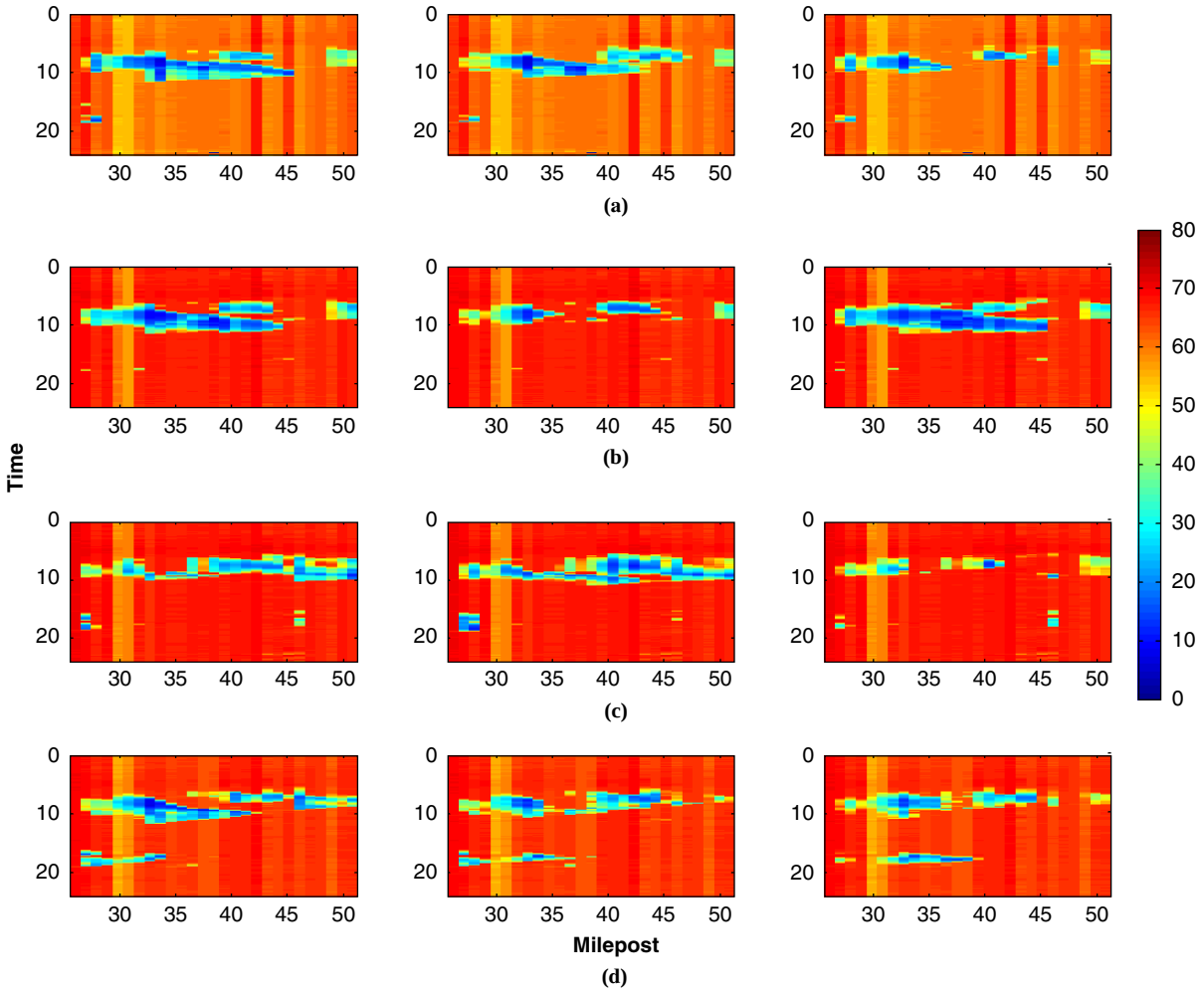

FIGURE 7 Velocity contours (in mph) from stochastic simulations for (*a*) May 11, (*b*) May 12, (*c*) May 26, and (*d*) May 27, 2009.

contours from April 2009 to April 2010 (excluding weekends and holidays) in PeMS indicate that some simulations can replicate observed congestion patterns, whereas others lead to congestion patterns not observed in practice. This finding is expected because other stochastic parameters (e.g., demand) are not modeled here. Notably, the bottlenecks triggered in the simulations corresponded to the common recurring bottlenecks observed in the measured velocity contours, but some were less common.

Figure 8 shows the variation of VMT versus VHT for stochastic simulations with capacity and congestion wave speed samples. These scatter plots indicate that stochastic variations of fundamental diagram parameters can have a significant effect on freeway performance. The plot consists of disjointed groups of points because the model did not account for the stochasticity of demand data.

The VHT–VMT scatter plots are shown in Figure 9; demand and split ratio data from May 11, 2009, were used for all three cases. The plots indicate that congestion wave speed variations have little effect on freeway performance compared with capacity variations. Also, capacity variations produce large variations in vehicle hours spent on the freeway (VHT) for similar numbers of vehicles serviced (VMT).

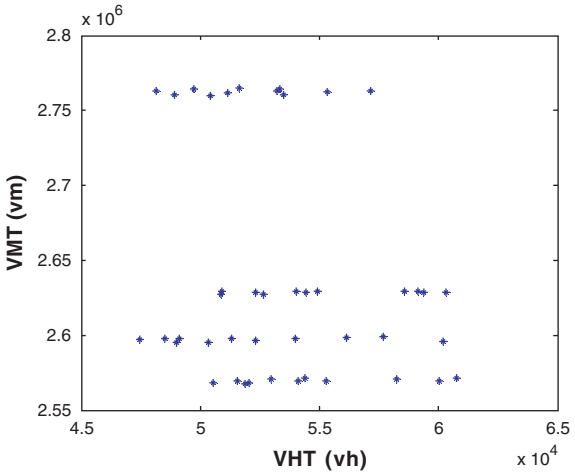

FIGURE 8 VHT–VMT scatter plot generated from stochastic simulations.

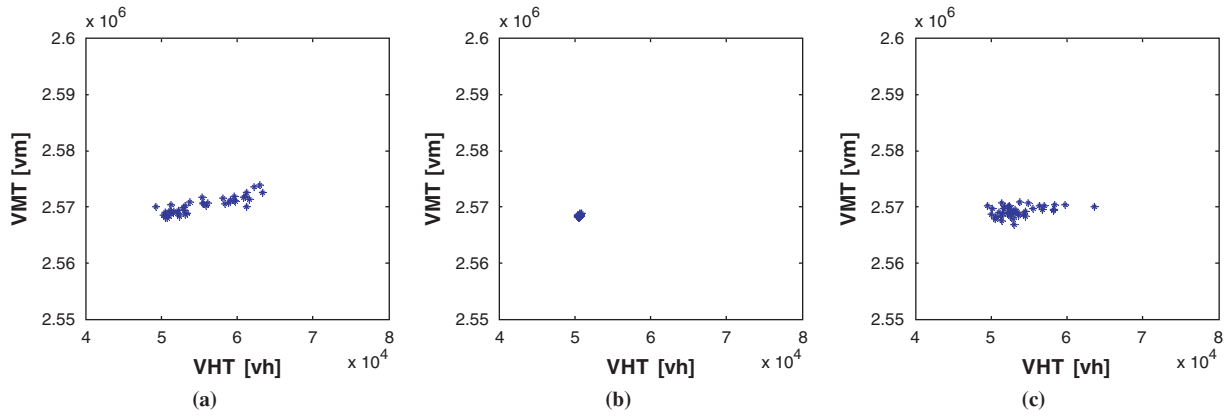

FIGURE 9 VHT–VMT scatter plots, January 23, 2008, with (*a*) capacity variation only, (*b*) congestion wave speed variation only, and (*c*) capacity and congestion wave speed variations.

CONCLUSIONS

The objective of this paper was to take a first step toward modeling the stochasticity of fundamental diagram parameters, with a view toward integrating them into simulation tools. A probabilistic graphical model was presented to represent the joint probability distribution of section capacities located along a freeway. Significant correlations in capacities observed in different sections along the freeway were highlighted. First-, second-, and third-order spatial Markov models were chosen as candidate distributions, and parameters of the distribution were learned with the expectation–maximization algorithm because the data were missing a significant number of observations. Cross-validation was used to conclude that a first-order spatial Markov model was the most appropriate model among the selection pool.

Congestion wave speed, which is another parameter in the fundamental diagram, exhibited independent variations. The congestion wave speed variations had a relatively insignificant effect on freeway performance compared with variations in capacity because although capacity directly affects the service rate of the freeway, the congestion wave speed has only an indirect effect. For example, consider a region of freeway upstream of a bottleneck. If congestion hits a section wave a little later (because of different congestion wave speed), the difference in service rate can be approximately attributed to the additional vehicles that have left the freeway through the upstream off-ramps because of free-flow conditions that existed during the extra time the section was in free flow. This effect is relatively insignificant compared with the effect of capacity on the service rate, given the range of capacity variations observed.

In this paper, maximum flow data points were excluded in sections when they were observed as a result of congestion spillback or obvious lack of demand (i.e., the section did not get congested during the day) because the CTM automatically models these situations. If the stochastic model is used for short-term prediction with models that do not capture these effects, the user may include these data points in the model learning procedure.

A binning and discretization approach also was used to model the distributions because candidate distributions did not conform to known probability distributions. However, the graphical model approach and the method applied here can be adapted even if one chooses to fit a continuous candidate function.

Finally, the framework used for introducing stochasticity in the triangular fundamental diagrams can be applied when other types of fundamental diagrams are considered. These models are not restricted to use in stochastic simulations. In the event that section capacities are missing, the capacity distribution model can estimate the most probable capacity. It can be used to completely specify deterministic models, replicating traffic measurements during a chosen day, even when some section capacities are not observed. These models also can be used for short-term prediction and ramp metering.

ACKNOWLEDGMENTS

This work was supported through a grant from the National Science Foundation. The authors thank Pravin Varaiya for ideas and discussion related to this paper and the Tools for Operations Planning group for their concerted efforts in developing the ideas and software tools that made the study possible. The authors also are grateful for the insightful comments and suggestions of the anonymous reviewers.

REFERENCES

1. Daganzo, C. F. The Cell Transmission Model: A Dynamic Representation of Highway Traffic Consistent with the Hydrodynamic Theory. *Transportation Research Part B*, Vol. 28, No. 4, 1994, pp. 269–287.
2. Muralidharan, A., G. Dervisoglu, and R. Horowitz. Freeway Traffic Flow Simulation Using the Link Node Cell Transmission Model. *Proc., American Control Conference*, St. Louis, Mo., 2009.
3. Performance Measurement System (PeMS). University of California, Berkeley. http://pems.eecs.berkeley.edu. Accessed Aug. 1, 2010.
4. *Highway Capacity Manual*. TRB, National Research Council, Washington, D.C., 2000.
5. Minderhoud, M. M., H. Botma, and P. H. L. Bovy. Assessment of Roadway Capacity Estimation Methods. In *Transportation Research Record 1572*, TRB, National Research Council, Washington, D.C., 1997, pp. 59–67.
6. Hall, F. L., and K. Agyemang-Duah. Freeway Capacity Drop and the Definition of Capacity. In *Transportation Research Record 1320*, TRB, National Research Council, Washington, D.C., 1991, pp. 91–98.
7. Banks, J. H. Review of Empirical Research on Congested Freeway Flow. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1802*, Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 225–232.

8. Cassidy, M. J., and R. L. Bertini. Some Traffic Features at Freeway Bottlenecks. *Transportation Research Part B,* Vol. 33, No. 1, 1999, pp. 25–42.
9. Chung, K., J. Rudjanakanoknad, and M. J. Cassidy. Relation Between Traffic Density and Capacity Drop at Three Freeway Bottlenecks. *Transportation Research Part B,* Vol. 41, No. 1, 2007, pp. 82–95.
10. Dervisoglu, G., G. Gomes, J. Kwon, R. Horowitz, and P. P. Varaiya. Automatic Calibration of the Fundamental Diagram and Empirical Observations on Capacity. Presented at 88th Annual Meeting of the Transportation Research Board, Washington, D.C., 2009.
11. Elefteriadou, L., F. Hall, W. Brilon, R. Roess, and M. Romana. Revisiting the Definition and Measurement of Capacity. Presented at 5th International Symposium on Highway Capacity and Quality of Service, Yokohama, Japan, 2006.
12. Brilon, W., J. Geistefeldt, and M. Regler. Reliability of Freeway Traffic Flow: A Stochastic Concept of Capacity. *Proc., 16th Symposium on Transportation and Traffic Theory,* College Park, Md., July 2005, pp. 125–144.
13. Polus, A., and M. A. Pollatschek. Stochastic Nature of Freeway Capacity and Its Estimation. *Canadian Journal of Civil Engineering,* Vol. 29, No. 6, 2002, pp. 842–852.
14. Ozbay, K., and E. E. Ozguven. A Comparative Methodology for Estimating the Capacity of a Freeway Section. *Proc., 2007 IEEE Intelligent Transportation Systems Conference,* Seattle, Wa., Sept. 30–Oct. 3, 2007, pp. 1034–1039.
15. Wu, X., P. Michalopoulos, and H. X. Liu. Stochasticity of Freeway Operational Capacity and Chance-Constrained Ramp Metering. *Transportation Research Part C,* Vol. 18, No. 5, 2010, pp. 741–756.
16. Koller, D., and N. Friedman. *Probabilistic Graphical Models: Principles and Techniques.* MIT Press, Cambridge, Mass., 2009.
17. Schafer, J. L. *Analysis of Incomplete Multivariate Data.* Chapman and Hall, London, 1997.
18. Muralidharan, A., and R. Horowitz. Imputation of Ramp Flow Data for Freeway Traffic Simulation. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2099,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 58–64.

The contents of this paper reflect the views of the authors and not necessarily the official views or policy of the National Science Foundation.

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