***Abstract*— The chance that a freeway will breakdown, transition from a free-flow to a congested state, is normally assumed to increase with an increase in traffic volume V (vehicles per unit time). In this paper, this assumption is challenged. Traffic density K (vehicles per unit length) proves to be a better predictor. Diffusion or stochastic differential equation (SDE) modeling is used to substantiate the claim. SDE modeling is especially useful in explaining the role that traffic noise (volatility) plays in breakdown. The SDE models take advantage of the unique properties of the geometric Brownian motion (*gBM*) and Ornstein-Uhlenbeck (*OU*) model structures. The breakdown probability model of $\pi(K)$ and delay models provide accurate forecasts.**

falls below the threshold speed of $u^* = 50$ mph and when the delay associated with the speed lose persists for one or more T time-intervals. Otherwise, the freeway is in a free-flow state. The freeway state for day d and time-of-day t is classified using the following test:

$$X_{dt} = \begin{cases} 0 & \text{if } u_{dt} > 50 \text{ mph}, \\ 1 & \text{if } u_{dt} \leq 50 \text{ mph}. \end{cases} \quad (1)$$

This information is stored in 109×96 state space matrices $\mathbf{X_{03}}$ and $\mathbf{X_{02}}$ for the bottleneck q03 and upstream q02 locations, respectively. When a string of contiguous congestion events is formed, the first traffic density and volume observations in the string are designated as *breakdown events* at q03, (k_0, q_0). Since no breakdown events are observed at q02 and all speeds less that u^* = 50 mph are due to spillback from q03, k_0 and q_0 at q02 are called *spillback events*, (k_0, q_0). The time-step where $u_t < u^*$ is observed at q03 and q02 is denoted as t_0 and stored in $\mathbf{T_{03}}$ and $\mathbf{T_{02}}$ vectors, respectively.

The easiest way to analyze spillback is with time-series plots and scatter plots. Panel (a) of Figures 1 and 2 show the breakdowns at the bottleneck can have very different upstream impacts. Consider them separately.

No spillback is observed for day d = 83. The speed at q02 does not fall below u^* = 50 mph at any time during the day. Interestingly, three breakdown events observed at the bottleneck q03 on this day.

On day d = 40, a single breakdown event at the bottleneck leads to spillback causing two separate speed reduction events of $u < 50$ mph at q02. Since the time t_0 for q03 is less than either of the times t_0 for q02 suggests that a queue formed and eventually grew in length exceeding 1.5 miles.

The scatter plots of panels (b) and (c) of Figures 1 and 2 show straight line $q-k$ relationships for observations in the free-flow state (open circles) and random patterns for those in the congestion state (solid circles). To be clear, the large open circles surrounding the solid circles in panels (b) are *breakdown events*. The large open circles surrounding the solid circles shown in panel (c) of Figure 2 are *spillback events*.

The easiest way to analyze spillback is with time-series plots and scatter plots. Panel (a) of Figures 1 and 2 show the breakdowns at the bottleneck can have very different upstream impacts. Consider them separately.

No spillback is observed for day d = 83. The speed at q02 does not fall below u^* = 50 mph at any time during the day. Interestingly, three breakdown events observed at the bottleneck q03 on this day.

On day d = 40, a single breakdown event at the bottleneck leads to spillback causing two separate speed reduction events of $u < 50$ mph at q02. Since the time t_0 for q03 is less than either of the times t_0 for q02 suggests that a queue formed and eventually grew in length exceeding 1.5 miles.

The scatter plots of panels (b) and (c) of Figures 1 and 2 show straight line $q-k$ relationships for observations in the free-flow state (open circles) and random patterns for those in the congestion state (solid circles). To be clear, the large

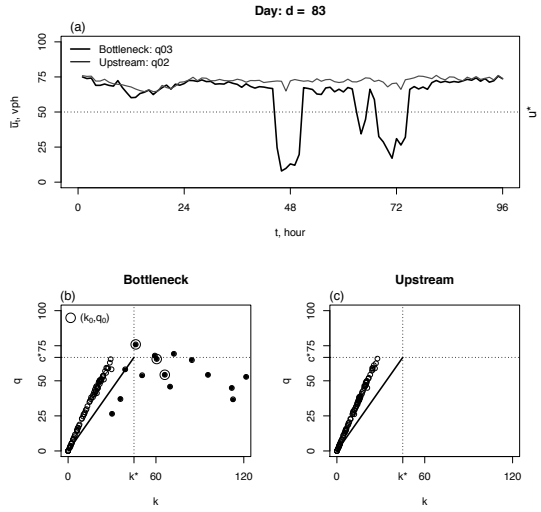

Fig. 1. Three breakdown events occur at the bottleneck. No spillback events are observed upstream.

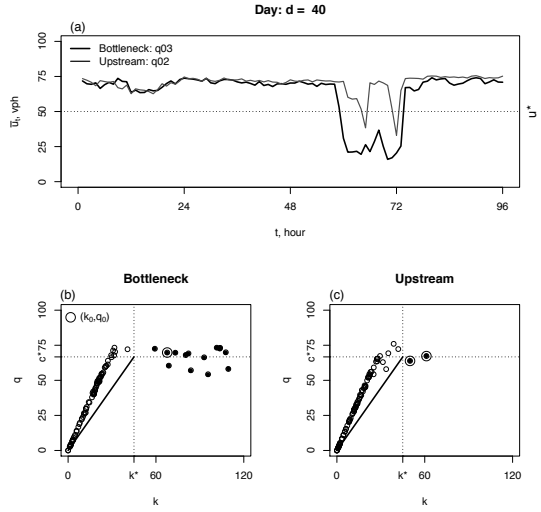

Fig. 2. One breakdown event occurs at the bottleneck. Two spillback events are observed upstream.

C. Freeway Capacity and Traffic Noise

Highway Capacity Manual (HCM) guidelines are used to construct the straight lines between $(0,0)$ and $(k^*, c^*) = (45, 67)$ shown in the (b) and (c) panels. The HCM defines roadway capacity with a density measure of k^* = 45 vehicles per mile (vpm) and specifies freeway capacity in vehicles per hour per lane. In this paper, the bottleneck location under study has two-lanes, thus the traffic volume capacity assumed to be twice the value given in the HCM of 2000 vehicles per hour for a single lane. A value of c^* = 67 vehicles per minute (vpm) is shown.

TABLE I

BREAKDOWN AND SPILLBACK COUNTS

	Bottleneck: q03	Upstream: q02
Free-flow	2	62
Breakdown	107	-
Spillback	-	47

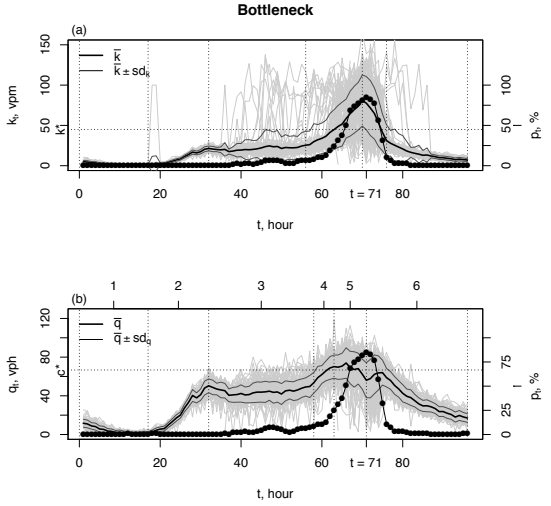

Fig. 3. Traffic flow data and congested state proportions p_t, shown with solid circles, at location q03. The light gray lines are traces for the individual traffic records.

All the solid circles in panels (b) of Figures 1 and 2 lie to the right of the vertical k^* line. These k values show a great deal of volatility, ranging from 67 to 120 vpm. In contrast, these same solid circles measured in q are clustered around the horizontal q^* line, showing relatively little volatility in comparison to the k measures. Inspecting the entire data set for all days reveals the volatility in K to be greater than V.

Delay time can be estimated by counting the number of solid circles and multiplying by T = 15 minutes. Bottleneck delay at q03 for days d = 83 and 40 is approximately three hours for each day. Upstream delay at q02 for day d = 40 is 30 minutes.

D. Breakdown and Spillback Counts and Proportions

Congestion at the study site is severe. The information given in Table I shows the bottleneck at location q03 was congested on 107 of 109 days. A queue that extends from locations q03 to q02 is common. These spillback events are estimated to occur at around 44% (p = 47/107) of the time.

All breakdown counts at q02 given in the table are initiated by spillback. For this reason, the discussion for the remainder of this section and the **Model** section focus on modeling the bottleneck at location q03. spillback will be discussed in the **Model Forecast** section.

E. Time-series Analyses

Proposition: As traffic demand increases, the chance of observing a congestion state is assumed to increase. That is, the proportion of congestion events is assumed to be in *direct* proportion to the traffic demand, or $p \propto q$ and $p \propto k$. Note that traffic demand is usually defined as a flow q. Since $k = q/u$ and k is a function of q, traffic density is also a measure of demand using a different scale. Inspection of the time-series plots of Figure 3 are used to test the validity of the claim that $p \propto q$ and $p \propto k$ are valid.

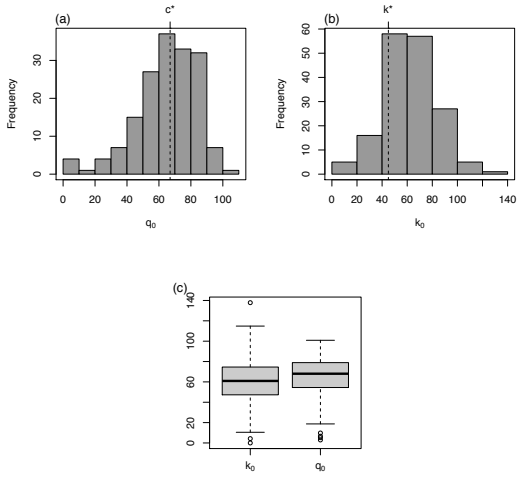

Fig. 4. Distributions of Q_0 and K_0 at location q03.

Before moving forward, casual inspection of the figure give further evidence that the site suffers from severe congestion. Panels (a) and (b) also show the process is extremely noisy or volatile. The volatilities of V and K are shown with the light gray lines and the 68% confidence intervals are shown as $\bar{k}_t \pm sd_k$ and $\bar{q}_t \pm sd_q$.

The percentage of time the freeway is in a congested state at time-step t is denoted as p_t. A p_t points is calculated as the ratio m_t/n where m_t is the breakdown count for time-step t and n = 109. The maximum value of p, $max(p) = p_{t=71}$ = 85%, occurs at time t_{max} = 71 or at 5:45 pm.

Since traffic flow q is a usually considered to be a good predictor of congestion, consider it first by inspecting panel (b). If the $p \propto q$ is in *direct* proportion, then the time when q reaches its maximum should to be equal to t_{max}. The two times do not match. Furthermore, the range of $q > c^*$ and range of peak p_t values do not match. A good match occurs only at low flows where $p_t < 50\%$.

In contrast, a similar inspection of panel (a) show $p \propto k$ to be direct proportion for all t. Unlike q, the time when $k = \bar{k}$ reaches its maximum value is t_{max} = 71. For the range between $65 < t < 75$, the $k = \bar{k}$ and p_t curves overlap almost perfectly.

Figure 4 gives another perspective of traffic breakdown. Here, all *breakdown events* (k_0, q_0), discussed above, are investigated for all $n = 109$ days. In other words, all the large open circles surrounding the solid circles in Figures 1(b) and 2(b) are included in the histograms and box plots. The plots show the Q_0 and K_0 data are volatile. The most striking feature of panels (a) and (b), the histograms, are the

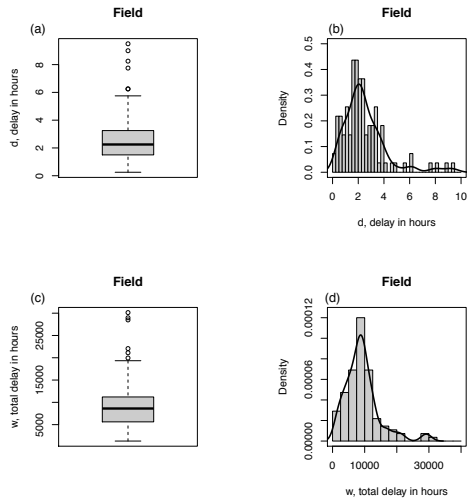

Fig. 5. Distributions of delay at location q03.

positions of c^* and k^*. In the case of Q_0, the q_0 values tend to evenly distribution around c^*. In the case of K_0, the majority of k_0 values lie to right of k^*.

The histograms prove useful in checking the validity of $p \propto q$ and $p \propto k$. Breakdown is assumed to occur when the demand exceeds capacity, $q > c^*$ and $k > k^*$. Applying this definition to histogram of Figure 4(a) for Q_0 will give a probability of breakdown approximately equal to 50%, the area of the bins to right of c^*. Obviously, this is a gross underestimate, particularly when $p_{t=71} = 85\%$ is known. The area of the bins for K_0 to the right of k^* is a good estimate of $p_{t=71} = 85\%$. Based of the two tests, "match tests" if you will, offers evidence that K is a better predictor of traffic breakdown than Q.

F. Delay

Delay d is defined as the length of time the freeway is in a congested state, $u < u^* = 50$ mph. The delay in hours and vehicle-hours for day d (subscript) are respectively calculated as:

$$\begin{aligned} d_d &= 0.25 \sum_{t=1}^{96} t_{dt}, && \text{from } \mathbf{T_{03}} \\ w_d &= 0.25 \sum_{t=1}^{96} q_{dt} t_{dt}, && \text{from } \mathbf{Q_{03}} \text{ and } \mathbf{T_{03}} \end{aligned} \quad (2)$$

Figure 5(a) and (b) show the distribution of d using box plots and histograms. These graphs show the delay is expected to be better than two hours per day. Again, the congestion at the site is severe.

The box plots of panels (a) and (c) and the histograms of panels (b) and (d) will be used in the "match tests" for checking the accuracy of the $\pi(V)$ and $\pi(K)$ models.

III. Model Forecastings

The overall approach used in this section for analyzing congestion and delay is similar to the presentation in the **Explanatory Data Analysis** section. Unlike above, computer simulation is used to forecast $\hat{\pi}(K)$ and $\hat{\pi}(Q)$ with the SDE models of K and V, respectively. To gain an appreciation of the importance of the V and K models, breakdown forecasting is introduced first. After the discussion of free-flow and congestion state determination is completed, model assignment, model calibration and delay forecasting are discussed. The section concludes with a goodness-of-testing, i.e., "match testing."

A. Breakdown Forecasting

The freeway state for day d and time-of-day t is identified as:

$$\hat{X}_{dt} = \begin{cases} 0 & \text{if } \hat{k}_{dt} > k^* = 45 \text{ vpm}, \\ 1 & \text{if } \hat{k}_{dt} \le k^* = 45 \text{ vpm}. \end{cases} \quad (3)$$

for the $\hat{\pi}(K)$ forecast model, and

$$\hat{X}_{dt} = \begin{cases} 0 & \text{if } \hat{q}_{dt} > c^* = 67 \text{ vpm}, \\ 1 & \text{if } \hat{q}_{dt} \le c^* = 67 \text{ vpm}. \end{cases} \quad (4)$$

for the $\hat{\pi}(V)$ forecast model. The breakdown probability forecasts for day d and time-of-day t are denoted as $\hat{\pi}(K_{dt}) = P(\hat{X}_{dt} = 1)$ and $\hat{\pi}(V_{dt}) = P(\hat{X}_{dt} = 1)$ and stored in matrices. For example, $\hat{\mathbf{X}}_{\mathbf{03}}$ will be used for $\hat{\pi}(K)$ forecasts and analyzed as described above.

B. Model Assignment

As can be gleaned from (3) and (4), the ability of forecast breakdown and congestion accurately depends on the K and V models in providing reliable $\hat{K}_{dt} = k_{dt}$ and $\hat{V}_{dt} = q_{dt}$ forecasts. To achieve this end, two parametric families of stochastic models are assigned. They are described in (5) and (6) using a generic random variable S:
The gBm Model:

$$dS_t = \theta_1 S_t dt + \theta_2 S_t dW_t \quad (5)$$

The OU Model:

$$dS_t = (\theta_1 - \theta_2 S_t) dt + \theta_3 dW_t, \quad (6)$$

Each dS_t model consists of well-defined *drift* and *diffusion* components. The *drift* and *diffusion* components for the *gBm* model are $\theta_1 S_t dt$ and $\theta_2 S_t dW_t$, respectively. For the *OU* model, they are $(\theta_1 - \theta_2 S_t)dt$ and $\theta_3 dW_t$. W represents *Brownian motion* or a *Wiener process*. The model parameters are θ_1, θ_2 and θ_3. Closed form solutions of S_t can be obtained where s_0 specifies the initial condition of S_t at $t = 0$.
The gBm Model:

$$S_t = S_0 e^{\left(\theta_1 - \frac{1}{2}\theta_2^2\right)t + \theta_2 W_t} \quad (7)$$

The OU Model:

$$S_t = \frac{\theta_1}{\theta_2} + \left(s_0 - \frac{\theta_1}{\theta_2}\right) e^{-\theta_2 t} + \theta_3 \int_0^t e^{-\theta_2 (t-u)} dW_u \quad (8)$$

Each model has an unique conditional probability distribution. The *gBm* and *OU* models have log-normal and normal distributions, respectively. For the purposes of model parameter estimation, the mean and variances of these functions prove to be most useful:

TABLE II

K MODEL PARAMETER ESTIMATES

Assignments			Zone Descriptors			Parameter Estimates		
Zone	Model	t range	$\bar{k}_{start}$	$\bar{k}_{end}$	sd_k	θ_1	θ_2	θ_3
1	*OU*	(0, 17)	5	2	0.1	0	0.22	0.065
2	*gBm*	(17, 32)	2	20	0.5	0.61	0.013	-
3	*gBm*	(32, 58)	20	25	12	0.034	0.18	-
4	*gBm*	(58, 70)	25	78	12	0.38	0.088	-
5	*gBm*	(70, 76)	78	30	5	-0.64	0.14	-
6	*OU*	(76, 96)	30	5	0.1	0	0.30	0.077

TABLE III

V MODEL PARAMETER ESTIMATES

Assignments			Zone Descriptors			Parameter Estimates		
Zone	Model	t range	$\bar{q}_{start}$	$\bar{q}_{end}$	sd_k	θ_1	θ_2	θ_3
1	*OU*	(0, 17)	12	2	0.1	0	0.42	0.092
2	*gBm*	(17, 32)	2	48	0.5	0.85	0.001	-
3	*gBm*	(32, 58)	48	50	10	0.006	0.077	-
4	*gBm*	(58, 70)	50	72	10	0.12	0.079	-
5	*gBm*	(70, 76)	72	72	15	0	0.17	-
6	*OU*	(76, 96)	72	12	0.1	0	0.36	0.085

The gBm Model:

$$\begin{aligned} m(t,s_0) &= s_0 e^{\theta_1 t}, & t \geq 0 \\ v(t,s_0) &= s_0^2 e^{2\theta_1 t}\left(e^{\theta_2^2 t}-1\right), & t \geq 0 \end{aligned} \tag{9}$$

The OU Model:

$$\begin{aligned} m(t,s_0) &= \frac{\theta_1}{\theta_2} + \left(s_0 - \frac{\theta_1}{\theta_2}\right)e^{-\theta_2 t}, & t \geq 0 \\ v(t,s_0) &= \frac{\theta_3^2\left(1-e^{-2\theta_2 t}\right)}{2\theta_2}, & t \geq 0 \end{aligned} \tag{10}$$

To describe traffic behavior for a single day or at times-of-day $t = 1,2,\ldots,96$, six individual time-zone S_t models are linked. The zones are identified by zone number $i = 1,2,\ldots,6$ as shown at the top edge of Figure 3(b). To reduce clutter, the zone numbers are not shown in panel (a).

The zone boundaries, which are shown with dotted vertical lines, are subjectively selected using a simple protocol. The slopes of the $\bar{k}$ and $\bar{q}$ lines within each zone decrease or increase at roughly the same rate. Zones 1 and 6 have decreasing $\bar{k}$ and $\bar{q}$ slopes and zones 2, 3 and 4 have increasing slopes. Zone 5, which have a slightly decreasing slope for K and zero slope for V, has relatively large volatilities as exhibited by the $\bar{k} \pm sd_k$ and $\bar{q} \pm sd_q$ boundary lines as compared to the other zones. Thus, zone 5 falls outside the simple protocol.

Next, *gBm* and *OU* models are assigned to each time zone. The *gBm* model is assigned to zones 2 through 5 and the *OU* model is assigned to zones 1 and 6. The θ_1 parameter of (6) is assumed to be equal to 0, the mean and variance of (10) are functions of θ_2 and θ_3 only.

The model assignments are summarized in Tables II and III.

C. Model Parameter Estimation

The next task is to estimate the parameters of the respective models of V and K. This is achieved by obtaining zonal descriptors and then substituting them into the $m(t,s_0)$ and $v(t,s_0)$ formulas and solving for the unknown θ parameters.

Take row 2 of Table II for example. The *start* and *end* points are specified under the t range column. The $\bar{k}_{start}$ = 2 and $\bar{k}_{end}$ = 20 are obtained from $\mathbf{K_{03}}$. Thus, $s_0 = \bar{k}_{start}$ = 2, $m(t,s_0) = \bar{k}_{end}$ = 20 where t = 32 - 17 = 15. Thus, $m(t,s_0) = m(15,2) = 20$. Substitute these values into the $m(t,s_0)$ of (3) and solve for θ_1. In this case, θ_1 = 0.61 as shown. Knowing θ_1 and $v(t,s_0) = sd_k$ = 0.5, it is an easy manner to substitute them into the $v(t,s_0)$ formula and to solve θ_2 = 0.013.

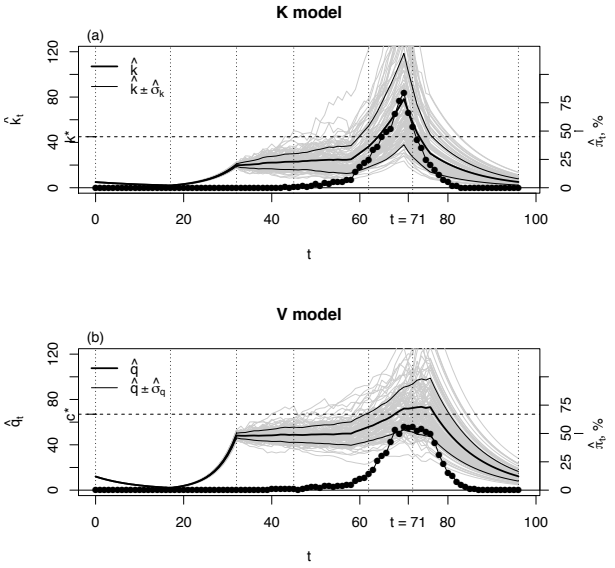

Fig. 6. Traffic demand and congestion probability forecasts at location q03.

It is a well recognized fact that model calibration, especially SDE model calibration, is a challenge [2]. The tricky part is assigning sd_k because it represents the zone, not a *start* and *end* point as in estimating θ_1. It is often difficult to anticipate how a model will respond. Testing, presented in the next section, provides assurances that the model estimation procedure is satisfactory.

D. Forecasting

Now that the model assignment and calibration procedures are completed, model simulation is initiated [3] [5]. Simulation is straight forward. A simulation is initiated by obtaining a $s_0 = q_0$ value for a V simulation and a $s_0 = k_0$ value for a K simulation. For V, a draw is made from $N(\bar{q}_1, sd_{q1})$ where the mean and standard deviation are calculated with the column of $\mathbf{Q_{03}}$. For K, a draw is made from $N(\bar{k}_1, sd_{k1})$ using $\mathbf{K_{03}}$.

To simplify the discussion, the generic S_t notation is used. Thus, a simulation starts with s_0 and forecasts for zone 1 are made with a *OU* model. The last value of the simulation for zone 1 is used as the starting value for zone 2. The procedure of linking one zone to the next one is used for all six zones. Simulation output or forecasts for one-hundred simulation runs are stored in $\hat{\mathbf{Q}}_{\mathbf{03}}$, $\hat{\mathbf{K}}_{\mathbf{03}}$ and $\hat{\mathbf{T}}_{\mathbf{03}}$ and shown in Figure 6.

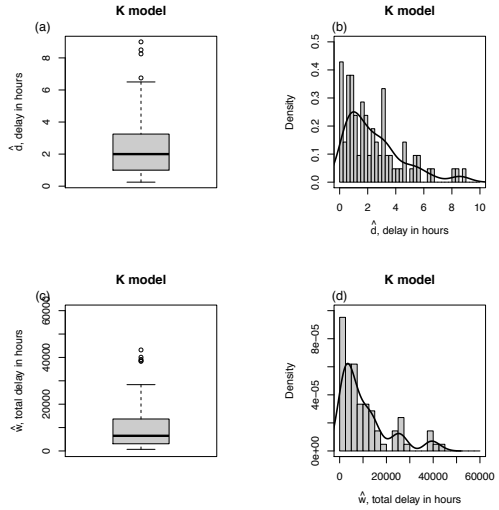

Fig. 7. Traffic delay forecasts at location q03 using the traffic density model.

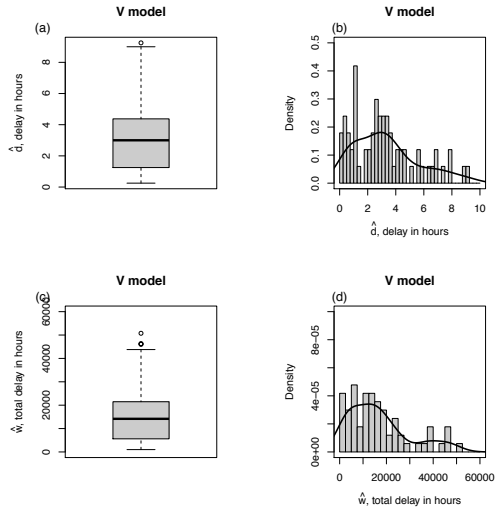

Fig. 8. Traffic delay forecasts at location q03 using traffic volume model.

E. Goodness-of-Fit

Goodness-of-fit tests for ordinary regression and categorical models offer evidence as to the overall quality of the model fit. More work is needed in this area for generalized models such as presented here [1]. A series of "match tests," comparing field and model forecasts, are used for model diagnosis. The first test compares deals with an evaluation the time-series models, followed by the evaluation of breakdown probability and delay models. The simulation results of delay for $\hat{V}$ and $\hat{K}$ are shown in Figures 7 and 8. These figures are generated by using the same steps and formulas that were used to produce the information described in the **Explanatory Data Analysis** section shown in Figure 5.

Both the $\hat{V}$ and $\hat{K}$ capture the overall trend and volatility displayed in Figure 3. The volatility in some individual k_t field records are much greater the most volatile $\hat{k}_t$ forecasts. Regardless, the $k_t \pm sd_k$ and $\hat{k}_t \pm \hat{\sigma}_k$ match well. The same can be said for the $q_t \pm sd_q$ and $\hat{q}_t \pm \hat{\sigma}_q$ match. These results give assurance that the model parameter estimates are satisfactory.

Inspection of the magnitudes of p_t and $\hat{\pi}_t$ of Figure 6(a) match well for the K model. For example, $max(p_t) = p_{t=71} = max(\hat{\pi}_t) = 85\%$. Secondly, the box plots and histograms of Figure 5 for delay derived from field measurements and Figure 7 derived from the K model show a good match.

The same cannot be said for V model for the forecasts given in Figures 6(b) and 8. For one, $max(\hat{\pi}_t) = 52\%$, severely underestimates $p_{t=71} = 85\%$. Figures 5 and 8 are poor matches. At the same time, panels (b) of Figures 3 and 7 show good matches of q_t and $\hat{q}_t$, suggesting the traffic breakdown criterion given in (4) is incorrect, especially in the rush hour period of the day when $p_t \geq 50\%$.

The "match tests" show favorable correspondence between k_t and $\hat{k}_t$ and between p_t and $\hat{\pi}_t$, shown in the time-series plots of Figures 3(a) and 6(a), and the delay box plots and histograms given in Figures 5 and 7. This is evidence that the traffic breakdown and congestion criterion given in (3) coupled with the $\hat{k}_t$ forecasts of the K model leads to reliable $\hat{\pi}_t$ and delay forecasts for all t.

F. Spillback

Queue formation initiated at the bottleneck q03 location causes serious delay upstream at q02. See Table I. Analyses using the K forecast model shows the the traffic to be volatile, causing a stop and go driving with an average estimated speed of 20 mph. The proportion of the time q02 is in a congested state with speeds less than u^* = 50 mph is forecast to be 25%. The model underestimates the proportion observed in the field, thus suggesting work is needed to improve it. Specifically, obtain reliable estimates of queue length, which is an important factor in traffic management.

IV. Summary

A critical feature of this SDE modeling approach is in explaining how traffic volume and traffic noise *work together* in triggering a breakdown event. It does so with diffusion models consisting of (1) a deterministic component, (2) a stochastic component, and (3) an initial condition. A most important finding in this study is that the $\hat{\pi}_t = P(K_t \geq k^*)$ model can provide accurate forecasts, a claim supported by "match" testing. The finding that K_t is a better predictor variable than V_t suggests vehicle operators are more sensitive to *space headway* than to *time headway* when faced with crowded conditions.

References

[1] Alan Agresti. *Categorical Data Analysis*. Wiley-Interscience, New York, first edition, 1990.

[2] Stefano Iacus. *Simulation and Inference for Stochastic Differential Equations: With R Examples*. Springer-Verlag, 2008.

[3] Stefano Maria Iacus. *sde: Simulation and Inference for Stochastic Differential Equations*, 2016. R package version 2.0.15.

[4] Paul J. Ossenbruggen and Eric M. Laflamme. Explaining freeway breakdown with a geometric brownian motion model. *Journal of Transportation Engineering*, In press, 2017.

[5] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.