Structure and Infrastructure Engineering, Vol. 00, No. 00, Month 200x, 1–26

RESEARCH ARTICLE

Estimating post-disaster traffic conditions using real-time data streams

Reece P. Otsuka[a], Daniel B. Work[a] and Junho Song[b*]

[a]*Department of Civil and Environmental Engineering, University of Illinois, Urbana-Champaign, IL 61801, USA*; [b]*Department of Civil and Environmental Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 151-742, Republic of Korea*

(*Received 00 Month 200x; final version received 00 Month 200x*)

This article presents a technique to determine the effects of an earthquake on road traffic conditions by linking seismic hazard and bridge fragility models with a traffic model and traffic sensor data. Using the earthquake characteristics as an input to the traffic model, the traffic conditions are sequentially estimated given traffic sensor measurements using an ensemble Kalman filter. The proposed algorithm is tested on a numerical experiment on Interstate 155 West near the New Madrid Seismic Zone in New Madrid, Missouri. The results show that the proposed technique improves the post-disaster traffic estimate. The supporting source code and data are available for download at `https://github.com/rotsuka/postDisaster_EnKF`.

Keywords: Traffic estimation; ensemble Kalman filtering; earthquake; post-disaster; data assimilation

1. Introduction

1.1 *Post-disaster traffic estimation as a sequential state estimation problem*

Earthquakes are devastating events that can have a number of social, economic, and structural consequences. One of the structural ramifications that can result from the propagation of earthquake shock waves is damage to transportation infrastructure. As observed in recent earthquake events, bridges are especially susceptible to damage (Akiyama et al. 2013, Basöz et al. 1999, Kosa 2012, Schanack et al. 2012). Structural damage to bridges can cause disconnections in the transportation network and reductions in traffic flow capacity, which in turn slows down post-hazard response efforts. If accurate traf-

*Corresponding author. Email: junhosong@snu.ac.kr

ISSN: 1573-2479 print/ISSN 1744-8980 online
© 200x Taylor & Francis
DOI: 10.1080/1573247YYxxxxxxxx
http://www.informaworld.com

fic estimates can be calculated immediately following an earthquake, it could be very helpful in expediting post-hazard rescue and response efforts. The problem addressed in this work is to estimate post-disaster traffic conditions using measurements from traffic sensors and information on the earthquake characteristics and infrastructure fragility. In this work, the magnitude of the earthquake and its proximity to bridge infrastructure results in changes to the traffic capacity of each bridge, affecting the traffic conditions of the network.

Typically, the traffic estimation problem is posed as a sequential state estimation problem. The traffic evolution equations are constructed from a macroscopic traffic model, denoted by $\boldsymbol{f}$. The function $\boldsymbol{f}$ describes how the *traffic state* $\boldsymbol{x}$, i.e. a vector of traffic densities (e.g. veh/km) at various locations along the roadway, evolves from time $n-1$ to time n. The evolution and observation equations (e.g., Simon 2006) are given respectively as

$$\boldsymbol{x}_n = \boldsymbol{f}\left(\boldsymbol{x}_{n-1}, \boldsymbol{\theta}, \boldsymbol{w}_n\right) \tag{1}$$

$$\boldsymbol{y}_n = \boldsymbol{h}_n\left(\boldsymbol{x}_n\right) + \boldsymbol{v}_n, \tag{2}$$

where $\boldsymbol{\theta}$ is a time-invariant vector which contains parameters of the road such as the maximum traffic capacity, the speed limit, and the jam density. The term $\boldsymbol{y}_n$ is a vector of traffic sensor measurements, $\boldsymbol{h}_n$ is an observation operator that relates the system state with the measurements, $\boldsymbol{w}_n$ is the model noise, and $\boldsymbol{v}_n$ is the measurement noise at time n. Given the model evolution (Equation (1)) and observation (Equation (2)), the traffic estimation problem can be posed as estimating the traffic state $\boldsymbol{x}_n$ given measurements $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_n$.

Following an earthquake, a standard traffic estimation algorithm will suffer poor performance because the forward model (Equation (1)) does not explicitly account for changes in the network topology due to the earthquake damage. Additionally, dedicated traffic sensors may be sparsely distributed throughout the network (due to the high installation and maintenance costs), which limits the ability of the measurements to correct the traffic state estimate. Thus, to improve the estimation performance, we augment the model evolution equation by feeding in real-time information about the earthquake at each time step as an input, $\boldsymbol{u}_n$, to the traffic model. As a result, Equation (1) can be rewritten as

$$\boldsymbol{x}_n = \boldsymbol{f}\left(\boldsymbol{x}_{n-1}, \boldsymbol{\theta}, \boldsymbol{u}_n, \boldsymbol{w}_n\right), \tag{3}$$

where $\boldsymbol{u}_n$ is the *earthquake input* at time step n defined as

$$\boldsymbol{u}_n = \begin{cases} \mathbf{1} & \text{if } n < n_{\text{EQ}} \\ \boldsymbol{\alpha} & \text{if } n \geq n_{\text{EQ}}. \end{cases} \tag{4}$$

The size of the vectors $\boldsymbol{u}_n$ and $\boldsymbol{\alpha}$ is equal to the number of bridges in the network, or more generally the number of elements of the infrastructure that may sustain damage. In Equation (4), n_{EQ} is the time step in which the earthquake occurs and the vector $\boldsymbol{\alpha}$ is

the *traffic capacity factor*, with elements $\alpha_i \in [0, 1]$. The elements in $\boldsymbol{\theta}$ that correspond to the maximum traffic capacity and the jam density are multiplied by the elements in $\boldsymbol{u}_n$ of the corresponding bridge to allow for the fact that the roadway properties may change following an earthquake. Due to the nonlinearity of the traffic model, we propose to use an ensemble Kalman filter (Evensen 1994) to solve the sequential state estimation problem with the earthquake input.

1.2 *Related work on post-disaster transportation networks and traffic state estimation*

Motivated by the number of natural disaster occurrences worldwide, many researchers have focused on the post-disaster resilience of infrastructure systems. Specifically, there have been several studies which employed investigative techniques to study the post-disaster performance of transportation systems. Chang and Nojima (2001) developed post-disaster performance measures and applied them to rail and highway transportation networks in Kobe, Japan following the 1995 Hyogoken-Nanbu earthquake. Nakanishi et al. (2013) performed travel demand modeling for the recovery phase following the 2011 Tohoku earthquake and applied it to the city of Ishinomaki, Japan.

Recently, software and simulation techniques have also been developed to analyze transportation systems following a disaster such as an earthquake. The *California Department of Transportation* (Caltrans) uses ShakeCast and ShakeMap (Wald et al. 2008), developed by the *United States Geological Survey* (USGS), to obtain the likelihood of damage to bridges and other structures following an earthquake. These systems provide Caltrans with insight on how to reroute traffic due to potential road closures. Nojima and Sugito (2000) proposed a model to simulate and evaluate the post-earthquake functional performance of a highway transportation system using *Monte Carlo* (MC) simulation and the *modified incremental assignment method* (MIAM) to simulate the traffic states. Similarly, Shizunoka et al. (2003) integrated bridge fragility models with an MC analysis of traffic flows to study the degradation of capacity of the Caltrans network following the 1994 Northridge earthquake.

System reliability analysis (SRA) techniques have also been developed to understand the post-disaster behavior of transportation networks. SRA methods are capable of integrating analyses across physical scales and combining models and data from multiple fields of science and engineering for quantifying risk at the system level. For example, Lee et al. (2011) evaluated the probabilities of various damage scenarios using a matrix-based reliability method (Kang et al. 2008, 2012, Song and Kang 2009) and a network flow analysis for a sample bridge transportation network. While these software, simulation, and SRA techniques have been successful in determining failures of bridges in networks and changes in flow capacity, they do not explicitly model the traffic dynamics. The present work considers the integration of a traffic flow model, and the approach can be extended to general road networks provided a network traffic flow model is available (Garavello and Piccoli 2006).

Filters are sequential estimation algorithms that produce best estimates (in the sense of minimal variance) or posterior distributions of the state given a noisy model and observations that appear sequentially in time. Sequential estimation algorithms for traffic applications date back several decades to the 1970s when Gazis and Knapp (1971) and Szeto and Gazis (1972) used the *Kalman filter* (KF) (Kalman 1960) and the *extended Kalman filter* (EKF) (Anderson and Moore 1979), respectively, to estimate the traffic density in the Lincoln Tunnel of New York City for optimal traffic control applications.

Recent uses of the KF and EKF for traffic applications include Jabari and Liu (2013), Wang and Papageorgiou (2005), and van Hinsbergen et al. (2012). The *mixture Kalman filter* (MKF) (Chen and Liu 2000) is an extension of the KF and has been applied in the transportation community using a switching mode traffic model (Sun et al. 2004) and GPS data from cell phones (Herrera and Bayen 2010) for traffic state estimation. Another filter of note is the *unscented Kalman filter* (UKF) (Julier and Uhlmann 1997), which was developed for highly nonlinear systems and has been used to estimate traffic on a Belgian highway (Mihaylova et al. 2006).

Recently, stochastic filtering methods (Evensen 1994, Gordon et al. 1993) have been applied to traffic estimation. The performance of the *particle filter* (PF) (Gordon et al. 1993) has been tested on a Belgian freeway using experimental and synthetic traffic data (Mihaylova et al. 2007) as well as Interstate 66 in the United States (Chen et al. 2011). A recent application of the PF is given in Chen and Rakha (2014) for real-time travel time prediction. The PF has been further expanded into the *parallelized particle filter* (PPF) and *parallelized Gaussian sum particle filter* (PGSPF) for use in large-scale traffic network systems (Mihaylova et al. 2012). Work et al. (2010) applied the *ensemble Kalman filter* (EnKF) (Evensen 2009) to San Francisco Bay Area highway networks to estimate traffic velocities. Using a similar approach, Hong and Fukuda (2012) implemented an EnKF system for the Tokyo Metropolitan Highway during rush hour that showed the effect of sensor spacing on the estimation results.

Note the specific choice of the filter depends on the model, the measurement equation, and the computational constraints needed for real time implementation. The interested reader is directed to Blandin et al. (2012) for a recent review of sequential estimation techniques for scalar traffic models. While sequential estimation techniques have been used with traffic dynamics to accurately estimate traffic states, they have not been integrated with seismic hazard and bridge fragility models in post-disaster scenarios.

1.3 *Outline and contributions of this article*

The main contribution of this article is the design and analysis of an estimation framework to determine the post-disaster traffic conditions of a transportation network. We propose combining a seismic hazard model, a bridge fragility model, and a traffic model with traffic sensor data in an EnKF framework to account for the nonlinearity of the model. Figure 1 illustrates the framework. Earthquake data is obtained and used as inputs into the hazard model. The hazard model quantifies the severity of the earthquake through the peak ground acceleration which is then used in the fragility model to determine the damage state of the bridge and the resulting traffic capacity. This information is then fed into the traffic model, which is combined with traffic sensor data in the filtering framework to get a best estimate of the traffic conditions of the system. This estimate is fed back into the traffic model at the next time step.

The remainder of this work is summarized as follows. In Section 2, the traffic evolution, earthquake hazard, and probabilistic fragility models are introduced. Section 3 gives a brief introduction to the KF and EnKF sequential state estimation algorithms and the observation equation used in the proposed framework. Section 4 evaluates the framework for several numerical experiments performed on Interstate 155 West, and quantifies the benefit of using both knowledge of the earthquake characteristics and traffic sensor data as inputs into the estimator.

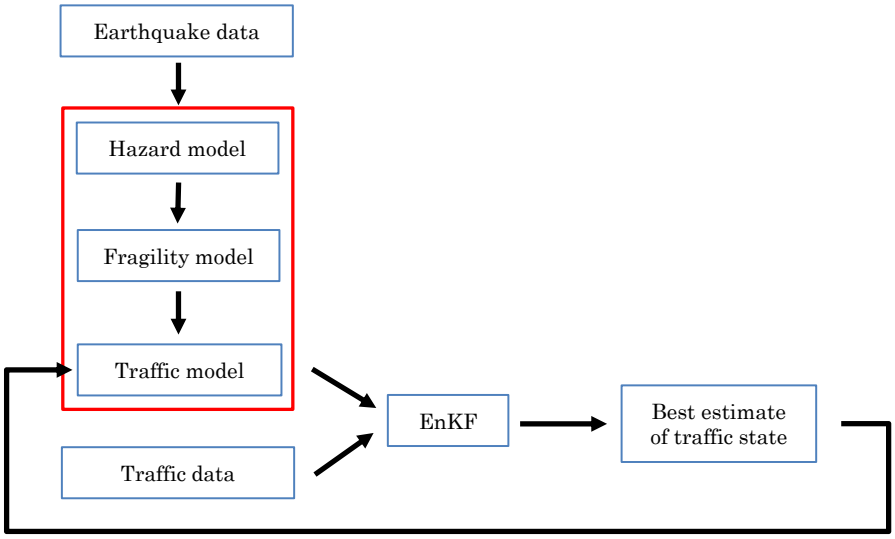


Figure 1. Proposed framework: The upper left boxed portion shows the three models. Earthquake data is used as an input into the forward model, which consists of the seismic hazard, bridge fragility, and traffic evolution models. These are integrated with traffic sensor data as inputs into the filter. The EnKF runs an error variance minimizing scheme to produce the best estimate of the traffic state. The estimated traffic states are fed into the forward model at the next time step.

2. Overview of traffic flow, seismic hazard, and earthquake fragility models

This section gives a brief introduction to the model of traffic dynamics, and the seismic hazard and bridge fragility models. Combining these elements is the basis for the forward model (Equation (3)).

2.1 *Macroscopic traffic model*

The *Lighthill-Whitham-Richards* (LWR) *partial differential equation* (PDE) (Lighthill and Whitham 1955, Richards 1956) is a conservation law used to model the evolution of the *traffic density*, $\rho(x,t)$, which is a measure of the number of vehicles per unit length at location x during time t (veh/km). It is a continuum model that describes the conservation of vehicles:

$$\frac{\partial \rho}{\partial t} + \frac{\partial Q(\rho)}{\partial x} = 0, \tag{5}$$

where $Q(\cdot)$ is the flux function, or *fundamental diagram*. The *traffic flow*, or *traffic flux*, is a measure of vehicle throughput per unit time (veh/h). The fundamental diagram is constructed from an empirical relationship between density and speed $v = V(\rho)$ as follows:

$$Q(\rho) = \rho V(\rho).$$

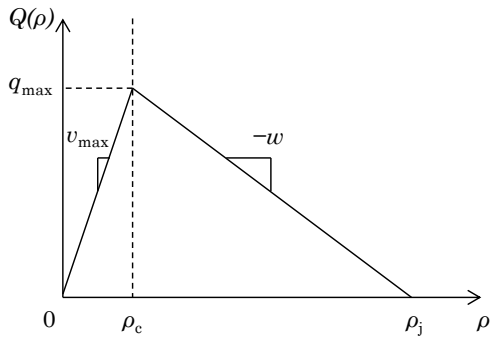


Figure 2. Newell-Daganzo flux function: Densities in the range $\rho \in [0, \rho_c]$ are in the free flow region and densities in the range $\rho \in [\rho_c, \rho_j]$ are in the congested region.

One widely used fundamental diagram is the triangular *Newell-Daganzo* flux function (Daganzo 1995, Newell 1993), shown in Figure 2. It is a piecewise linear function of the density, with different slopes in the free-flow and congestion regions:

$$Q(\rho) = \begin{cases} \rho v_{\max} & \text{if } \rho \in [0, \rho_c] \\ \frac{q_{\max}(\rho_j - \rho)}{\rho_j - \frac{q_{\max}}{v_{\max}}} & \text{if } \rho \in [\rho_c, \rho_j], \end{cases} \tag{6}$$

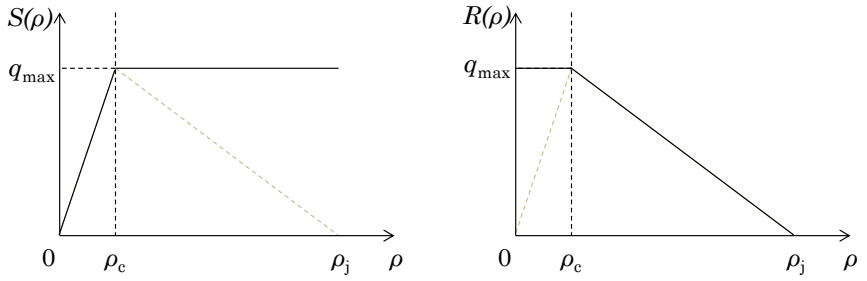
where $v_{\max}$ is the *free-flow speed*, ρ_c is the *critical density* where traffic transitions from free flow to congestion, and ρ_j is the *jam density* where traffic is fully congested and vehicles are stationary. The *maximum flux* is given by $q_{\max} = \rho_c v_{\max}$ and the maximum backward propagating wave speed is given by $w = q_{\max}/(\rho_j - \rho_c)$. From Equation (6), it is observed that the three variables $q_{\max}, v_{\max}$, and ρ_j are sufficient to define the triangular fundamental diagram (Figure 2).

For numerical implementation, Equation (5) is discretized in time and space using a *Godunov scheme* (Godunov 1959) as described below. The time and space domains are discretized by introducing a discrete time step ΔT, indexed by $n \in \{0, \cdots, n_{\max}\}$, and a discrete space step Δx, indexed by $i \in \{0, \cdots, i_{\max}\}$. The space discretization yields *cells*. The density at the next time step $(n+1)$ for cell i is computed as

$$\rho_{n+1}^i = \rho_n^i + \frac{\Delta T}{\Delta x}\left(q\left(\rho_n^{i-1}, \rho_n^i\right) - q\left(\rho_n^i, \rho_n^{i+1}\right)\right), \tag{7}$$

where the numerical flux $q(\cdot, \cdot)$ is a function used to compute flow across the boundary of two adjacent cells. Boundary conditions are imposed through ghost cells at -1 and $i_{\max} + 1$ (Herrera and Bayen 2010). Equation (7) is commonly referred to as the *cell transmission model* (CTM) (Daganzo 1994, 1995) for triangular fundamental diagrams. The quantity $q\left(\rho_n^i, \rho_n^{i+1}\right)$ is the flow between cell i and cell $i+1$ during time n and time $n+1$, which can be computed as

$$q\left(\rho^i, \rho^{i+1}\right) = \min\left\{S\left(\rho^i\right), R\left(\rho^{i+1}\right)\right\}, \tag{8}$$



(a) Sending function (b) Receiving function

Figure 3. Sending and receiving functions of the Newell-Daganzo flux function: The fundamental diagram is shown as the light dotted line. In free flow conditions, the sending function shares the same behavior as the fundamental diagram. In congested conditions, the receiving function shares the same behavior as the fundamental diagram.

where the *sending function*, $S(\cdot)$, and *receiving function*, $R(\cdot)$, represent the traffic that is sent by the upstream cell and the traffic that is received by the downstream cell, respectively, and ρ^i and ρ^{i+1} represent the densities of the upstream and downstream cells. The sending and receiving functions are constructed from the fundamental diagram as shown in Figure 3. Intuitively, as the upstream density increases, more vehicles are available to be sent into the downstream cell, resulting in an increased flow, up to the maximum flow, $q_{\max}$. Beyond the critical density, ρ_c, even if the density continues to increase, the flow sent to the downstream cell cannot be greater than $q_{\max}$ (Figure 3(a)). Similarly, when there are few vehicles in the downstream cell, the downstream cell can receive a flow up to $q_{\max}$. This is true for any $\rho^{i+1} \leq \rho_c$. If the density continues to increase, the flow that can be received decreases (Figure 3(b)).

2.1.1 Non-additive noise model

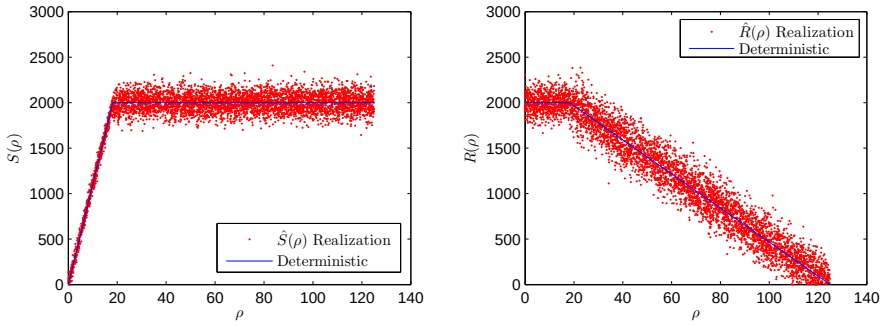
Often, the evolution equation (Equation (1)) is formulated with additive noise $\boldsymbol{w}_n$, which conserves mass in expectation (Simon 2006). However, in this work, a non-additive noise model is assumed for Equation (3). The noise is instead embedded in the empirical fundamental diagram (Equation (6)) and thus directly affects the sending and receiving functions. This results in mass being conserved always, not just in expectation. A normal distribution is used for the noise and it is added to the fundamental diagram to generate more realistic sending and receiving functions consistent with experimental data from inductive loops (Caltrans 2014):

$$\hat{S}(\rho) = \begin{cases} S(\rho) + \varepsilon_S & \text{if } \rho \in [0, \rho_c] \\ S(\rho) + \varepsilon_Q & \text{if } \rho \in [\rho_c, \rho_j], \end{cases} \tag{9}$$

and

$$\hat{R}(\rho) = \begin{cases} R(\rho) + \varepsilon_Q & \text{if } \rho \in [0, \rho_c] \\ R(\rho) + \varepsilon_R & \text{if } \rho \in [\rho_c, \rho_j], \end{cases} \tag{10}$$

where $\varepsilon_S \sim \mathcal{N}\left(\mu_S, \sigma_S^2\right)$ and $\varepsilon_R \sim \mathcal{N}\left(\mu_R, \sigma_R^2\right)$ are noises associated with the free flow



(a) Noisy sending function (b) Noisy receiving function

Figure 4. Illustration of noise models of the sending and receiving functions. There are 5000 sample points on each graph to illustrate the relative magnitudes of the means and variances associated with each function.

region of the sending function and the congested region of the receiving function, respectively, and μ and σ^2 are the mean and variance of the distribution. A third noise, $\varepsilon_Q \sim \mathcal{N}\left(\mu_Q, \sigma_Q^2\right)$, is introduced to represent the noise on the maximum flow. To introduce unbiased noise, we let $\mu_S = \mu_R = \mu_Q = 0$. The noise on the receiving function is assumed to be greater than that on the sending function to be consistent with data from inductive loops (Caltrans 2014). An example of noisy sending and receiving functions is shown in Figure 4, where the solid line represents the deterministic sending and receiving functions.

Note that for densities near zero or the jam density, Equations (9) and (10) may result in a negative flow. To avoid negative flows across cell boundaries, the numerical flow is taken as the maximum of the minimum value of Equations (9) and (10), and 0. Thus, Equation (8) is rewritten as

$$q\left(\rho^i, \rho^{i+1}\right) = \max\left\{\min\left\{\hat{S}\left(\rho^i\right), \hat{R}\left(\rho^{i+1}\right)\right\}, 0\right\}.$$

2.2 *Seismic hazard and fragility models*

Two main characteristics of an earthquake event are its magnitude and location. Using these characteristics and site conditions of the structure, peak ground parameters are computed through the use of attenuation relationships. By mapping these parameters to existing fragility models, the probability of a bridge being in prescribed damage states is determined. The maximum traffic capacity for each bridge based on a realization of the bridge damage distribution is then obtained.

2.2.1 Attenuation relationship

The *peak ground acceleration* (PGA) exerted on a structure from an earthquake is dependent on a number of factors, including the magnitude of the earthquake, the distance of the structure from the epicenter, the site conditions (i.e., soil profile), and the fault type. PGA is related to these factors by attenuation relationships. In this work, the attenuation relationship by Campbell (1997) is used as an example:

$$\begin{aligned}\ln A_{\mathrm{H}} = &-3.512 + 0.904M - 1.328 \ln \sqrt{r^2 + \left(0.149 \exp\left(0.647M\right)\right)^2} \\ &+ (0.440 - 0.171 \ln r) s_{\mathrm{sr}} + (0.405 - 0.222 \ln r) s_{\mathrm{hr}} \\ &+ (1.125 - 0.112 \ln r - 0.0957M) F_{\mathrm{t}},\end{aligned} \tag{11}$$

where A_{H} is the horizontal component of the PGA (in g), M is the magnitude, r is the distance to the epicenter (in km), F_{t} is the fault type, and s_{sr} and s_{hr} define the local site conditions. Equation (11) can be written as follows:

$$A_{\mathrm{H}} = f_{\mathrm{e}}\left(M, r, s_{\mathrm{sr}}, s_{\mathrm{hr}}, F_{\mathrm{t}}\right). \tag{12}$$

2.2.2 Fragility curves

A fragility curve defines the conditional probability of exceeding a prescribed *limit state* (LS) as a function of a selected *intensity measure* (IM) (e.g., horizontal peak ground acceleration) of ground motions. The fragility is often modeled by a lognormal *cumulative distribution function* (CDF), where the *demand* D and *capacity* C are assumed to be independent lognormally distributed variables. The *safety factor* F (Rosenblueth and Esteva 1972) can then be computed as

$$F = \ln C - \ln D,$$

where $F \leq 0$ represents the failure event. Since C and D are both assumed to be lognormal variables, their logarithms are both normal, and thus F is also normal. Since $F \sim \mathcal{N}\left(\mu_F, \sigma_F^2\right)$, the variable $Z = \left(F - \mu_F\right)/\sigma_F$ follows the standard normal distribution, i.e., $Z \sim \mathcal{N}\left(0, 1\right)$. The conditional probability of exceedance given an intensity measure is given as

$$\begin{aligned}P\left(F \leq 0 | IM\right) &= P\left(Z \leq -\frac{\mu_F}{\sigma_F}\Big| IM\right) \\ &= \Phi\left(-\frac{\mu_F\left(IM\right)}{\sigma_F\left(IM\right)}\right),\end{aligned} \tag{13}$$

where $\Phi(\cdot)$ is the standard normal CDF. Since there are multiple limit states possible (based on different levels of drift, peak ground accelerations, etc.), the conditional probability is computed for each limit state. The fragility curve used in this work is a univariate model where the probability of exceedance is only dependent on the PGA in the horizontal direction, A_{H} (in g). As a result, Equation (13) can be rewritten in the following form (Nielson and DesRoches 2007):

$$P\left(F_i \leq 0 | A_{\mathrm{H}}\right) = \Phi\left(\frac{\ln A_{\mathrm{H}} - \ln\left(\tilde{x}_i\right)}{\zeta_i}\right),$$

where F_i is the safety factor, $\tilde{x}_i$ is the median capacity of the structure in terms of PGA,

Table 1. Median and dispersion values in g for seismic fragility curves of bridges (Nielson and DesRoches 2007).

Bridge Class	Median PGA (g)			ζ
	Slight	Moderate	Complete	
MSC concrete	0.15	0.52	1.03	0.70
MSC steel	0.18	0.31	0.50	0.55
MSSS concrete	0.20	0.57	1.17	0.65
MSSS steel	0.24	0.44	0.82	0.50

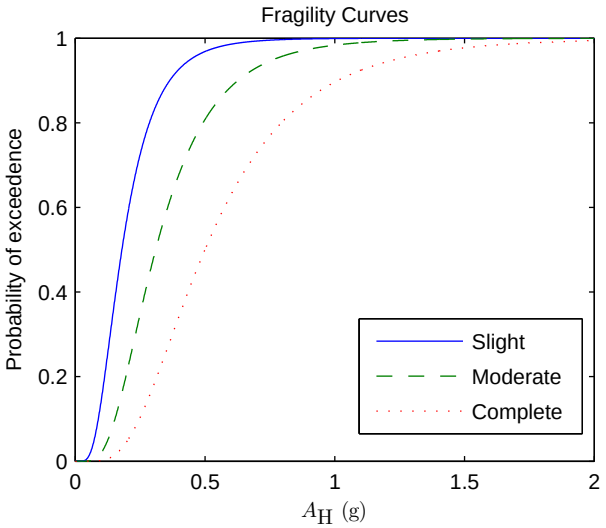


Figure 5. Fragility model for MSC steel bridges: The fragility curves represent the three limit states.

and ζ_i is the dispersion value belonging to the ith limit state. Table 1 provides the fragility curve parameters for several bridge classes typical to the Central and Southeastern United States, excerpted from Nielson and DesRoches (2007).

For simplicity, we assume three limit states, *slight*, *moderate*, and *complete*, as shown in Table 1 for *multispan continuous* (MSC) steel girder bridges (Nielson and DesRoches 2007). The fragility curve corresponding to each limit state is given in Figure 5. Intuitively, the probability of exceedance of a less severe limit state is greater than that of a more severe limit state. Also, as A_H increases, the probability of exceedance of any limit state increases monotonically.

2.2.3 Computing the probability of damage states

Once the fragility curves have been modeled, it is possible to determine the probabilities of being in different damage states by computing the difference between adjacent fragility curves for a given A_H. The three limit states described previously yield four *damage states* (DS) defined as $\mathcal{D} = \{insignificant, medium, high, total\}$, following a similar procedure to Bai et al. (2009).

Consider the probabilities of exceedance of the three limit states denoted P_S (*slight*), P_M (*moderate*), and P_C (*complete*), which are listed in increasing severity. The second column of Table 2 shows how the probability of occurrences of each damage state is computed using the fragilities for a given A_H. If damage events $d \in \mathcal{D}$ are defined such that they are *mutually exclusive and collectively exhaustive* (MECE), for a given A_H the following holds:

Table 2. Probability of being in different damage states and their associated traffic capacity factors, α (Mackie and Stojadinović 2006).

Damage state	Probability	Traffic capacity factor, α
insignificant	$1 - P_{\mathrm{S}}$	1
medium	$P_{\mathrm{S}} - P_{\mathrm{M}}$	0.75
high	$P_{\mathrm{M}} - P_{\mathrm{C}}$	0.5
total	P_{C}	0

$$\sum_{d \in \mathcal{D}} P\left(d|A_{\mathrm{H}}\right) = 1.$$

Assuming damage states are MECE events is important because it allows $\mathcal{D}$ to define the full distribution of damage states given the earthquake magnitude and location, and site conditions.

2.2.4 Mapping damage states to traffic capacity

We assume quarter-based capacities are used to relate the bridge damage state to traffic capacity, similar to the assumption used by Mackie and Stojadinović (2006) and Murachi et al. (2003). The third column of Table 2 shows the mapping between the damage state and traffic capacity factor, α. Note that this mapping is used as an example and more detailed models could be constructed.

2.2.5 Fragility model as an input to the traffic model

At a given time step $n \geq n_{\mathrm{EQ}}$ (i.e., after the occurrence of the earthquake), the probability distribution of the damage states (Section 2.2.3) using the earthquake magnitude and location, and site conditions can be computed. This distribution of damage states maps to a distribution of the traffic capacity factor (Section 2.2.4) which can be included explicitly as an input into the CTM (Equation (3)). In this way, the existence of an earthquake and its subsequent damage to the transportation infrastructure has a direct influence on the maximum flow $q_{\max}$, and jam density ρ_j, at a given instance in time.

3. Sequential state estimation

This section briefly reviews the development of the ensemble Kalman filter (Evensen 1994) sequential state estimation algorithms used in this work. The observation (Equation (2)) is also defined.

3.1 *Kalman filter*

When the model evolution and observation equations are linear with additive white noise, Equations (1) and (2) can be written as the following:

$$\boldsymbol{x}_n = \mathbf{F}_n \boldsymbol{x}_{n-1} + \boldsymbol{w}_n$$

$$\boldsymbol{y}_n = \mathbf{H}_n \boldsymbol{x}_n + \boldsymbol{v}_n,$$

where $\mathbf{F}_n$ is the linear model operator and $\mathbf{H}_n$ is the linear measurement matrix, which is also indexed by n to integrate the possibility of moving sensors (e.g. in vehicles) and intermittently operating fixed sensors. The model and measurement noises are $\boldsymbol{w}_n \sim \mathcal{N}(0, \mathbf{W}_n)$ and $\boldsymbol{v}_n \sim \mathcal{N}(0, \mathbf{R}_n)$, respectively, where $\mathbf{W}_n$ is the model error covariance matrix and $\mathbf{R}_n$ is the measurement error covariance matrix at time n. The Kalman filter (Kalman 1960) computes the posterior state, $\boldsymbol{x}_{n|n}$, given $\boldsymbol{x}_{n-1|n-1}$ and measurements $\boldsymbol{y}_n$, and is given as

$$\text{Prediction:} \begin{cases} \boldsymbol{x}_{n|n-1} = \mathbf{F}_n \boldsymbol{x}_{n-1|n-1} \\ \mathbf{P}_{n|n-1} = \mathbf{F}_n \mathbf{P}_{n-1|n-1} \mathbf{F}_n^{\mathrm{T}} + \mathbf{W}_n \end{cases} \tag{14}$$

$$\text{Update:} \begin{cases} \mathbf{K}_n = \mathbf{P}_{n|n-1} \mathbf{H}_n^{\mathrm{T}} \left(\mathbf{H}_n \mathbf{P}_{n|n-1} \mathbf{H}_n^{\mathrm{T}} + \mathbf{R}_n \right)^{-1} \\ \boldsymbol{x}_{n|n} = \boldsymbol{x}_{n|n-1} + \mathbf{K}_n \left(\boldsymbol{y}_n - \mathbf{H}_n \boldsymbol{x}_{n|n-1} \right) \\ \mathbf{P}_{n|n} = \mathbf{P}_{n|n-1} - \mathbf{K}_n \mathbf{H}_n \mathbf{P}_{n|n-1}. \end{cases} \tag{15}$$

The model prediction step (Equation (14)) propagates the mean ($\boldsymbol{x}_{n-1|n-1}$) and covariance ($\mathbf{P}_{n-1|n-1}$) of the state at the previous time $n-1$ forward through $\mathbf{F}_n$ to obtain $\boldsymbol{x}_{n|n-1}$ and $\mathbf{P}_{n|n-1}$. The measurement update step (Equation (15)) computes the posterior mean ($\boldsymbol{x}_{n|n}$) and covariance ($\mathbf{P}_{n|n}$) by taking into account observations given up to step n. The Kalman gain, $\mathbf{K}_n$, is chosen such that the resulting filter is a *best linear unbiased estimator* (BLUE) of the state for linear systems, where best is in the sense of minimal variance of the posterior state error covariance (Simon 2006).

3.2 *Ensemble Kalman filter*

Due to the nonlinearities of the CTM (Equation (7)), a nonlinear extension of the KF, known as the ensemble Kalman filter (Evensen 1994), is used in this work. The EnKF is also computationally cheap compared to the particle filter, and has improved accuracy compared to extended Kalman filters (Blandin et al. 2012), thereby motivating the use of the EnKF in the present work.

The EnKF is a sample (ensemble) approximation of the KF, which represents the prior and posterior distributions by ensembles. Rather than obtaining the covariance matrix through $\mathbf{F}_n$ (Equation (14)), the EnKF evolves the covariance matrix indirectly through the individual ensemble evolutions of the prior distribution. The ensemble representations of the model states and observations indexed by $e \in \{1, \cdots, N\}$ are

$$\boldsymbol{x}_n^e = \boldsymbol{f}\left(\boldsymbol{x}_{n-1}^e, \boldsymbol{\theta}^e, \boldsymbol{u}_n^e, \boldsymbol{w}_n^e \right)$$

$$\boldsymbol{y}_n^e = \boldsymbol{y}_n + \boldsymbol{v}_n^e, \tag{16}$$

where $\boldsymbol{w}_n^e$ and $\boldsymbol{v}_n^e$ are the realizations of the model noise, $\boldsymbol{w}_n$, and measurement noise, $\boldsymbol{v}_n$. The variables $\boldsymbol{u}_n^e$ and $\boldsymbol{\theta}^e$ are also indexed by e to show that each ensemble is associated with its own realization from the damage state distribution. In the limit of an infinite number of ensembles, the EnKF converges to the KF for linear systems with Gaussian additive noises (Mandel et al. 2011). The model prediction and measurement update steps for the EnKF are given as

$$\text{Prediction: } \begin{cases} \boldsymbol{x}_{n|n-1}^e = \boldsymbol{f}\left(\boldsymbol{x}_{n-1|n-1}^e, \boldsymbol{\theta}^e, \boldsymbol{u}_n^e, \boldsymbol{w}_n^e\right) \\ \boldsymbol{x}_{n|n-1} = \frac{1}{N}\sum_{e=1}^{N} \boldsymbol{x}_{n|n-1}^e \\ \mathbf{P}_{n|n-1} = \frac{1}{N-1}\sum_{e=1}^{N}\left(\boldsymbol{x}_{n|n-1}^e - \boldsymbol{x}_{n|n-1}\right)\left(\boldsymbol{x}_{n|n-1}^e - \boldsymbol{x}_{n|n-1}\right)^{\mathrm{T}} \end{cases} \tag{17}$$

$$\text{Update: } \begin{cases} \mathbf{K}_n = \mathbf{P}_{n|n-1}\mathbf{H}_n^{\mathrm{T}}\left(\mathbf{H}_n\mathbf{P}_{n|n-1}\mathbf{H}_n^{\mathrm{T}} + \mathbf{R}_n\right)^{-1} \\ \boldsymbol{x}_{n|n}^e = \boldsymbol{x}_{n|n-1}^e + \mathbf{K}_n\left(\boldsymbol{y}_n^e - \mathbf{H}_n\boldsymbol{x}_{n|n-1}^e\right) \\ \mathbf{P}_{n|n} = \mathbf{P}_{n|n-1} - \mathbf{K}_n\mathbf{H}_n\mathbf{P}_{n|n-1}. \end{cases} \tag{18}$$

The formulation of the EnKF shown above is identical to that used in (Blandin et al. 2012), with the exception that the algorithm presented here is generalized to include the dependency of the forward model $\boldsymbol{f}$ on the input $\boldsymbol{u}$ and the parameters $\boldsymbol{\theta}$. After the posterior ensembles from the previous time ($\boldsymbol{x}_{n-1|n-1}^e$) have been propagated forward in time through the nonlinear model $\boldsymbol{f}$ to obtain the prior ensembles, $\boldsymbol{x}_{n|n-1}^e$, the sample mean and the state error covariance matrix are computed from the ensembles. In the analysis step (Equation (18)), a posterior ensemble, $\boldsymbol{x}_{n|n}^e$ is computed using the prior ensemble (Equation (17)) and the ensemble of observations (Equation (16)). By updating each ensemble member using a perturbed observation $\boldsymbol{y}_n^e$, a new set of ensembles is created which has the correct error statistics after the update step (Evensen 2003). The posterior state mean, $\boldsymbol{x}_{n|n}$, can be computed similarly to $\boldsymbol{x}_{n|n-1}$. Notice the posterior error covariance, $\mathbf{P}_{n|n}$, and Kalman gain, $\mathbf{K}_n$, are computed in exactly the same way as that of the KF (Equation (15)).

Looking at Equation (16), it is possible that measurements may be perturbed outside the physical bound of the quantity to be measured. One could throw away these samples if desired, or add a non Gaussian measurement noise to prevent non-physical perturbed measurements.

Equations (17) and (18) illustrate the structural similarities to the linear KF (Equations (14) and (15)). For the numerical experiments shown next, we use a more efficient implementation described in Evensen (2003) and Evensen (2009). The implementation avoids building the covariance matrix explicitly, which reduces the computational cost of the algorithm. The implementation is used purely for numerical efficiency and is therefore not explained in this work; the interested reader is directed to Evensen (2003) and Evensen (2009) for further details and to the supplementary source code available at `https://github.com/rotsuka/postDisaster_EnKF`.

Note it is also possible to augment the state vector, $\boldsymbol{x}$, by the model parameters (e.g., q_{max}) and estimate the parameters directly through the filter. Wang and Papageorgiou (2005) used the extended Kalman filter to perform joint traffic state and parameter estimation, and capacity reductions due to incidents is explored in Wang and Work (2014) assuming the capacity reduction is a discrete variable. The main constraint is

the capacity drop is not always observable, which can cause poor performance of the filter unless good models of the parameter evaluation exist to assure convergence. As a result, Equation (16) treats the capacity reductions as noisy inputs to the traffic model to improve the traffic estimates.

3.3 *Observation equation*

We assume that the traffic density is observed at every time step from sensors located at various positions on the road network. Thus, the operator $\boldsymbol{h}$ in Equation (2) that relates the traffic state $\boldsymbol{x}_n = \left(\rho_n^0, \cdots, \rho_n^{i_{\max}}\right)^{\mathrm{T}}$ at a time n to the sensor measurements must be defined. The observation operator $\boldsymbol{h}$ is defined as follows:

$$\boldsymbol{h}_n\left(\boldsymbol{x}_n\right) = \mathbf{H}_n \boldsymbol{x}_n.$$

For traffic sensors indexed by $l \in \{1, \cdots, l_{\max}\}$, the measurement operator, $\mathbf{H}_n \in \{0,1\}^{l_{\max} \times (i_{\max}+1)}$, is defined as

$$\mathbf{H}_{ij} = \begin{cases} 1 & \text{if the } i\text{th measurement is obtained in cell } j+1 \\ 0 & \text{otherwise.} \end{cases}$$

Assuming measurement errors v_n are independent, the measurement error covariance matrix $\mathbf{R}_n$ is diagonal and given as

$$\mathbf{R}_n = \operatorname{diag}\left(\sigma_{n,1}^2, \cdots, \sigma_{n,l_{\max}}^2\right), \tag{19}$$

where $\sigma_{n,l}^2$ is the variance of sensor l at time n.

4. Numerical experiments

This section describes the numerical experiments to quantify the potential improvements achieved by linking the seismic hazard and bridge fragility models to the traffic model updated by traffic sensor data. First, a transportation network based on Interstate 155 West near the New Madrid Seismic Zone in New Madrid, Missouri is constructed and the discretization of the network is described. Using prescribed initial and boundary conditions, the *true model* is run using a set of *true parameters*, including a reduction in capacity due to an earthquake, to generate the *true state* to be estimated. The true state serves to assess the accuracy of the estimates and it is also used to generate the synthetic measurements in the filter.

In the estimation algorithm, a separate model called the *approximate model* is used, which has different initial and boundary conditions and traffic parameters. To test the value of various information sources (earthquake information, traffic sensors), several estimation algorithms are run with different sensor measurements and inputs. The four estimation algorithms are the following:

(1) Neither earthquake information nor traffic sensor data available

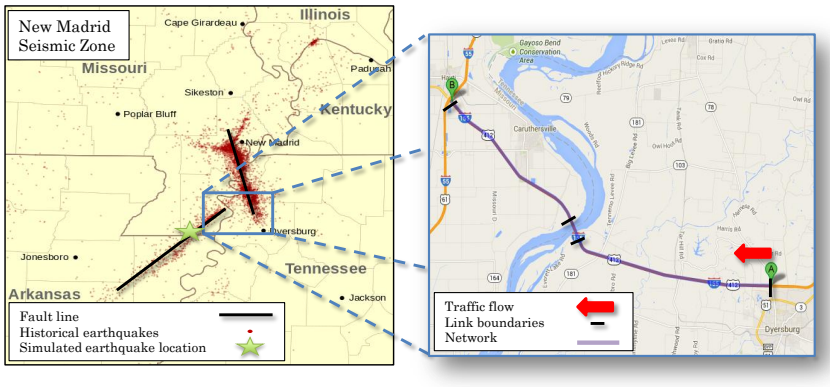


Figure 6. Location of network: The picture on the left shows the location of the network, as well as the location of the earthquake. The dots represent all locations of earthquakes from 1974-2011. The figure also shows the approximate orientations of the fault lines in the New Madrid Seismic Zone. The picture on the right shows a close-up of the transportation network, the network discretization into links, and the direction of traffic flow (Points A to B).

(2) No earthquake information, but traffic sensor data available
(3) Earthquake information available, but no traffic sensor data
(4) Both earthquake information and traffic sensor data available

Estimation algorithms that do not account for traffic sensor data are *open loop* models while those that do account for traffic sensor data are *filters*. The approximate models, regardless if they are open loop or filters, provide estimates. To quantify performance of the filter and to compare improvements from one estimation algorithm to others (e.g. filter vs. open loop, or earthquake vs. no earthquake), we use the *Bayesian estimation error quotient* (BEEQ) (Li and Zhao 2005). The BEEQ is also used to compare filter performance in terms of computational cost when different numbers of ensembles are used.

4.1 *Discretization of transportation network*

As shown in Figure 6, we simulate a 37.5 km stretch of roadway from the junction of Interstate 155 (I-155) and Tennessee State Route (SR) 78 in Dyersburg, TN across the Caruthersville Bridge to the intersection of I-155 and I-55 in Hayti, MO. For simplicity, we assume ramp inflows and outflows are negligible, although the framework can be easily extended to support them (Daganzo 1995). This network contains one bridge, the Caruthersville Bridge, which spans the Mississippi River. The network lies in the *New Madrid Seismic Zone* (NMSZ), as shown in Figure 6.

There are indeed many other pieces for infrastructure (e.g., overpasses, underpasses) which may cause disruptions in traffic flow. These infrastructure elements could be included in the model provided that fragility models exist and that the location of the infrastructure is integrated into the model. For simplicity and pedagogy, the numerical example only considers the dependence on the bridge.

Figure 6 also shows the discretization of the network into links used in the experiments while Figure 7 shows the linear representation of this network, which contains three links of varying length. On each link the parameters of the roadway are constant. Links are further subdivided into cells and the density in each cell creates the network traffic state vector, $\boldsymbol{x}$. The total number of cells in the network is 80.

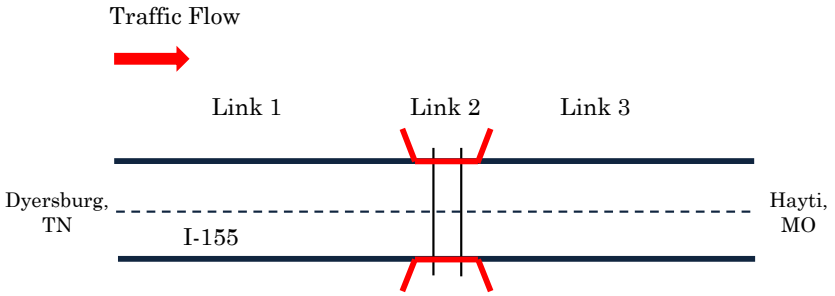


Figure 7. Schematic of model simulation: The roadway is composed of three links of varying length and number of cells.

Table 3. Discretization of road network.

Link	1	2	3
Length (km)	19.7	2.2	15.6
Cells	42	4	34
Δx (km)	0.469	0.550	0.459
Lanes	2	2	2

The Caruthersville Bridge is represented by Link 2. It is assumed there are nine traffic sensors located throughout I-155 West spaced approximately four km apart, and there are no sensors along the bridge. Table 3 describes the discretization in detail. The simulation is run for a duration ($t_{\max}$) of 40 minutes with a time step (ΔT) of 15 seconds (0.25 minutes), resulting in 161 ($n_{\max}$) time steps starting at $t = 0$ minutes. Since the speed limit does not vary along I-155 West, these links are discretized into cells of approximately equal length as shown in Table 3. The discretization obeys the *Courant-Friedrichs-Lewy* (CFL) condition (Courant et al. 1967), $\Delta x \geq v_{\max}\Delta T$, which is required for numerical stability of the CTM (Equation (7)).

4.2 *True model parameters and assumptions*

4.2.1 Bridge fragility and damage scenarios

The Caruthersville Bridge is an MSC steel bridge, following the classification of Nielson and DesRoches (2007). The site condition (soil profile) is assumed to be alluvium, thus $s_{\text{sr}} = s_{\text{hr}} = 0$. The bridge is located 15 km from the earthquake epicenter (i.e. $r = 15$), as shown in Figure 6.

In this work, different scenarios for the *true model* are defined by the resulting damage state of the Caruthersville Bridge following an earthquake event. In addition to earthquakes which cause insignificant damage (i.e., flow is uninterrupted), we assume there are three types of earthquakes for the true model, corresponding to the various damage states of the bridge, which are shown in Table 4. We assume the earthquake originates along the New Madrid fault, which has a strike-slip mechanism ($F = 0$).

When the values corresponding to the site conditions ($s_{\text{sr}}, s_{\text{hr}}$), the fault ($F$), and Table 4 are substituted into Equation (12), the resulting PGA value when mapped to the fragility model (Figure 5) results in the damage state defined by those variables to have the highest probability of occurrence.

Table 4. Earthquake characteristics for the true and approximate models.

Magnitude, M		Distance (km), r		Damage
True	Approximate	True	Approximate	
6.5	$\mathcal{N}(6.5, 0.3^2)$	15	$\mathcal{N}(15, 3^2)$	*medium*
7.5	$\mathcal{N}(7.5, 0.3^2)$	15	$\mathcal{N}(15, 3^2)$	*high*
8.5	$\mathcal{N}(8.5, 0.3^2)$	15	$\mathcal{N}(15, 3^2)$	*total*

Table 5. Initial and boundary density conditions in veh/km and traffic parameters for the true and approximate models.

Link	True	Approximate	Error (%)
1	30	10	66.7
2	30	10	66.7
3	30	10	66.7
UBC	40	30	25
DBC	20	10	50
$v_{\max}$ (km/h)	110	$\mathcal{N}(110, 3.3^2)$	-
$q_{\max}$ (veh/h/lane)	2000	$\mathcal{N}(2000, 100^2)$	-
ρ_{j} (veh/km/lane)	125	$\mathcal{N}(125, 6.25^2)$	-

4.2.2 Parameters and assumptions

The model parameters are given in Table 5, which is split into two sections: initial and boundary density conditions and link traffic parameters. Note that the densities are given in units of veh/km, while the parameters are given in units of veh/h/lane and veh/km/lane. The parameters of the true model are given in the second column of Table 5. The traffic parameters along I-155 are constant across the links. The fundamental diagram (Figure 2) is computed using $v_{\max}$, $q_{\max}$, and ρ_{j}.

The true model is initialized with the traffic in free flow and the *downstream boundary condition* (DBC) also in free flow. The *upstream boundary condition* (UBC) is a congested density, which given the free flow initial condition, generates maximum inflow on the upstream boundary. At $t = 10$ minutes, an earthquake occurs of magnitude M a distance r away from the bridge (Table 4). A realization of α is drawn from the damage state distribution and this is used to compute a realization of $q_{\max}$ to propagate the forward deterministic model in Equation (7). This is the true state which is used to compare the accuracy of the traffic estimates. The density contour in the time-space domain is shown in Figure 8(a) for an earthquake which resulted in a 50% reduction in traffic capacity of the bridge (i.e. *high* damage). The true state will be further analyzed in subsequent sections.

4.3 *Approximate model parameters and assumptions*

In the approximate models, an earthquake magnitude and epicenter location is drawn for each ensemble at each time step from normal distributions whose means are the true characteristics (Table 4). For the following experiments, the estimation algorithms use $N = 200$ ensembles. We assume accurate knowledge of the earthquake magnitude and location are available once the earthquake occurs. Therefore, the variances of the distribution assigned to the noises for the magnitude and distance are fairly small. In the approximate model, the parameters that compose the fundamental diagram are drawn for each ensemble from distributions centered around the true values from the distributions defined in Table 5. Thus, each link for each ensemble has a different realization of the

Table 6. Filter parameters of the approximate model.

σ_{sensor} (veh/km)	10
ε_S (veh/h)	$\mathcal{N}(0, 50^2)$
ε_R (veh/h)	$\mathcal{N}(0, 100^2)$
ε_Q (veh/h)	$\mathcal{N}(0, 150^2)$

fundamental diagram. Table 4 shows the distribution of the earthquake parameters (M, r) used in the approximate model.

In the approximate model, the measurements are synthesized by computing the true state and adding the measurement noise $\boldsymbol{v}_n$. We assume that the traffic sensors in the road network are fixed, thus $\mathbf{H}_n = \mathbf{H}$ (i.e., $\mathbf{H}$ is time-invariant). We also assume that each traffic sensor has the same accuracy and the accuracy of traffic sensors do not change after the occurrence of the earthquake. Thus, each entry in Equation (19) is defined with a constant standard deviation, σ_{sensor}, shown in Table 6. The noise parameters associated with the fundamental diagram (Equations (9) and (10)) are also shown in Table 6.

We initialize the approximate model with the traffic and boundary conditions given in the third column of Table 5. Each link for each ensemble is associated with a different realization of the fundamental diagram. The mean values (Table 5, second column) of the initial and boundary conditions for approximate models are chosen in free flow. The table also illustrates the relative error of the approximate model initial and boundary conditions compared to those of the true model. The relative error then quantifies the amount of error in the parameters used in the true model, and the parameters used in the approximate model in the estimator. The values in Table 5 show that the approximate model is initialized with fewer vehicles than the true model. Thus, there is less chance for congestion in the approximate model if the draw from the damage distribution does not result in a reduced traffic capacity of the bridge.

4.4 *Results and discussion*

In the true state (Figure 8(a)), at the start of the simulation ($t < 10$ minutes), there is no congestion in the network, because the network is initialized in free flow and there are no merges to create bottlenecks in the network. At time $t = 10$ minutes, the earthquake occurs and damages the bridge (Link 2) such that its traffic capacity is reduced by 50% ($\alpha = 0.5$). This results in a backward propagating shock wave which starts at the end of Link 1. Due to this bottleneck, downstream of the shock wave the vehicle density decreases because the outflow from the bridge is decreased due to the damage.

The subplots of Figure 8 show the various estimation results using the four algorithms described at the beginning of this section. As a baseline of the estimator, an open loop model with no knowledge of the earthquake input is used (Figure 8(b)). Figures 8(c) and 8(d) show the estimation results when there is either traffic data or knowledge of the earthquake input, but not both. Figure 8(e) shows the estimation results when there is both traffic data and knowledge of the earthquake input.

Since the model is initialized in free flow, when there is no input from the earthquake in the forward model (Figure 8(b)), the traffic conditions will remain in free flow. Having no earthquake input into the model is numerically similar to having no earthquake event or one that only results in *insignificant* damage to the bridge, thus no disruption in traffic flow. When there is no earthquake input, but there is information fed to the filter via traffic sensors (Figure 8(c)), congestion is detected in cells where sensors are located,

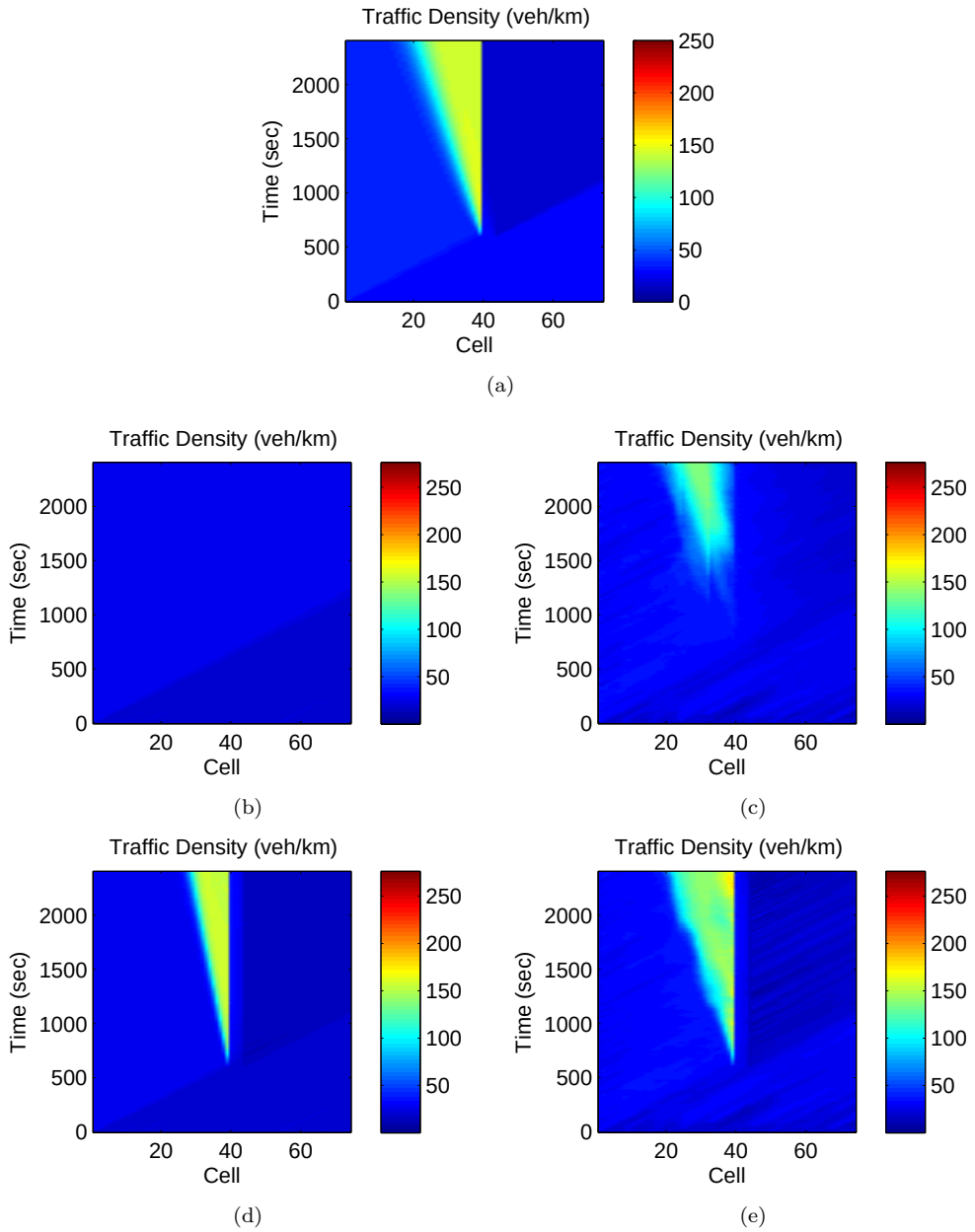


Figure 8. Time-space density plots for *high* damage: (a) Before the earthquake occurs there is no congestion in the network. When an earthquake occurs that results in a 50% reduction of traffic capacity of the bridge, a shock forms at the end of Link 1. (b) Without an earthquake input into the model or traffic sensor data, the model fails to predict any congestion. (c) With sensor data, the model accurately predicts congestion and backward propagating shockwave is detected. (d) Supplementing the model with data from the earthquake improves the model even if there is no sensor data. (e) The estimator in (d) performs better when both traffic sensor and earthquake data are also used.

and the shock wave propagates backward. However, the congestion is not detected by the model until the shockwave is detected by a sensor. Figure 8(c) illustrates that the sensor did not detect congestion until 25 minutes into the simulation, whereas the true state (Figure 8(a)) has congestion starting from 10 minutes.

Figures 8(d) and 8(e) show the results of the approximate model when the earthquake

is used as an input to the traffic model. Figure 8(d) shows that by having knowledge of the earthquake characteristics, the congestion due to the bridge capacity reduction is detected. However, there are two errors in the traffic estimates of the open loop model with the earthquake input. First, the speed at which the congestion wave travels in the open loop model is slower than in the true model. Thus, the region of congestion is smaller than in the true model. Second, the congestion levels are slightly overestimated. With the addition of the traffic sensor data (Figure 8(e)), the estimation of the state is significantly improved, and the duration, location, and magnitude of the shock wave are estimated accurately. This is an improvement on the models that have either traffic sensor data (Figure 8(c)) or the earthquake input (Figure 8(d)) and is a significant improvement on the open loop model (Figure 8(b)). The traffic estimate is still slightly incorrect, because the vehicle density on the bridge itself is unaffected, whereas in the true model the density decreases.

4.5 *Quantifying model performance*

To quantify the performance of open loop estimators and filters, the model performance is evaluated using the BEEQ error score (Li and Zhao 2005). The BEEQ of an estimator is a measure of improvement for an N_r number of independent estimates indexed by $r \in \{1, \cdots, N_\mathrm{r}\}$. It is defined by the *geometric mean*, $\beta = \left(\prod_{r=1}^{N_\mathrm{r}} \beta_r\right)^{1/N_\mathrm{r}}$, which is computed through its logarithm for numerical reasons:

$$\log \beta = \frac{1}{N_\mathrm{r}} \sum_{r=1}^{N_\mathrm{r}} \log \beta_r, \quad \beta_r = \frac{\|\boldsymbol{s}_r - \hat{\boldsymbol{s}}_r\|}{\|\boldsymbol{s}_r - \bar{\boldsymbol{s}}_r\|}, \tag{20}$$

where $\boldsymbol{s}_r$ represents the true state, $\bar{\boldsymbol{s}}_r$ represents the prior state, and $\hat{\boldsymbol{s}}_r$ represents the estimation state for the rth simulation run. For our purposes, $\bar{\boldsymbol{s}}_r$ is the open loop solution with no knowledge of the earthquake input and $\hat{\boldsymbol{s}}_r$ is one of three solutions: the filter solution with no earthquake input, the open loop solution with the earthquake input, or the filter solution with the earthquake input. All three solution quantities in Equation (20), $\boldsymbol{s}_r, \hat{\boldsymbol{s}}_r, \bar{\boldsymbol{s}}_r \in \Re^{i_\mathrm{max} n_\mathrm{max}}$, are vectors that contain the ensemble mean densities of the cells for all time steps. The formulation of the BEEQ as the geometric mean ensures that error amplification and error reduction are balanced (Li and Zhao 2005). A BEEQ of 1 indicates that the error of the estimation state is no different than that of the open loop model and a BEEQ less than 1 indicates that the estimation state has a smaller error than that of the open loop model.

Table 7 shows the BEEQ using each definition of $\hat{\boldsymbol{s}}$ stated previously for earthquakes that result in *medium*, *high*, and *total* damage to the bridge. We use $N_\mathrm{r} = 100$ simulation runs. From the table, it is observed that having traffic sensors (second column) or the earthquake input (third column) gives an improvement over the open loop model with no earthquake input and this is consistent with the graphical solutions (Figures 8(c) and 8(d)) corresponding to these estimators.

Table 7 illustrates several interesting results. For an earthquake that results in *medium* damage to the bridge, the BEEQs imply that knowledge of the earthquake is less beneficial than obtaining data from traffic sensors. For more severe earthquake events, the importance of the earthquake characteristics increases relative to the importance of the traffic sensor data, which is suggested by a decreasing BEEQ when there is knowledge of

Table 7. BEEQ for various estimators for different earthquake scenarios averaged over 100 experiments.

Damage	No earthquake Filter	Earthquake Open loop	Earthquake Filter
medium	0.6296	0.7722	0.3093
high	0.7304	0.5570	0.1932
total	0.8485	0.5341	0.1402

Table 8. BEEQ and running time for using different amounts of ensembles for the estimator that has the earthquake input and traffic sensor data.

Number of ensembles, N	Running time (min/sim)	β	σ_{g}
100	22.4	0.2139	1.0863
150	33.3	0.2002	1.0870
200	45.2	0.1932	1.0687
300	67.1	0.1846	1.0736
500	111.8	0.1824	1.0690
1000	223.1	0.1801	1.0674

the earthquake input. This is due to the fact that the approximate models are initialized in free flow. Thus, the closer the magnitude of the shock wave congestion is to the initial densities, the more accurate the model will be when there is no earthquake input. Note that this trend is specific to the setup of our numerical simulation (e.g. road geometry, initial and boundary conditions, noise models) and might not be observed in different scenarios.

The fourth column of Table 7 suggests that having both the earthquake input and traffic sensor data makes for the best estimator as the BEEQ for all three damage scenarios is 0.30 or less. Put another way, using this estimator results in at least a 70% reduction of error compared to the open loop solution with no earthquake input, which is a significant improvement in performance. Using the BEEQ as a measure of filter performance, the estimator using the earthquake input and traffic sensor data is numerically validated for the real-time prediction of post-disaster traffic conditions.

4.6 *Influence of the number of ensembles*

The performance of the EnKF is dependent on the number of ensembles used. Using too few ensembles might fail to represent the distribution of the state estimate, but using too many could result in very high computational costs. Thus, the performance of the proposed algorithm is also tested with different numbers of ensembles, N. Table 8 shows the running times per simulation, BEEQ, and geometric standard deviations, σ_{g}, for $N_{\mathrm{r}} = 100$ simulations for the estimator that has the earthquake input and traffic sensor data. The assumed damage state of the bridge due to the earthquake is *high*.

The second column shows that the running time per simulation grows approximately linearly with the number of ensembles. Thus, the number of ensembles used in the EnKF is a key factor to reduce computational cost. The third column is the BEEQ computed using Equation (20). The fourth column is the computed geometric standard deviation of the N_{r} error ratios used to compute the BEEQ. Unlike the standard deviation, the geometric standard deviation is a dimensionless, multiplicative factor. Thus, a data set with no geometric variance will give a σ_{g} of 1.

From the table, it is observed that for $N > i_{\max}$ (i.e. the number of ensembles is greater than the number of cells in the network) β has a small variance, suggested by a σ_g near unity. Thus, using $N = 200$ ensembles in Section 4.3 is reasonable for both

accurate and efficient traffic estimation for the traffic network. As N increases, there is a reduction in β, which is intuitive since the more ensembles that are used, the better estimate the filter can give for the traffic states. However, for this particular numerical experiment, for an increasing number of ensembles, the decrease in β is small compared to the increase in computation cost. If the network were to change (i.e. expand), a larger number of ensembles would have to be used or localization approaches (Evensen 2009) are necessary since $i_{\max}$ would also increase.

All simulation results in this work were produced using MATLAB®. The experiments performed in Sections 4.5 and 4.6 were done on a server running four Intel® Xeon® E5620 2.40 GHz CPUs and 24 GB of RAM. From Table 8, a single simulation took 22 minutes for 100 ensembles up to 223 minutes for 1000 ensembles, running on a single core. For efficiency, the code is parallelized at the simulation level (each of four cores runs) when multiple simulations are run to generate the results in Tables 7 and 8.

5. Conclusions

This article illustrated a proof of concept estimation technique for determining the post-disaster traffic behavior due to a reduction of traffic capacity from an earthquake. The standard additive noise model of the model prediction step of sequential estimators is altered using a non-additive model to allow for the conservation of vehicles of the macroscopic cell transmission model. An ensemble Kalman filter is proposed to predict traffic conditions using a stochastic process which draws earthquake characteristics (magnitude, distance) centered around the true earthquake characteristics for each ensemble. Synthetic traffic data using the true state is used to propagate the observation equation in time. Our results show that there are moderate improvements when either the earthquake input or traffic sensor data is integrated into the model. We show that having both of these elements results in a significant improvement in filter performance. The error in the filter is approximately 30% that of the error in the open loop model with no earthquake input across the three earthquake objects we tested.

There are several areas which are open for further exploration. In addition to state estimation, a joint state and parameter estimation problem could result in a better estimate of the model parameters. This joint state and parameter problem is known as *parameter estimation* and requires the state vector to be augmented by the desired parameters.

The fragility model is another aspect that could be expanded. Though we use a univariate model that is only dependent on the peak ground acceleration, there are also bivariate models which make use of fragility surfaces (Huang et al. 2010). Using bivariate fragility surfaces may give a more accurate representation of the probability of being in different damage states.

In terms of a practical application, increasing the network size would result in higher computational costs and scaling techniques such as the distributed local Kalman consensus filter (Sun and Work 2014) or the parallelized filter (Mihaylova et al. 2012) may be required.

The methodology presented in this article assumes that sensor data will be available following an earthquake event. Damage to sensors following an earthquake could result in inaccurate observations and is a potential hurdle to overcome for a practical implementation. However, the intent of this work is to show that if traffic sensors are available and if earthquake data is available, it might provide better information in post disaster

response environments.

Acknowledgment

The authors would like to thank the US National Science Foundation for funding under grant number CMMI 1031318. The third author was supported by the grant (14CCTI-A052531-07-000000) from the Ministry of Land, Infrastructure and Transport of Korean government through the Core Research Institute at Seoul National University for Core Engineering Technology Development of Super Long Span Bridge R&D Center, and the grant Development of cutting edge technologies for the multi-faceted representation of design earthquake ground motions based on analyses of acceleration records (NEMA-ETH-2013-09) from the Earthquake and Tsunami Hazard Mitigation Research Group, National Emergency Management Agency of Korea.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

References

Akiyama, M., Frangopol, D. M., Arai, M., and Koshimura, S. (2013). Reliability of bridges under tsunami hazards: emphasis on the 2011 Tohoku-Oki earthquake. *Earthquake Spectra*, 29(S1):S295–S314.

Anderson, B. and Moore, J. (1979). *Optimal Filtering*. Prentice-Hall.

Bai, J.-W., Hueste, M. B. D., and Gardoni, P. (2009). Probabilistic assessment of structural damage due to earthquakes for buildings in Mid-America. *Journal of Structural Engineering*, 135(10):1155–1163.

Basöz, N. I., Kiremidjian, A. S., King, S. A., and Law, K. H. (1999). Statistical analysis of bridge damage data from the 1994 Northridge, CA, earthquake. *Earthquake Spectra*, 15(1):25–54.

Blandin, S., Couque, A., Bayen, A., and Work, D. (2012). On sequential data assimilation for scalar macrosopic traffic flow models. *Physica D: Nonlinear Phenomena*, 241(17):1421–1440.

Caltrans (2014). Caltrans performance measurement system (PeMS). `http://pems.dot.ca.gov`. Online; accessed 7-April-2014.

Campbell, K. W. (1997). Empirical near-source attenuation relationships for horizontal and vertical components of peak ground acceleration, peak ground velocity, and pseudo-absolute acceleration response spectra. *Seismological Research Letters*, 68(1):154–179.

Chang, S. E. and Nojima, N. (2001). Measuring post-disaster transporation system performance: the 1995 Kobe earthquake in comparative perspective. *Transporation Research Part A: Policy and Practice*, 35(6):475–494.

Chen, H. and Rakha, H. A. (2014). Real-time travel time prediction using particle filtering with a non-explicit state-transition model. *Transportation Research Part C: Emerging Technologies*.

Chen, H., Rakha, H. A., and Sadek, S. (2011). Real-time freeway traffic state prediction: a particle filter approach. In *Proceedings of the 14th IEEE Conference on Intelligent Transportation Systems*, pages 626–631.

Chen, R. and Liu, J. S. (2000). Mixture Kalman filters. *Journal of the Royal Society: Series B (Statistical Methodology)*, 62(3):493–508.

Courant, R., Friedrichs, K., and Lewy, H. (1967). On the partial difference equations of mathematical physics. *IBM Journal of Research and Development*, 11(2):215–234.

Daganzo, C. F. (1994). The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, 28B(4):269–287.

Daganzo, C. F. (1995). The cell transmission model, part II; network traffic. *Transportation Research Part B: Methodological*, 29B(2):79–93.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5):10143–10162.

Evensen, G. (2003). The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53:343–367.

Evensen, G. (2009). *Data Assimilation: The Ensemble Kalman Filter*. Springer.

Garavello, M. and Piccoli, B. (2006). *Traffic Flow on Networks*. American Institute of Mathematical Sciences.

Gazis, D. C. and Knapp, C. H. (1971). On-line estimation of traffic densities from time-series of flow and speed data. *Transportation Science*, 5(3):283–301.

Godunov, S. K. (1959). A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, 89(3):271–306.

Gordon, N. J., Salmond, D. J., and Smith, A. F. M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian estimation. *Radar and Signal Processing, IEE Proceedings F*, 140(2):107–113.

Herrera, J. C. and Bayen, A. M. (2010). Incorporation of Lagrangian measurements in freeway traffic state estimation. *Transporation Research Part B: Methodological*, 44(4):460–481.

Hong, Z. and Fukuda, D. (2012). Effects of traffic sensor location on traffic state estimation. In *Proceedings of the 15th Conference of the Euro Working Group on Transportation*, pages 1–11.

Huang, Q., Gardoni, P., and Hurlebaus, S. (2010). Probabilistic seismic demand models and fragility estimates for reinforced concrete highway bridges with one single-column bent. *Journal of Engineering Mechanics*, 136(11):1340–1353.

Jabari, S. E. and Liu, H. X. (2013). A stochastic model of traffic flow: Gaussian approximation and estimation. *Transporation Research Part B: Methodological*, 47:15–41.

Julier, S. J. and Uhlmann, J. K. (1997). A new extension of the Kalman filter to nonlinear systems. In *Proceedings of the SPIE 11th Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls*, pages 182–193.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45.

Kang, W.-H., Lee, Y.-J., Song, J., and Gencturk, B. (2012). Further development of matrix-based system reliability method and applications to structural systems. *Structure and Infrastructure Engineering*, 8(5):441–457.

Kang, W.-H., Song, J., and Gardoni, P. (2008). Matrix-based system reliability method and applications to bridge networks. *Reliability Engineering and Systen Safety*, 93(11):1584–1593.

Kosa, K. (2012). Damage analysis of bridges affected by tsunami due to great East Japan earthquake. In *Proceedings of the International Symposium on Engineering Lessons*

Learned from the 2011 Great East Japan Earthquake, pages 1386–1397.

Lee, Y.-J., Song, J., Gardoni, P., and Lim, H.-W. (2011). Post-hazard flow capacity of bridge transportation network considering structural deterioration of bridges. *Structure and Infrastructure Engineering*, 7(7-8):509–521.

Li, X. R. and Zhao, Z. (2005). Relative error measures for evaluation of estimation algorithms. In *Proceedings of the 7th IEEE International Conference on Information Fusion*, pages 211–218.

Lighthill, M. J. and Whitham, G. B. (1955). On kinematic waves II. a theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 229(1178):317–345.

Mackie, K. R. and Stojadinović, B. (2006). Post-earthquake functionality of highway overpass bridges. *Earthquake Engineering and Structural Dynamics*, 35:77–93.

Mandel, J., Cobb, L., and Beezley, J. D. (2011). On the convergence of the ensemble Kalman filter. *Applications of Mathematics*, 56:533–541.

Mihaylova, L., Boel, R., and Hegyi, A. (2006). An unscented Kalman filter for freeway traffic estimation. In *Proceedings of the IFAC Symposium of Control in Transportation Systems*, pages 31–36.

Mihaylova, L., Boel, R., and Hegyi, A. (2007). Freeway traffic estimation within particle filtering framework. *Automatica*, 43(2):290–300.

Mihaylova, L., Hegyi, A., Gning, A., and Boel, R. K. (2012). Parallelized particle and Gaussian sum particle filters for large-scale freeway traffic systems. *IEEE Transactions on Intelligent Transportation Systems*, 13(1):36–48.

Murachi, Y., Orikowski, M. J., Dong, X., and Shinozuka, M. (2003). Fragility analysis of transportation networks. In *Proceedings of the SPIE Conference on Smart Structures, Materials, and Nondestructive Evaluation for Civil Infrastructures*, pages 655–663.

Nakanishi, H., Matsuo, K., and Black, J. (2013). Transportation planning methodologies for the post-disaster recovery in regional communities: the East Japan earthquke and tsunami 2011. *Journal of Transport Geography*, 31:181–191.

Newell, G. F. (1993). A simplified theory of kinematic waves in highway traffic, part II: queueing at freeway bottlenecks. *Transportation Research Part B: Methodological*, 27B(4):289–303.

Nielson, B. G. and DesRoches, R. (2007). Analytical seismic fragility curves for typical bridges in the Central and Southeastern United States. *Earthquake Spectra*, 23(3):615–633.

Nojima, N. and Sugito, M. (2000). Simulation and evaluation of post-earthquake functional performance of transportation network. In *Proceedings of the 12th World Conference on Earthquake Engineering*.

Richards, P. I. (1956). Shock waves on the highway. *Operations Research*, 4(1):42–51.

Rosenblueth, E. and Esteva, L. (1972). Reliability basis for some Mexican codes. In *Probabilistic Design of Reinforced Concrete Buildings*. American Concrete Institute.

Schanack, F., Valdebenito, G., and Alvial, J. (2012). Seismic damage to bridges during the 27 February 2010 magnitude 8.8 Chile earthquake. *Earthquake Spectra*, 28(1):301–315.

Shizunoka, M., Murachi, Y., Dong, X., Zhou, Y., and Orlikowski, M. J. (2003). Seismic performance of highway transportation networks. In *Proceedings of the China-US Workshop on Protection of Urban Infrastructure and Public Buildings against Earthquakes and Manmade Disasters*.

Simon, D. (2006). *Optimal State Estimation: Kalman, H-Infinity, and Nonlinear Ap-*

proaches. Willy-Interscience.

Song, J. and Kang, W.-H. (2009). System reliability and sensitivity under statistical dependence by matrix-based system reliability method. *Structural Safety*, 31(2):148–156.

Sun, X., Muñoz, L., and Horowitz, R. (2004). Mixture Kalman filter based highway congestion mode and vehicle density estimator and its application. In *Proceedings of the American Control Conference*, pages 2098–2103.

Sun, Y. and Work, D. B. (2014). A distributed local Kalman consensus filter for traffic estimation. In *Proceedings of the IEEE Conference on Decision and Control*.

Szeto, M. W. and Gazis, D. C. (1972). Application of Kalman filtering to the surveillance and control of traffic systems. *Transportation Science*, 6(4):419–439.

van Hinsbergen, C. P. I. J., Schreiter, T., Zuurbier, F. S., van Lint, J. W. C., and van Zuylen, H. J. (2012). Localized extended Kalman filter for scalable real-time traffic state estimation. *IEEE Transcations on Intelligent Transportation Systems*, 13(1):385–394.

Wald, D. J., Lin, K.-W., Porter, K., and Turner, L. L. (2008). ShakeCast: automating and improving the use of ShakeMap for post-earthquake decision-making and response. *Earthquake Spectra*, 24(2):533–553.

Wang, R. and Work, D. B. (2014). Interactive multiple model ensemble Kalman filter for traffic estimation and incident detection. In *Proceedings of the IEEE Conference on Intelligent Transporation Systems*.

Wang, Y. and Papageorgiou, M. (2005). Real-time freeway traffic state estimation based on extended Kalman filter: a general approach. *Transportation Research Part B: Methodological*, 39(2):141–167.

Work, D. B., Blandin, S., Tossavainen, O.-P., Piccoli, B., and Bayen, A. M. (2010). A traffic model for velocity data assimilation. *Applied Mathemtics Research eXpress*, 1(1):1–35.