ESTIMATING POST-DISASTER TRAFFIC CONDITIONS USING REAL-TIME DATA STREAMS

BY

REECE POMAIKAI OTSUKA

THESIS

Submitted in partial fulfilment of the requirements
for the degree of Master of Science in Civil Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Advisers:

Associate Professor Junho Song
Assistant Professor Daniel Work

ABSTRACT

This thesis 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 through numerical experiments and the results show that integrating the seismic hazard and bridge fragility model with the traffic model and traffic sensor data improves the post-disaster traffic estimate. The supporting source code and data used in this thesis are available for download at `https://github.com/rotsuka/UIUCthesis`.

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisers, Professor Junho Song and Professor Daniel Work, for giving me the opportunity to be a part of this multidisciplinary project and for the guidance and support they have given me along the way, academically and personally. They are both individuals who live up to high standards and who have pushed me to go beyond my limits. I feel truly honored to have worked with both of them and been a part of their research groups.

I also would like to thank my parents, who have given me their unconditional love and unwavering support from day one. They have been my linchpin when I faced rough times throughout my life and they taught me the value of perseverance, dedication, and hard work.

Contents

List of Tables

List of Figures

Chapter 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 could 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 recovery efforts. If accurate traffic estimates can be calculated immediately following an earthquake, it could be very helpful in expediting post-hazard rescue and recovery efforts. The problem addressed in this thesis is to estimate post-disaster traffic conditions using measurements from traffic sensors and information about the earthquake's characteristics and bridge fragility. In this work, the magnitude of the earthquake and its proximity to the bridge corresponds to changes in bridge traffic capacity, affecting the traffic conditions of the network.

To this end, the post-diaster traffic estimation problem is posed as a sequential state estimation problem as follows. Suppose the traffic evolution equations are

constructed from a scalar 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 (Simon, 2006) are given respectively as

$$\boldsymbol{x}_n = \boldsymbol{f}\left(\boldsymbol{x}_{n-1}, \boldsymbol{\theta}, \boldsymbol{w}_n\right) \tag{1.1}$$

$$\boldsymbol{y}_n = \boldsymbol{h}_n\left(\boldsymbol{x}_n\right) + \boldsymbol{v}_n, \tag{1.2}$$

where $\boldsymbol{\theta}$ is a time-invariant vector which contains the parameters of road such as the maximum traffic capacity, the speed limit, and the maximum backward propagating wave speed. 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. Note that the model noise here is not additive, which allows us to keep the mass conservation of $\boldsymbol{f}$.

Given the model evolution (Equation (1.1)) and observation (Equation (1.2)) equations, the post-disaster traffic estimation problem can be posed as estimating the traffic state $\boldsymbol{x}_n$ given measurements $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_n$. However, following an earthquake, an online implementation would not be able to detect any discrepancies in the network until the congestion generated by the capacity reduction is detected by the traffic sensor. Additionally, dedicated traffic sensors may be sparsely distributed throughout the network due to the high cost of installation and maintenance. Thus, to improve the estimation performance, we augment the model evolution equation by feeding in real-time information about the earthquake immediately after the occurrence at time step n_{EQ}, and every time step thereafter. We then rewrite Equation (1.1) as

$$\boldsymbol{x}_n = \boldsymbol{f}\left(\boldsymbol{x}_{n-1}, \boldsymbol{\theta}, u_n, \boldsymbol{w}_n\right), \tag{1.3}$$

and u_n is the *earthquake input factor* defined as

$$u_n = \begin{cases} 1 & \text{if } n < n_{\text{EQ}} \\ \alpha & \text{if } n \geq n_{\text{EQ}}, \end{cases} \tag{1.4}$$

where α is the *traffic capacity ratio*, $0 \leq \alpha \leq 1$. Though the lane parameters $\boldsymbol{\theta}$ are time-invariant, u_n is multiplied by the element in $\boldsymbol{\theta}$ that corresponds to the maximum traffic capacity of the bridge to allow for the fact that the roadway capacity may change following an earthquake. Due to the nonlinearity of the traffic model, we propose to use an ensemble Kalman filter to solve the sequential state estimation problem.

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

With the number of natural disaster occurrences worldwide, researchers are focusing 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-diaster 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), software 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 aid in understanding 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 attempts have been successful in determining failures of bridges in networks and changes in flow capacity, they do not explicitly model the traffic dynamics.

Sequential estimation in traffic applications dates back several decades to the 1970s when Gazis and Knapp (1971) used the *Kalman filter* (KF) (Kalman, 1960) and Szeto and Gazis (1972) used the *extended Kalman filter* (EKF) (Anderson and Moore, 1979) 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) and Wang and Papageorgiou (2005). 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, 2009; 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. 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 thesis

The main contribution of this thesis is the design and analysis of an estimation framework to determine the post-diaster 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 1.1 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 input 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, creating an iterative process.

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

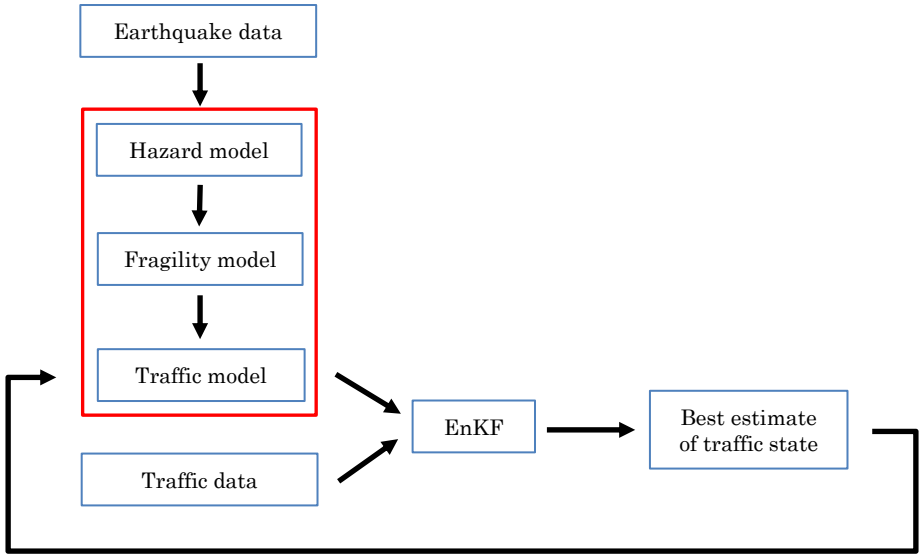


Figure 1.1: Proposed model framework: The upper left boxed portion shows the three dynamic models our framework utilizes. Earthquake data is used as an input into the forward model, which consists of seismic hazard, bridge fragility, and traffic evolution models. These are integrated with traffic sensor data as inputs into the filter. The EnKF runs an error variance minimizing scheme to produce the best estimate of the traffic state. The traffic states are fed into the forward model at the next time step.

Chapter 2

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

This chapter 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 (1.3)).

2.1 Macroscopic traffic model

The *Lighthill-Whitham-Richards* (LWR) *partial differential equation* (PDE) (Lighthill and Whitham, 1955) is a conservation law used to model the evolution of the *traffic density*, ρ, which is a measure of the number of vehicles on a given length of road (veh/km). It is a continuum model that describes the conservation of vehicles:

$$\frac{\partial \rho}{\partial t} + \frac{\partial Q(\rho)}{\partial x} = 0, \tag{2.1}$$

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

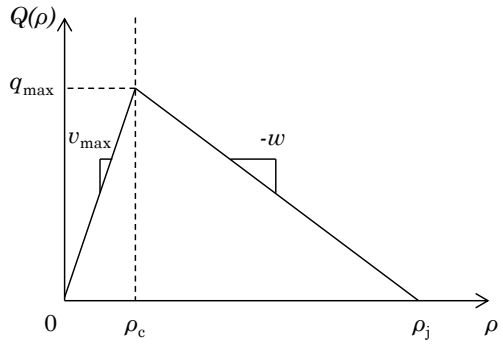


Figure 2.1: Newell-Daganzo flux function: Densities in the range $\rho \in [0, \rho_c]$ are in the free flow region and densities in the range $\rho \in [\rho_c, \rho_j]$ are in the congested region.

$v = V(\rho)$ as follows:

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

One widely used fundamental diagram is the triangular *Newell-Daganzo* flux function (Daganzo, 1995; Newell, 1993), shown in Figure 2.1. It is a piecewise linear function of the density, with different slopes in the free-flow and congestion regions:

$$Q(\rho) = \begin{cases} \rho v_{\max} & \text{if } \rho \in [0, \rho_c] \\ w\,(\rho_j - \rho) & \text{if } \rho \in [\rho_c, \rho_j], \end{cases} \tag{2.2}$$

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 $w = v_{\max}\rho_c/(\rho_j - \rho_c)$. From Equation (2.2), it is observed that the three variables $q_{\max}, v_{\max}$, and w are adequate to define the triangular fundamental diagram (Figure 2.1).

For numerical implementation, Equation (2.1) 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

$$\rho_{n+1}^{i} = \rho_{n}^{i} + \frac{\Delta T}{\Delta x}\left(q\left(\rho_{n}^{i-1}, \rho_{n}^{i}\right) - q\left(\rho_{n}^{i}, \rho_{n}^{i+1}\right)\right), \tag{2.3}$$

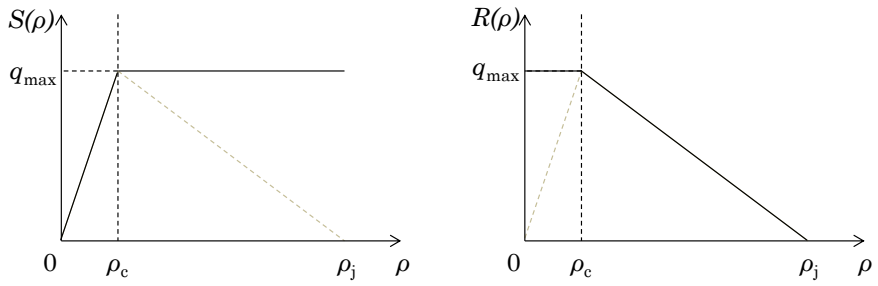
where the numerical flux $q(\cdot, \cdot)$ is a function used to compute flow across the boundary of two adjacent cells. Equation (2.3) is commonly referred to as the *cell-transmission model* (CTM) (Daganzo, 1994, 1995) for triangular fundamental diagrams. The flow, $q(\cdot, \cdot)$, can be computed as

$$q\left(\rho^{\mathrm{u}}, \rho^{\mathrm{d}}\right) = \min\left\{S\left(\rho^{\mathrm{u}}\right), R\left(\rho^{\mathrm{d}}\right)\right\}, \tag{2.4}$$

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^{u} and ρ^{d} represent the densities of the upstream and downstream cells. The sending and receiving functions are constructed from the fundamental diagram as shown in Figure 2.2. 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 2.2a). 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 $\rho^{\mathrm{d}} < \rho_{\mathrm{c}}$. If the density continues to increase, the flow that can be received decreases (Figure 2.2b).

2.1.1 Non-additive noise model

As mentioned previously, the model evolution equation does not have additive noise. The noise is instead embedded in the fundamental diagram and thus directly affects



(a) Sending function (b) Receiving function

Figure 2.2: Sending and receiving functions of Newell-Daganzo flux: 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.

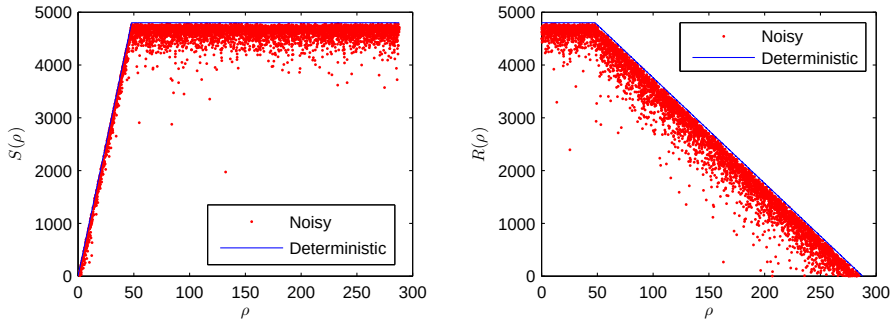
the sending and receiving functions. We treat the fundamental diagram as an upper bound on the flow that can be sent across a cell boundary at any given time. A lognormal distribution is used for the noise and it is subtracted from the fundamental diagram to generate more realistic functions consistent with experimental data from inductive loops (Caltrans, 2014):

$$\hat{S}(\rho) = \begin{cases} S(\rho) - \varepsilon_S & \text{if } \rho \in [0, \rho_c] \\ S(\rho) - \varepsilon_Q & \text{if } \rho \in [\rho_c, \rho_j], \end{cases} \tag{2.5}$$

and

$$\hat{R}(\rho) = \begin{cases} R(\rho) - \varepsilon_Q & \text{if } \rho \in [0, \rho_c] \\ R(\rho) - \varepsilon_R & \text{if } \rho \in [\rho_c, \rho_j], \end{cases} \tag{2.6}$$

where $\varepsilon_S \sim \mathcal{LN}(\lambda_S, \xi_S^2)$ and $\varepsilon_R \sim \mathcal{LN}(\lambda_R, \xi_R^2)$ are noises associated with the free flow region of the sending function and the congested region of the receiving function, respectively, and λ and ξ are the respective location and scale parameters. We introduce a third noise, $\varepsilon_Q \sim \mathcal{LN}(\lambda_Q, \xi_Q^2)$ to represent the noise on the maximum flow. The noise on the receiving function is greater than that on the sending function.



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

Figure 2.3: Noisy sending and receiving functions: Using the sending and receiving functions derived from the fundamental diagram as an upper bound, logarithmic noise is subtracted to obtain the flow. There are 5000 sample points on each graph to illustrate the relative magnitudes of the means and variances associated with each function.

An example of noisy sending and receiving functions is shown in Figure 2.3. The solid line represents the deterministic sending and receiving functions.

Note that for densities near zero or the jam density, Equations (2.5) or (2.6) 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 (2.5) or (2.6), and 0. Thus, we rewrite Equation (2.4) as

$$q\left(\rho^{\mathrm{u}}, \rho^{\mathrm{d}}\right) = \max\left\{\min\left\{\hat{S}\left(\rho^{\mathrm{u}}\right), \hat{R}\left(\rho^{\mathrm{d}}\right)\right\}, 0\right\}. \tag{2.7}$$

2.2 Seismic hazard and fragility models

Two main characteristics of an earthquake 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 probabilities of a bridge being in different damage states are determined. From knowing the damage states, the percentage of maximum flow capacity that is allowed for each bridge based on a probabilistic

draw of the bridge damage distribution is 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 (soil profile), and the fault type. PGA is related to the mentioned factors by attenuation relationships. In this work, we use the attenuation relationship by Campbell (1997) as an example:

$$\begin{aligned}\ln A_{\mathrm{H}} = &-3.512 + 0.904M - 1.328 \ln \sqrt{r^2 + \left(0.149 \exp\left(0.647M\right)\right)^2} \\ &+ (0.440 - 0.171 \ln r) s_{\mathrm{sr}} + (0.405 - 0.222 \ln r) s_{\mathrm{hr}} \\ &+ (1.125 - 0.112 \ln r - 0.0957M) F_{\mathrm{t}},\end{aligned} \tag{2.8}$$

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. We can write Equation (2.8) as follows:

$$A_{\mathrm{H}} = f_{\mathrm{e}}\left(M, r, s_{\mathrm{sr}}, s_{\mathrm{hr}}, F_{\mathrm{t}}\right) \tag{2.9}$$

For simplicity, we assume that we have a slip-strike fault ($F_{\mathrm{t}} = 0$) and alluvium or firm soil ($s_{\mathrm{sr}} = s_{\mathrm{hr}} = 0$). For simplicity, but without loss of generality, in this work A_{H} is assumed to depend only on the magnitude and the distance from the structure to the epicenter, which are both dependent on intrinsic properties of the earthquake.

2.2.2 Fragility curves

Fragility curves define the conditional probability of exceeding certain *limit states* (LS) as a function of a selected *intensity measure* (IM) of ground motions. Fragility

is often modeled by a lognormal cumulative distribution function, where the *demand* D and *capacity* C are assumed to be lognormally distributed variables. The safety factor F (Rosenblueth and Esteva, 1972) can then be computed as

$$F = \ln C - \ln D, \tag{2.10}$$

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 $F \sim \mathcal{N}(\mu_F, \sigma_F^2)$, let the standard normal variable $Z = (F - \mu_F)/\sigma_F$. The conditional probability of exceedance given an intensity measure is given as

$$\begin{aligned} P(F \leq 0|IM) &= P\left(Z \leq -\frac{\mu_F}{\sigma_F}\middle| IM\right) \\ &= \Phi\left(-\frac{\mu_F(IM)}{\sigma_F(IM)}\right), \end{aligned} \tag{2.11}$$

where $\Phi(\cdot)$ is the normal *cumulative distribution function* (CDF). Since there are multiple limit states possible (based on different levels of drift, peak ground parameters, etc.), the conditional probability needs to be 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}. As a result, we can write Equation (2.11) in the following form (Nielson and DesRoches, 2007):

$$P(F_i \leq 0|A_{\mathrm{H}}) = \Phi\left(\frac{\ln A_{\mathrm{H}} - \ln(\tilde{x}_i)}{\zeta_i}\right), \tag{2.12}$$

where F_i is the failure event, $\tilde{x}_i$ is the median capacity of the structure in PGA, and ζ_i is the dispersion value belonging to the ith limit state. Table 2.1 provides the fragility curve parameters for several bridge classes typical to the Central and Southeastern United States, excerpted from 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

Table 2.1: Median and dispersion values in g for seismic fragility curves of bridges

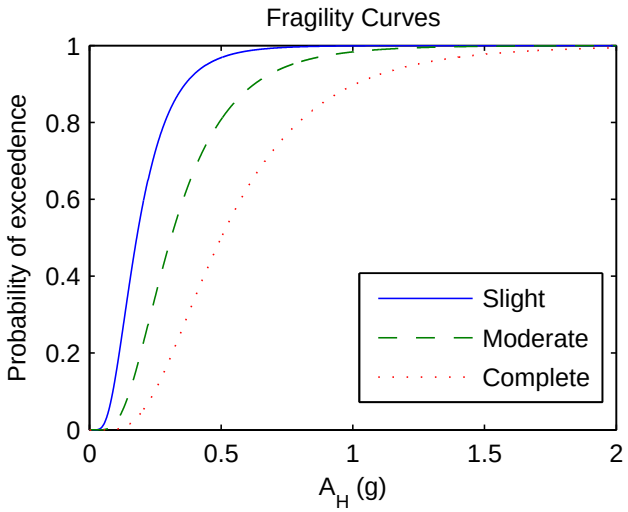


Figure 2.4: Fragility model for MSC steel bridges: The fragility curves represent the three limit states.

For simplicity, we assume three limit states, *slight*, *moderate*, and *complete*, as shown in Table 2.1 for *multispan continuous* (MSC) steel girder bridges (Nielson and DesRoches, 2007). The fragility curve corresponding to each limit state is given in Figure 2.4. The probability of exceedance of a less severe limit state is greater than the probability of exceedance of a more severe limit state. Also, as A_{H} increases the probability of exceedance of any limit state increases monotonically. Both of these observations are intuitive considering the mathematical formulation of the fragility curves.

2.2.3 Creating probability of damage states tables

Once the fragility curves have been modeled, it is possible to determine the probability of being in different damage states by looking at the difference between fragility curves for a given A_{H}. The three limit states stated 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: P_{S} (*slight*), P_{M} (*moderate*), and P_{C} (*complete*), which are listed in increasing severity. The second column of Table 2.2 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:

$$\sum_{d \in \mathcal{D}} P\left(x_d | A_{\mathrm{H}}\right) = 1, \tag{2.13}$$

where x_d is the actualization of a damage state. 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.

2.2.4 Mapping damage states to traffic capacity

For numerical simulation purposes, 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 2.2 shows the mapping between the damage state and traffic capacity ratio. Note that this mapping is used as an example and more detailed models could be constructed.

Damage State	Probability	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

Table 2.2: Probability of being in different damage states and their associated traffic capacity ratios, α

2.2.5 Fragility model as an input to the traffic model

At a given time $n \geq n_{\mathrm{EQ}}$, we can compute the probability distribution of the damage states (Chapter 2.2.3) using the earthquake magnitude and location. This distribution of damage states maps to a distribution of traffic capacity (Chapter 2.2.4) which can be included explicitly as an input into the CTM (Equation (1.3)). In this way, the existence of an earthquake and its subsequent damage to the transportation infrastructure has a direct influence on the maximum flow, $q_{\max}$, at a given instance in time.

Chapter 3

Sequential state estimation

This chapter briefly reviews sequential state estimation. It also introduces the Kalman filter and ensemble Kalman filter sequential state estimation algorithms used in this work. The observation equation (Equation (1.2)) is also defined.

3.1 Kalman filter

When the model evolution and observation equations are linear with additive white noise, Equations (1.1) and (1.2) can be written as the following:

$$\boldsymbol{x}_n = \mathbf{F}_n \boldsymbol{x}_{n-1} + \boldsymbol{w}_n \tag{3.1}$$

$$\boldsymbol{y}_n = \mathbf{H}_n \boldsymbol{x}_n + \boldsymbol{v}_n, \tag{3.2}$$

where $\mathbf{F}_n$ is the linear model operator and $\mathbf{H}_n$ is the modeled 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 KF (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

$$\text{Prediction:} \begin{cases} \boldsymbol{x}_{n|n-1} = \mathbf{F}_n \boldsymbol{x}_{n-1|n-1} \\ \mathbf{P}_{n|n-1} = \mathbf{F}_n \mathbf{P}_{n-1|n-1} \mathbf{F}_n + \mathbf{W}_n \end{cases} \tag{3.3}$$

$$\text{Update:} \begin{cases} \mathbf{K}_n = \mathbf{P}_{n|n-1} \mathbf{H}_n^{\mathrm{T}} \left(\mathbf{H}_n \mathbf{P}_{n|n-1} \mathbf{H}_n^{\mathrm{T}} + \mathbf{R}_n \right)^{-1} \\ \boldsymbol{x}_{n|n} = \boldsymbol{x}_{n|n-1} + \mathbf{K}_n \left(\boldsymbol{y}_n - \mathbf{H}_n \boldsymbol{x}_{n|n-1} \right) \\ \mathbf{P}_{n|n} = \mathbf{P}_{n|n-1} - \mathbf{K}_n \mathbf{H}_n \mathbf{P}_{n|n-1}. \end{cases} \tag{3.4}$$

The model prediction step (Equation (3.3)) 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 (3.4)) 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).

Due to the nonlinearities of the CTM (Equation (2.3)), a nonlinear extension of the KF, the EnKF, is used in this work. The EnKF is a Monte Carlo approximation of the KF, which represents the prior and posterior distributions by randomly generated ensembles. Rather than obtaining the covariance matrix through $\mathbf{F}_n$ (Equation (3.3)), the EnKF evolves the covariance matrix indirectly through the individual ensemble evolutions of the prior distribution. The EnKF is described in the following section.

3.2 Ensemble Kalman filter

The ensemble representations of the model states and observations indexed by $e \in \{1, \cdots, N\}$ are given as

$$\boldsymbol{x}_n^e = \boldsymbol{f}\left(\boldsymbol{x}_{n-1}^e, \boldsymbol{\theta}, u_n^e, \boldsymbol{w}_n^e\right) \tag{3.5}$$

$$\boldsymbol{y}_n^e = \boldsymbol{y}_n + \boldsymbol{v}_n^e, \tag{3.6}$$

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 variable u_n^e is 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 (Mandel et al., 2011). The model prediction and measurement update steps for the EnKF are given as

$$\text{Prediction:} \begin{cases} \boldsymbol{x}_{n|n-1}^e = \boldsymbol{f}\left(\boldsymbol{x}_{n-1|n-1}^e, \boldsymbol{\theta}, u_n^e, \boldsymbol{w}_n^e\right) \\ \boldsymbol{x}_{n|n-1} = \frac{1}{N}\sum_{e=1}^{N} \boldsymbol{x}_{n|n-1}^e \\ \mathbf{P}_{n|n-1} = \frac{1}{N-1}\sum_{e=1}^{N}\left(\boldsymbol{x}_{n|n-1}^e - \boldsymbol{x}_{n|n-1}\right)\left(\boldsymbol{x}_{n|n-1}^e - \boldsymbol{x}_{n|n-1}\right)^{\mathrm{T}} \end{cases} \tag{3.7}$$

$$\text{Update:} \begin{cases} \mathbf{K}_n = \mathbf{P}_{n|n-1}\mathbf{H}_n^{\mathrm{T}}\left(\mathbf{H}_n\mathbf{P}_{n|n-1}\mathbf{H}_n^{\mathrm{T}} + \mathbf{R}_n\right)^{-1} \\ \boldsymbol{x}_{n|n}^e = \boldsymbol{x}_{n|n-1}^e + \mathbf{K}_n\left(\boldsymbol{y}_n^e - \mathbf{H}_n\boldsymbol{x}_{n|n-1}^e\right) \\ \mathbf{P}_{n|n} = \mathbf{P}_{n|n-1} - \mathbf{K}_n\mathbf{H}_n\mathbf{P}_{n|n-1}. \end{cases} \tag{3.8}$$

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 (3.8)), a posterior ensemble, $\boldsymbol{x}_{n|n}^e$ is computed using the prior ensemble (Equation (3.7)) and the ensemble observation (Equation (3.6)). 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 (3.4)).

Note that Equations (3.7) and (3.8) illustrate the structural similarities to the linear KF (Equations (3.3) and (3.4)). For the numerical experiments shown next, we use a more efficient implementation described in Evensen (2009). Rather than obtaining each ensemble state vector recursively, a matrix form of the analysis equation for fast implementation is employed to reduce the computational cost of the algorithm. A detailed explanation of the matrix scheme and algorithm is given in Evensen (2009).

3.3 Observation equation

In our model, 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 (1.2) that relates the traffic state $\boldsymbol{x}_n = (\rho_n^0, \cdots, \rho_n^{i_{\max}})^{\mathrm{T}}$ at a time n to the sensor measurements must be defined. The observation operator $\boldsymbol{h}$ is defined as follows:

$$\boldsymbol{h}_n(\boldsymbol{x}_n) = \mathbf{H}\boldsymbol{x}_n, \tag{3.9}$$

where $\mathbf{H}$ is the measurement operator which indicates locations where sensor measurements are obtained. Note that $\mathbf{H}$ is time-invariant because the sensors always produce measurements and their positions are fixed in the following numerical experiments.

For traffic sensors indexed by $l \in \{1, \cdots, l_{\max}\}$, the measurement error covariance matrix, $\mathbf{R}_n$, is given as

$$\mathbf{R}_n = \mathrm{diag}\left(\sigma_{n,1}^2, \cdots, \sigma_{n,l_{\max}}^2\right) \tag{3.10}$$

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

Chapter 4

Numerical experiments

In this chapter, we describe 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 simple transportation network is constructed and the macroscopic traffic model parameters on each link are described. Using prescribed initial and boundary conditions, the forward model (Equation (2.3)) is run once. This *true model* has a prescribed damage state associated with the bridge, which allows the evolution of the traffic state over time to be solved deterministically, resulting in the *true solution* to be estimated. In addition to the true model, there is an *approximate model*, which has its own set of initial and boundary conditions and is used in the traffic estimation algorithm. The approximate model may or may not have the earthquake input (u) and may or may not use the EnKF algorithm and traffic sensor data ($\boldsymbol{y}$). Approximate models that do not account for traffic sensor data are *open loop* models and those that do account for traffic sensor data are *filters*. The approximate models, regardless if they are open loop or filters, provide *estimation solutions*. The true solution serves to assess the accuracy of the estimation solutions and it is also used to generate the synthetic measurements in the filter. To quantify performance of the filter and to compare improvements from one model to others (e.g. filter vs. open loop and earthquake vs. no earthquake), we use the *Bayesian Estimation Error Quotient* (BEEQ) (Li

Link	1	2	3	4
Length (km)	2	1	0.5	1
Cells	14	7	3	7
Δx (km)	0.143	0.143	0.167	0.143
Lanes	3	2	2	2
$v_{\max}$ (km/h)	100			
$q_{\max}$ (veh/h/lane)	2400			
w (km/h)	20			
ρ_{j} (veh/km/lane)	144			

Table 4.1: Macroscopic traffic model parameters for the true and approximate models

and Zhao, 2005). The BEEQ is also used to compare filter performance against computational cost when different numbers of ensembles are used.

4.1 Parameters of the macroscopic traffic model

As shown in Figure 4.1, we simulate a 4.5 km stretch of roadway which contains four *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, $\boldsymbol{x}$.

We assume the parameters for each link between the true and approximate models are the same and are given in Table 4.1. The transition from Link 1 to Link 2 is a merge from three lanes to two, which creates a bottleneck. The simulation is run for a duration ($t_{\max}$) of 40 minutes (2400 seconds) with a time step (ΔT) of 5 seconds, resulting in 481 ($n_{\max}$) time steps starting at $t = 0$ seconds. Since the speed limit, $v_{\max}$, does not vary between links, each link is discretized into cells of approximately equal length as shown in Table 4.1. This discretization obeys the *Courant-Friedrichs-Lewy* (CFL) condition (Courant et al., 1967), which is required for numerical stability of the CTM (Equation (2.3)). The fundamental diagram (Figure 2.1) is computed using $v_{\max}$, $q_{\max}$, and w.

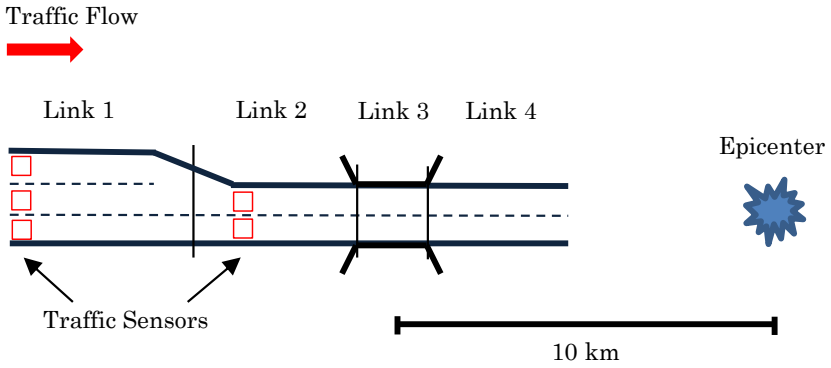


Figure 4.1: Schematic of model simulation: The roadway is composed of four links of varying length and number of cells. Squares denote locations of inductive loop detectors and the earthquake epicenter is located 10 km downstream from the center of Link 3, the bridge.

Link 3 represents a bridge susceptible to damage from an earthquake, as shown in Figure 4.1. Inductive loop detector sensors are located in cells 1 and 15 of the discretized model and traffic flows from left to right. Although Figure 4.1 shows a transition from three lanes to two lanes, the CTM computationally treats this transition as a point.

4.2 True model assumptions

In addition to earthquakes which cause insignificant damage (i.e. flow is uninterrupted), there are three main types of earthquakes for the *true model* which are outlined in Table 4.2. When the values for M and r from Table 4.2 are substituted into Equation (2.9), the resulting PGA value when mapped to the fragility model (Figure 2.4) results in the damage state defined by those M and r values to have the highest probability of occurrence. The third column shows the earthquake is located at the coordinates (13.25,0). This is due to the upstream end of Link 1 being initialized at (0,0), which, in turn, results in the center of Link 3 being located at (3.25,0). Thus, the distance between the bridge center and the earthquake epicenter

Damage	Magnitude, M	Epicenter (x, y)	Distance, r (km)
medium	6.0	(13.25,0)	10
high	7.0	(13.25,0)	10
total	8.0	(13.25,0)	10

Table 4.2: Earthquake characteristics

is 10 km, as shown in Figure 4.1 and confirmed in Table 4.2.

We initialize the model with the traffic state in free flow and the *upstream and downstream boundary conditions* (UBC, DBC) are both in free flow. The initial and boundary conditions are given in the second column of Table 4.4. At $t = 10$ minutes (600 seconds), an earthquake occurs. We draw a realization of $q_{\max}$ from the damage state distribution and use this to propagate the forward deterministic model in Equation (2.3). This is the true solution with which we compare the accuracy of our estimates. The density contour in the time-space domain is shown in Figure 4.2a for an earthquake which resulted in a 50% reduction in traffic capacity of the bridge (i.e. *high* damage). The true solution will be analyzed further in subsequent sections.

4.3 Approximate model assumptions

In the *approximate model*, we draw earthquake characteristics for each ensemble at each time step from normal distributions whose means are the true characteristics of the earthquake (Table 4.2). We assume we are able to get relatively accurate details of the earthquake such as the magnitude and location once the earthquake occurs. Therefore, the variance of the distribution we assign to the noises is fairly small. Table 4.3 shows the standard deviations of the earthquake parameters $(\sigma_M, \sigma_x, \sigma_y)$ as well as the other noise parameters used in the approximate model. In a real world implementation, the traffic measurements would be taken directly from the sensors. However, in numerical experiments, the measurements are synthesized by computing

σ_M	0.3
σ_x (km)	5
σ_y (km)	5
σ_{sensor} (veh/km)	5
ε_S (veh/h)	$\mathcal{LN}(4.82, 0.61^2)$
ε_R (veh/h)	$\mathcal{LN}(5.44, 0.73^2)$
ε_Q (veh/h)	$\mathcal{LN}(5.08, 0.67^2)$

Table 4.3: Noise model parameters of the approximate model

Link	True	Approximate	Absolute error (%)
1	40	20	50
2	35	20	42.9
3	35	30	16.7
4	35	30	16.7
UBC	60	40	33.3
DBC	20	20	0

Table 4.4: Initial and boundary conditions for the true and approximate models in veh/km

the true solution and adding the measurement noise, $\boldsymbol{v}_n$. We further assume that the each traffic sensor has the same accuracy and the accuracy of the traffic sensors do not change after the occurrence of the earthquake. Thus, each entry in Equation (3.10) is defined with a constant standard deviation, σ_{sensor}. The noise associated with the sending and receiving functions (Equations (2.5) and (2.6)) are also shown in Table 4.3.

In order to test the efficacy of our model, we initialize the approximate model in free flow with both boundary conditions also in free flow, with values as given in the third column of Table 4.4. The table also illustrates the absolute relative error of the approximate model initial and boundary conditions compared to those of the true model. The values in Table 4.4 show that the approximate model is initialized with less 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. The initial values of the approximate model

are chosen such that if an ensemble predicts an earthquake whose resulting damage state is that of the true model, a queue will start to form from the end of Link 2. We use $N = 100$ ensembles for the approximate model.

4.4 Results and discussion

As mentioned previously, whether the cell is in free flow or congestion depends on the value of the critical density, ρ_c, which in this simulation is 24 veh/km/lane for both the true and approximate models. In the true solution (Figure 4.2a), congestion occurs near the start of the simulation ($t \approx 0$) at cell 14, before the occurrence of the earthquake. It occurs at the boundary of Link 1 and Link 2 and is due to the fact that there is a merge from three lanes to two lanes. The outflow from Link 1 exceeds the capacity of Link 2, thus congestion forms on Link 1 and a shock wave propagates backward on the link. At time $t = 10$ minutes (600 seconds), the earthquake occurs and damages the bridge (Link 3) such that its traffic capacity is reduced by 50% ($\alpha = 0.5$). This results in another backward propagating shock wave which starts at the end of Link 2. Initially, the congestion caused by the reduction in traffic capacity is less severe than the congestion caused by the merge, which is represented by the differences in magnitude and also by the differing wave speeds. When the second backward propagating congestion wave reaches the merge, the congestion severely increases, as shown by the instantaneous transition from the light blue congestion region to the orange heavy congestion region. This is a result of the fact that the initial congestion is combined with the congestion resulting from the earthquake damage to the bridge. Link 4 remains in free flow. The density of vehicles in Link 4 reduces because the outflow from the bridge is decreased due to the damage.

The subplots of Figure 4.2 show the various estimation results with or without the earthquake input and/or with or without traffic sensor data. As a baseline of

the estimator, an open loop model with no knowledge of the earthquake input is used (Figure 4.2b). Figures 4.2c and 4.2d show the estimation results when there is either traffic data or knowledge of the earthquake input, but not both. Figure 4.2e 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 4.2b), the traffic conditions will continue to be in free flow. Having no earthquake input into the model is equivalent to having no earthquake event or one that only results in *insignificant* damage to the bridge, thus no disruption in traffic flow. Since traffic sensor data is not used to correct the model (open loop) with knowledge about the congestion in the true solution (Figure 4.2a), there is also no detection of the initial congestion region at the merge. When there is no earthquake input, but there is information fed to the filter via traffic sensors (Figure 4.2c), the correct amount of congestion is detected in cells where the sensors are located and the shockwave caused by the merge is detected. Congestion detection results in the backward propagation of shocks. However, there are two major errors in the traffic estimates of the filter. First, the magnitude of the backward propagating shock is underestimated because there is no knowledge of the congestion from the earthquake-damaged bridge. Second, since no congestion is detected after cell 15, the estimator is unable to capture the second shock wave.

Figures 4.2d and 4.2e show the results of the approximate model when the earthquake is used as an input to the traffic model. Figure 4.2d shows that by having knowledge of the earthquake characteristics, the shock due to the bridge capacity reduction is detected. However, since no traffic sensor data is integrated into the model, the congestion wave which is a result of the merge is no longer detected due to the poor estimates of the initial and boundary conditions. With the addition of the traffic sensor data (Figure 4.2e), the estimation of the state is significantly

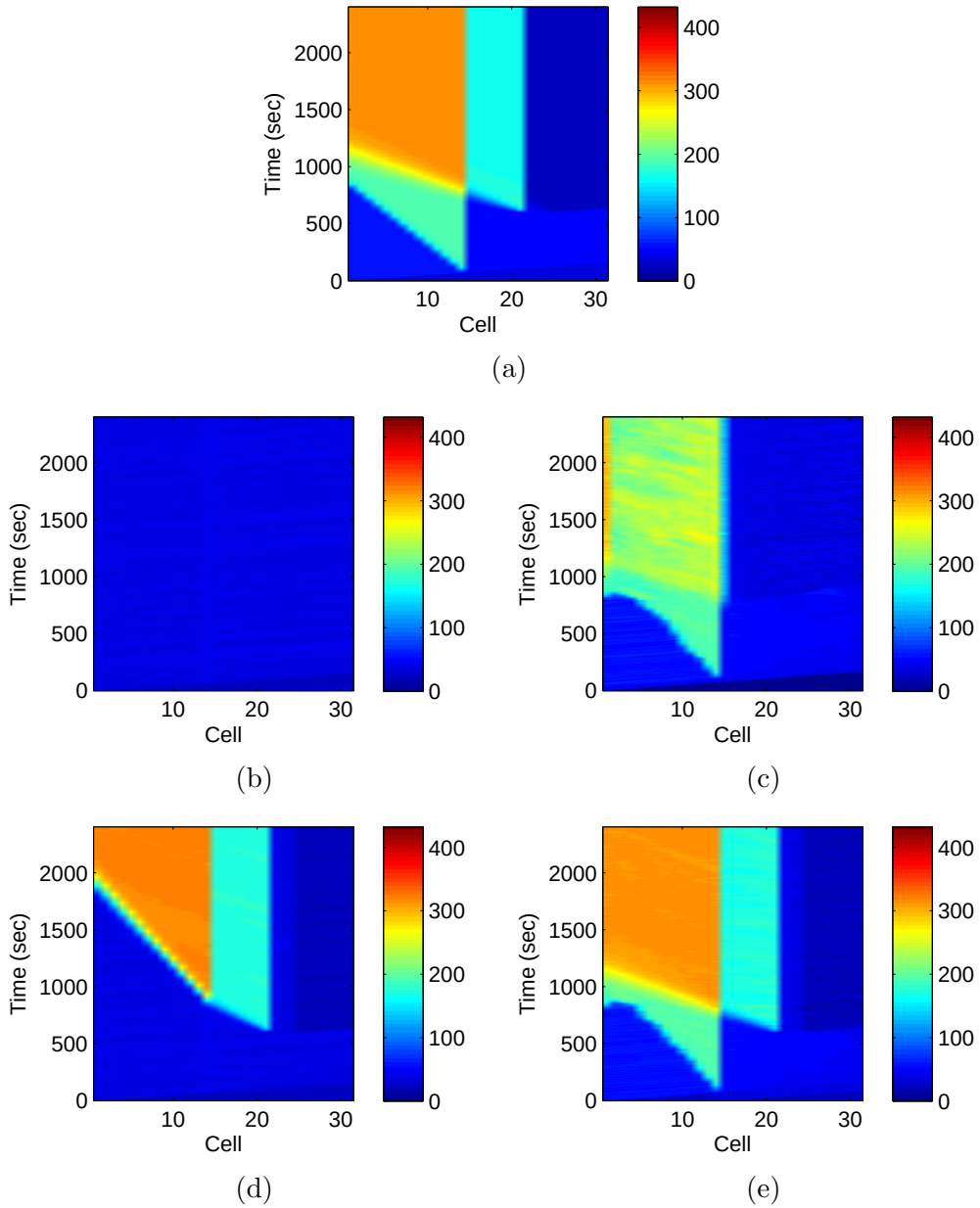


Figure 4.2: Time-space density plots for *high* damage: (a) Before the earthquake occurs there is a shock that begins at the end of Link 1 due to the merge. When the earthquake occurs that results in a 50% reduction of traffic capacity of the bridge, another shock forms at the end of Link 2. (b) Without 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 the congestion level at places where there are sensors and detects the shock wave due to the merge. (d) Supplementing the model with data from the earthquake significantly improves the model even if there is no sensor data. (e) The estimator in (d) performs better when traffic sensor data is also used.

improved, and both shocks are detected. This is an improvement on the models that have either traffic sensor data (Figure 4.2c) or the earthquake input (Figure 4.2d) and is a significant improvement on the open loop model (Figure 4.2b). It is still slightly incorrect, because the initial shock wave detected is travelling at a slightly slower rate than that of the true solution (Figure 4.2a), however, both major backward propagating shocks are successfully detected.

4.5 Changing the earthquake characteristics

In this section, model results are presented and analyzed when a less severe earthquake event occurs and when a more severe event occurs. Figures 4.3 and 4.4 show the solutions for the models used to produce the results in Figure 4.2, except with earthquake characteristics in the true model (Table 4.2) that result in *medium* and *total* damage to the bridge, respectively. The parameters given in Tables 4.1, 4.3, and 4.4 are applicable regardless of the resulting damage state of the bridge due to the earthquake.

Figure 4.3a shows the true solution when the earthquake results in only a 25% traffic capacity reduction (i.e. *medium* damage,$\alpha = 0.75$) of the bridge. Note that this is less of a reduction than if there were *high* damage to the bridge. This is apparent in Figure 4.3a because the congestion caused by the earthquake is of a lesser magnitude than in Figure 4.2a. However, the initial congestion caused by the transition from three lanes to two is exactly the same, because the initial and boundary conditions are the same and the earthquake has not yet occurred. The open loop model with no earthquake input (Figure 4.3b) does not result in any congestion detection, because the network is initialized in free flow with a flow lower than the merge bottleneck capacity. When traffic sensor data is obtained and the filter is run (Figure 4.3c), the estimates are correct for cells where sensors are located.

Also note that since the true solution shows the congestion due to the earthquake and that due to the merge are similar in magnitude, the traffic state estimates for Link 1 appear fairly accurate. However, the second, albeit less severe, shock wave due to the earthquake is undetected, so this is still an incorrect representation of the traffic state. When the earthquake input is used in the open loop model (Figure 4.3d) the shock wave caused by the earthquake is detected, however, the shock wave due to the merge is nonexistent. When the shock wave propagates backward to the merge, the resulting shock wave is of a lower magnitude and slower speed than in the case which results in *high* damage to the bridge (Figure 4.2d). Finally, when the traffic sensor data is added to the model (Figure 4.3e), the traffic state estimates are improved drastically.

Figure 4.4a shows the true solution when the earthquake results in a 100% traffic capacity reduction (i.e. *total* damage,$\alpha = 0$) of the bridge. Note that this is a much greater reduction than if there were *high* damage to the bridge. This is apparent in Figure 4.4a because the area of greatest congestion is of a much greater magnitude than in Figure 4.2a. It is actually at the jam density value of 432 veh/km (144 veh/km/lane multiplied by 3 lanes). Thus, the vehicles in the network come to a complete standstill, which intuitively makes sense, considering that the earthquake results in total destruction of the bridge, rendering it completely unusable. It may be observed that at the end of the simulation there are still vehicles in Link 3, the bridge. Although the CTM is unable to numerically capture what happens on the bridge itself once it is damaged, it does capture the inflows and outflows of the bridge. This explains why Link 4 is empty shortly after the occurrence of the earthquake; the outflow from the bridge is reduced immediately to zero. The initial shock wave due to the merge is of the same magnitude as seen previously. The open loop model with no earthquake input (Figure 4.4b) does not result in any congestion detection, because the network is initialized in free flow with a flow

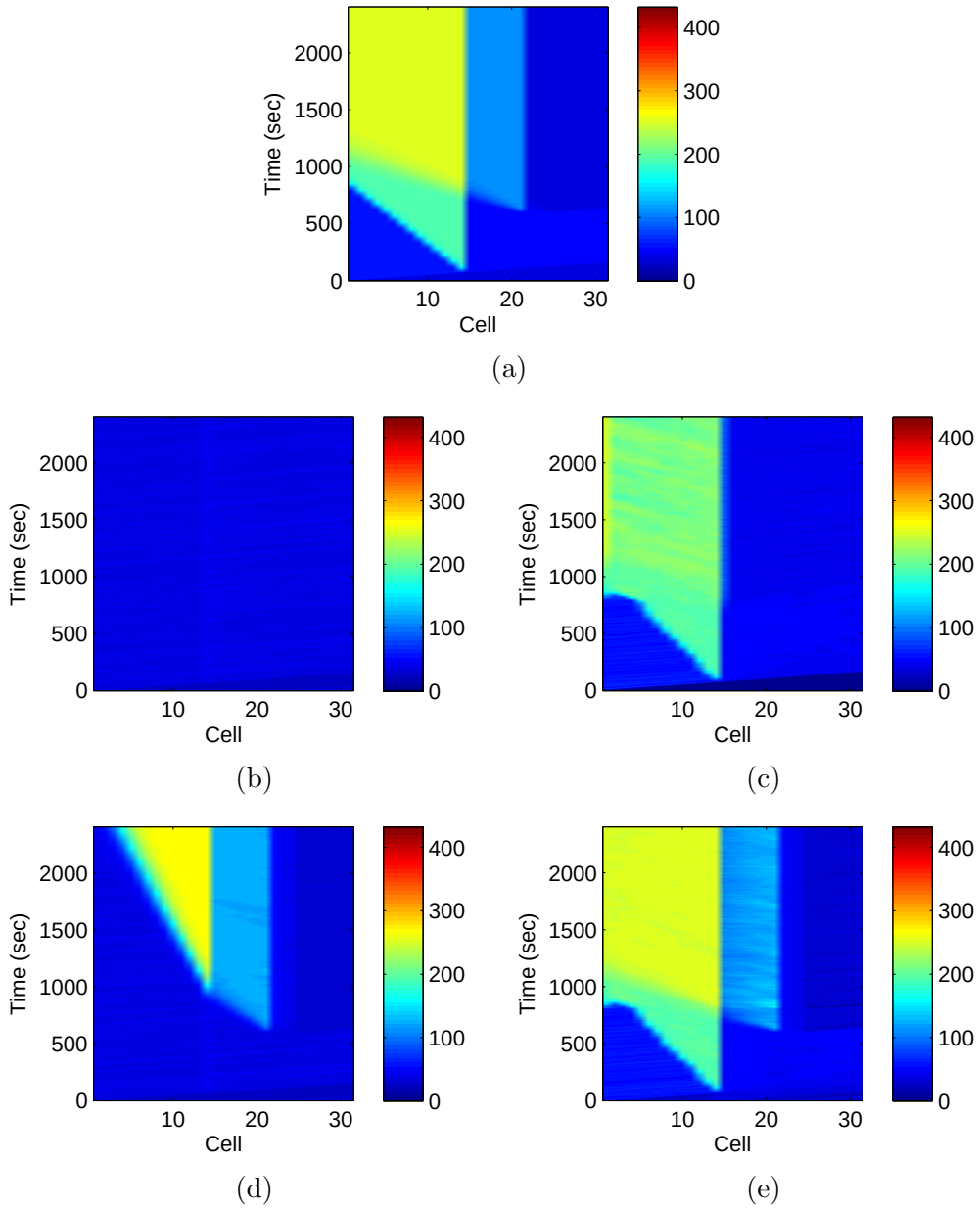


Figure 4.3: Time-space density plots for *medium* damage: (a) Before the earthquake occurs there is a shock that begins at the end of Link 1 due to the merge. When the earthquake occurs that results in a 25% reduction of traffic capacity of the bridge, another shock forms at the end of Link 2. (b) Without 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 the congestion level at places where there are sensors and detects the shock wave due to the merge. (d) Supplementing the model with data from the earthquake significantly improves the model even if there is no sensor data. (e) The estimator in (d) performs better when traffic sensor data is also used.

lower than the merge bottleneck capacity. With the addition of traffic sensor data (Figure 4.4c), congestion levels are accurately detected where traffic sensors are located and the shock wave due to the merge is detected. However, the magnitude of the traffic state estimate in cell 15, where the second traffic sensor is located, is underestimated. This is due to the difference in the density magnitudes of the sensor measurements, which are based on the true solution (Figure 4.4a), and the open loop solution (Figure 4.4b). Computationally, this results in a discrepancy in the estimation. Also, the backward propagation of congestion information can clearly be seen in Figure 4.4c. When the earthquake input is used in the open loop model (Figure 4.4d), the shock wave caused by the earthquake is detected, however, the shock wave due to the merge is nonexistent. When the shock wave propagates backward to the merge, the resulting shock wave is of a higher magnitude and faster speed than in the case which results in *high* damage to the bridge (Figure 4.2d). When traffic sensor data is added to the model (4.4e), the traffic estimates are improved significantly. However, downstream of the merge (Link 2), the traffic states appear to be underestimated compared to the true solution. This could possibly be attributed to a wider distribution of damage state draws (i.e. *total* and *high* damage states) given the true characteristics of the earthquake (Table 4.2). The filter uses the traffic sensor data to correct some of the underestimated traffic states downstream of the merge.

In comparing the results of the three experiments with different earthquake scenarios (Figures 4.2, 4.3, and 4.4), several interesting observations are noted. Due to the same road geometry and initial and boundary conditions, the true solution of the three experiments is exactly the same before the earthquake occurs. The open model with no earthquake input provides no hint of congestion in any of the experiments. The stronger the earthquake, the greater the traffic capacity reduction of the bridge, and the larger the magnitude of the resulting congestion, which is intuitive.

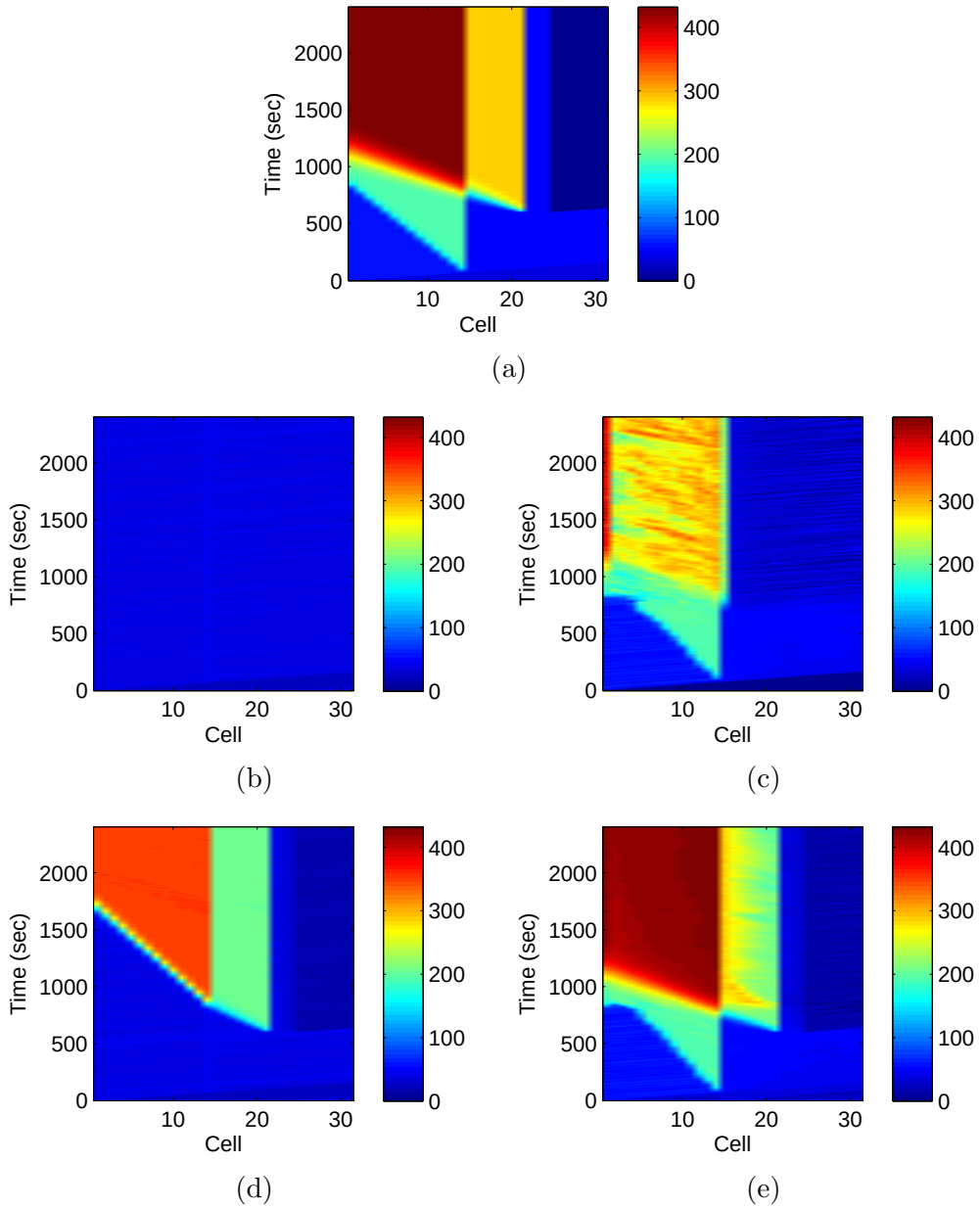


Figure 4.4: Time-space density plots for *total* damage: (a) Before the earthquake occurs there is a shock that begins at the end of Link 1 due to the merge. When the earthquake occurs that results in a 100% reduction of traffic capacity of the bridge, another shock forms at the end of Link 2. (b) Without 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 the congestion level at places where there are sensors and detects the shock wave due to the merge. (d) Supplementing the model with data from the earthquake significantly improves the model even if there is no sensor data. (e) The estimator in (d) performs better when traffic sensor data is also used.

Having either knowledge of the earthquake input or traffic sensor data is beneficial for detecting the congestion caused by either the merge or the earthquake, but not both. Finally, having both the earthquake input and traffic sensor data provides the best traffic state estimates in all three cases, however when the bridge traffic capacity is reduced to zero (i.e. *total* damage), traffic states downstream of the second sensor are slightly underestimated.

4.6 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 score of recursive estimation improvement for an N_r number of independent simulation runs indexed by $r \in \{1, \cdots, N_\mathrm{r}\}$. It is defined by the *geometric mean* of a data set, $\beta = \left(\prod_{r=1}^{N_\mathrm{r}} \beta_r\right)^{1/N_\mathrm{r}}$, which is computed through its logarithm for numerical reasons:

$$\log \beta = \frac{1}{N_\mathrm{r}} \sum_{r=1}^{N_\mathrm{r}} \log \beta_r, \quad \beta_r = \frac{\|\boldsymbol{s}_r - \hat{\boldsymbol{s}}_r\|}{\|\boldsymbol{s}_r - \bar{\boldsymbol{s}}_r\|}, \tag{4.1}$$

where $\boldsymbol{s}_r$ represents the true solution, $\bar{\boldsymbol{s}}_r$ represents the prior solution, and $\hat{\boldsymbol{s}}_r$ represents the estimation solution 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 $\hat{\mathbf{s}}_r$ 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. Although the open loop model with the earthquake input is not technically an estimator, it is a potential improvement on the open loop model without the earthquake input, thus we treat it as an estimator. All three solution quantities in Equation (4.1), $\boldsymbol{s}_r, \hat{\boldsymbol{s}}_r, \bar{\boldsymbol{s}}_r \in \mathfrak{R}^{mn_{\max}}$, are vectors that contain the 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). Since the BEEQ is essentially a ratio of the error of the estimation solution to the error of the prior solution summed over all simulation runs, a good estimator will have a BEEQ significantly less than unity (Li and Zhao, 2005). A BEEQ value of 1 indicates that the error of the estimation solution is no different than that of the open loop model and a BEEQ greater than 1 indicates that the estimation solution has a greater error than that of the open loop model.

Table 4.5 shows the computed BEEQ values 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 data (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 4.2c and 4.2d) corresponding to these estimators.

Table 4.5 illustrates several interesting results. For an earthquake that results in *medium* damage to the bridge, the BEEQ values suggest that having 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 the earthquake input. This is due to the fact that, in this specific example, the magnitude of the congestion due to the merge from Link 1 to Link 2 is of a similar magnitude to the congestion of the true solution. The increasing severity of damage due to the earthquake does not affect the magnitude of the initial shock wave until at least time $t = 10$ minutes (600 seconds), at the occurrence of the earthquake. Note that this trend is specific to the setup of our numerical simulation (e.g. road geometry, initial and boundary

Damage	No earthquake Filter	Earthquake Open loop	Earthquake Filter
medium	0.2882	0.7785	0.1159
high	0.4127	0.5552	0.0906
total	0.5476	0.4932	0.1227

Table 4.5: Computed BEEQ values for various estimators for different earthquake scenarios averaged over 100 experiments

conditions, noise models) and might not be observed in different scenarios.

The fourth column of Table 4.5 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 around 0.10. Put another way, using this estimator results in about 10% as much error as 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.

4.7 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 capture the whole distribution, but using too many could result in very high computational costs, especially for complex systems. Thus, the performance of the proposed algorithm is also tested with different numbers of ensembles, N. Table 4.6 shows the running times per simulation, BEEQ values, 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 for each doubling of the number of ensembles used, the running time approximately doubles. Thus, the number of ensembles

Number of ensembles, N	Running time (min/sim)	β	σ_{g}
25	3.4	0.0931	1.2532
50	6.6	0.0887	1.2578
100	13.0	0.0906	1.2165
200	26.5	0.0917	1.1372
500	68.6	0.0895	1.1141

Table 4.6: BEEQ values and running time for using different amounts of ensembles for the estimator that has the earthquake input and traffic sensor data

used is a key factor to focus on when looking to reduce computational cost. The third column is the BEEQ computed using Equation (4.1). 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 as the number of ensembles increases, σ_{g} generally decreases. For this particular numerical experiment, the number of ensembles has little effect on the efficacy of the estimator, which means a lower number of ensembles can be used to produce similar results for a smaller computational cost. However, since the real-time implementation only runs the algorithm once ($N_{\mathrm{r}} = 1$), by using a lower number of ensembles one risks obtaining an outlier since the resulting geometric standard deviation is higher compared to that obtained when using a higher number of ensembles.

All simulation results in this thesis were produced using MATLAB®. The experiments performed in Sections 4.6 and 4.7 were done on an Intel® Core™ i7-4770 3.40 GHz CPU and 8 GB of RAM. Parallel computing using four cores was utilized to run simulations simultaneously, though the authors recognize that the ensembles themselves could have also been parallelized.

Chapter 5

Conclusions

In this thesis, we show proof of concept for an estimation technique for determining the post-disaster traffic behavior due to a reduction of flow from a natural disaster such as 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, location) centered around the true earthquake characteristics for each ensemble. Synthetic traffic data using the true solution 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 is in the filter is approximately 10% 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, there are extensions to consider. Although the results presented in this thesis were only illustrated on a single stretch of roadway, the forward model we implemented is designed to handle networks including merges and diverges, which may occur due to intersections or highway on-ramps and off-ramps. The scalability of the algorithm should also be tested since increasing the network size would result in higher computational costs and possibly reduced efficacy of the traffic state estimation.

References

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

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

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

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

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

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

K. W. Campbell. 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, 1997.

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

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

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

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

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

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

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

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

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

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

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

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

Z. Hong and D. Fukuda. 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, 2012.

Q. Huang, P. Gardoni, and S. Hurlebaus. 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, 2010.

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

S. J. Julier and J. K. Uhlmann. 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, 1997.

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

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

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

K. Kosa. 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, 2012.

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

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

M. J. Lighthill and G. B. Whitham. 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, 1955.

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

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

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

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

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

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

H. Nakanishi, K. Matsuo, and J. Black. 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, 2013.

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

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

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

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

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

M. Shizunoka, Y. Murachi, X. Dong, Y. Zhou, and M. J. Orlikowski. 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*, 2003.

D. Simon. *Optimal State Estimation: Kalman, H-Infinity, and Nonlinear Approaches*. Willy-Interscience, 2006.

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

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

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

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

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

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