Estimating probabilistic impacts of catastrophic network damage from earthquakes using an activity-based travel model

Mahalia Miller\textsuperscript{1}, Samuel Cortes\textsuperscript{1}, Dave Ory\textsuperscript{2}, Jack W. Baker\textsuperscript{1}

ABSTRACT

In this paper, we present a framework for a probabilistic risk assessment of catastrophic network damage from earthquakes using a practical activity-based travel model. The travel model captures variable demand, non-car travel modes, and interdependencies between the road and transit network, features often missing in probabilistic network risk studies to date. We have adapted this model for estimating network performance and regional impacts after catastrophic network damage. We demonstrate our framework with a case study of earthquakes in the San Francisco Bay Area, including damage to the extensive road and transit networks. For the case study, we partnered with the local Master Planning Organization (MPO) to demonstrate the method with an activity-based model used in practice. In addition, while most work to date in the transportation field focuses on a single deterministic earthquake event, we apply state-of-the-art seismic risk assessment including a stochastic set of earthquake events, spatially-correlated ground-motion intensity maps, damage maps for bridges and other structures, practical network performance measures, and the probabilistic risk to different communities and households. One result of this framework is quantifying the annual rate of exceeding various levels of loss in accessibility for each community and main socio-economic group in the entire region. The presented method offers a rare level of detail for understanding the potential impacts of earthquakes to travel-related outcomes, and thus informing mitigation efforts and emergency planning.

INTRODUCTION

An event-based probabilistic loss estimation model considers realizations of relevant random variables, such as earthquake magnitude and component damage state, sampled from corresponding distributions. For each event realization, the impact to the infrastructure network is evaluated (e.g., Figure 1). This general risk analysis framework is common in academic literature and in practice in the seismic risk assessment field \cite{1,2}. However, infrastructure networks can be complex and computationally expensive to evaluate, so it is common in the seismic risk assessment field to consider a simplified network \cite{3}. For example, the complex web of highways, local roads, bus lines, and trains is often represented as a simplified graph of a few dozen nodes and edges. Furthermore, in transport planning the problem is often simplified by considering a single scenario event instead of the full probabilistic risk assessment \cite{4}.

In this paper, we present an event-based framework for estimating the probabilistic impacts of catastrophic network damage from earthquakes using a practical activity-based travel model. For illustration, we consider the seismic risk to the San Francisco Bay Area transportation network...
network. After collecting and processing case study data, we generate spatially-correlated ground-motion intensity maps that capture the level of ground shaking intensity at each location of interest. We then use these maps to sample realizations of damage to the bridges and other components in the transportation network, which are considered particularly vulnerable in the event of an earthquake. The component damage states enable generating the damage map for the full transportation network, including public transportation. The damage map has a realization of the damage state (functional or closed) for each component of interest. These damage states are then translated to the road and transit networks, impacting the functionality of different road segments and transit line segments. Then, we represent the damaged transportation network in an activity-based travel model that combines a variable demand model and various travel modes, such as walking, driving, biking, ride sharing, and taking public transportation. We show how this model facilitates calculating a wealth of performance measures more complex than in prior work. Furthermore, we evaluate the annual likelihood of exceeding a certain loss of network performance, a common metric in seismic risk assessment and the insurance industry (1). This activity-based model of a reasonably-comprehensive network adds more realistic models to seismic risk assessment. Additionally, the event-based probabilistic framework enables transport planning to transition from largely scenario-by-scenario assessment to a risk-consistent approach. Thus, it represents a collaboration between seismic risk assessment and transport planning. In addition, this work provides a method that is only a harbinger of the potential range of applications activity-based models enable.

BACKGROUND

Performance Measures

For seismic risk assessment, many performance measures have been proposed to model the complex array of impacts an earthquake may have on transportation infrastructure and the surrounding communities. Two common performance measures for road networks are connectivity (5) and flow capacity (6). Connectivity captures whether one can travel from a start node to a destination node. Flow capacity refers to the physical amount of cars or people who can travel between a start node and a destination node, often modeling only the road network and considering some key roads as network elements with a known capacity.

In order of increasing general computational cost, other measures to capture road network performance include the percentage of bridges damaged, weighted-shortest path between locations of interest (7), fixed-demand travel-time (8), and the economic impacts from increased travel time and bridge repairs (9). Fixed-demand travel time delay and its variants have become particularly popular in current literature. Fixed-demand travel time is found by assigning predetermined trip demand to a potentially damaged graph assuming user equilibrium. In other words, the model assumes each driver (user) has perfect information about trip times and chooses the fastest or shortest route. The problem is solved iteratively until a steady-state is found. While this performance measure offers insight into aggregate travel behavior, most studies have three shortcomings: 1) travel demand (thought of as a volume or amount) is a constant, 2) the varied desirability and importance of various destinations is ignored, and 3) travel mode choices remain fixed, usually assuming all trips are by car. In contrast, in this study, we will consider a more sophisticated performance measure that overcomes these limitations, accessibility, as measured by logsums from representative destination choice models, also modeling variable demand and travel mode choice.
Network Representation

For various reasons including computational expense, many seismic risk assessment studies to date have considered a portfolio of transportation network structures or highly aggregated networks. For example, Wang et al. (10) evaluated the seismic performance of the Bay Area Rapid Transit (BART) aerial structures, but did not extend the study to the network impacts. Moreover, while Wakabayashi et al. (11) did consider a network representation, they simplified the San Francisco Bay Area transportation network to a set of 31 road segments; their aim was to compare historical performance in the 1989 Loma Prieta Earthquake with their proposed transportation model. A slight increase in complexity is found in Chang et al. (12), who modeled the Los Angeles road network by a set of 202 road segments. More recently, work has considered networks aggregated to dozens of links, such as (6). Two notable exceptions are the assessments in (8) and (13) that modeled the San Francisco Bay Area transportation network as 586 and 26,522 road segments respectively. Neither considered public transportation.

Furthermore, while many scenario-based studies have been performed to improve our understanding of the impacts of earthquakes on transportation, very few have considered a probabilistic assessment to capture the distribution of possible losses of a wide range of damaged networks, particularly in transport planning. For example, B. Baker and R. Miller considered the ground shaking intensity with a 10% in 50 years exceedance probability, the “design level earthquake,” at 42 Seattle-area bridges, in order to do a cost-benefit analysis of seismic retrofit (14). However, that study considered neither a stochastic set of earthquake events nor uncertainty in the performance of the bridges given a certain level of shaking. Furthermore, by assuming that each bridge experienced the 10% in 50 years ground shaking simultaneously, the potential network impact, and thus the benefits, may be overestimated.

This work represents a departure from the prior work through three contributions: to transport planning, it demonstrates the use of a practical activity-based travel model for probabilistic risk analysis; and to seismic risk assessment, this study proposes both a reasonably comprehensive transportation network, and a more realistic transportation model with variable travel demand and shifts in transport mode.

MODELING NETWORK PERFORMANCE

Network Description and Model Representation

This study applies seismic risk assessment to the transportation network of the San Francisco Bay Area. We consider the accessibility losses to measure how easy it is to get to desirable destinations following a probabilistic set of earthquake events. We use a practical, agent-based transportation model, Travel Model One (Version 0.3), used by the Metropolitan Transportation Commission (MTC), the regional metropolitan planning organization (MPO) for the nine county San Francisco Bay Area. This will be described in more detail below.

This study considers a road network of 11,921 nodes representing road intersections and 32,858 edges representing road segments and centroid connectors, and a detailed transit network representing service provided by 43 agencies. The road network contains highways and key local roads. We have linked it with 1743 bridges and other components (including footbridges and overpasses, e.g.) for which we model possible seismic damage. Furthermore, we model damage to some key transit lines, including the local light rail and heavy rail lines. 60 road-network
components and 1409 additional components are modeled as potentially impacting these transit lines. Readers are referred to \((15)\) for more details about the network and components.

**Ground-Motion Intensity Map Models**

Ground-motion intensity maps express shaking intensities following a possible, future earthquake. In this study, we measure the intensity with the 5%-damped pseudo absolute spectral acceleration \((\text{Sa})\) at a period \(T=1\)s, which is the required input below. This spectral acceleration value represents the maximum acceleration over time that a linear oscillator with 5% damping and a period of 1 second will experience from a given ground motion. We calculate these values at each component location (each bridge and structure location). We use the UCERF2 seismic source model \((16)\), OpenSHA Event Set Calculator \((17)\), Wald and Allen topographic slope model for the shear wave velocity \(V_{s30}\) \((18)\), and the Boore and Atkinson ground-motion prediction equation \((19)\).

We simulate the sets of maps by combining the mean ground-motion intensity terms from the Event Set Calculator and spatially-correlated residual terms of the ground-motion intensity, using \((20)\), according to the basic ground-motion model, Equation 1.

\[
\ln Sa_{ij} = \ln Sa(M_j, R_{ij}, V_{s30,ij}, \ldots) + \sigma_{ij} \varepsilon_{ij} + \tau_j \eta_j
\]

where \(j\) is the ground-motion intensity map index \((j = 1, \ldots, m\) where \(m\) equals the total number of spatially-correlated ground-motion intensity maps); \(\varepsilon_{ij}\) is the normalized within-event residual in \(\ln Sa\) representing location-to-location variability; \(\eta_j\) is the normalized between-event residual in \(\ln Sa\); \(M_j\) is the magnitude; \(R_{ij}\) is the source-to-site distance; \(\sigma_{ij}\) and \(\tau_j\) are the standard deviation terms of the residuals; and the other variables are defined above. Both \(\varepsilon_{ij}\) and \(\eta_j\) are normal random variables with zero mean and unit standard deviation. The vector of \(\varepsilon_{ij}\) can be modeled by a spatially-correlated multivariate normal distribution and the \(\eta_j\) by a standard univariate normal distribution.

**Damage Map Models**

Damage maps capture the level of functionality for the components of interest, bridges and other structures in this case. Fragility functions translate ground shaking intensity into possible structural damage. The damage state is a discrete random variable that in this case represents if the component is functional or not functional one week after an earthquake event. We chose the performance one week after an earthquake after discussions with the California Department of Transportation (Caltrans) because, at this point in time, minor road surface repairs will be finished, but major reconstruction, which depends on various social, economic, and political factors, has not yet been completed. The damage state is conditioned on a realization, \(sa\), of the random variable \(Sa_{ij}\), the ground-motion intensity at the \(i^{th}\) site and \(j^{th}\) ground-motion intensity map. By sampling damage states at each component for a given ground-motion intensity map, we build a damage map.

Functional percentage relationships link the component damage state to the functionality of network elements. For example, when a bridge collapses, the traffic flow capacity of the road it carries and it crosses can be modeled as reduced to zero. These relationships are often derived from a combination of observation and expert opinion, often due to data scarcity \((21)\). Furthermore, the relationships are typically deterministic for a certain component damage state.
and restoration time (21). Thus, in this paper, each damage map corresponds to a functionality state for every element of the network, such as a road segment or part of a transit lines.

**Accessibility Performance Measure**

We demonstrate one example of a sophisticated performance measure, accessibility. This accessibility measure is computed by taking the log value of the sum of a function of the utilities of each destination over all possible destinations and travel modes from representative destination choice models, where the utility decreases if getting to that destination is more costly or time-intensive (22). Readers are referred to Miller (15) for more details.

In contrast to prior work that has used more simplified metrics for measuring post-catastrophe impact, this performance measure captures the fact that some destinations and trips have higher value than others. Furthermore, this measure allows analyzing individual communities and demographic groups without aggregating region-wide.

**IMPLEMENTATION WITH AN ACTIVITY-BASED TRAVEL MODEL**

The activity-based travel model used by the Metropolitan Transportation Commission (MTC) includes an explicit representation of every freeway and major local road in the Bay Area. Further, it includes an explicit representation of every transit route by time of day, including rail, ferry, and bus service. Simple representations of bicycle and pedestrian infrastructure are also included. Taking as input household demographic and location information, and employment type and location information, the model makes detailed and validated predictions of travel-related outcomes, including automobile ownership, activity schedules and locations, travel mode choice, and travel route choice. The model allows each of these aspects of travel demand to change in response to an earthquake. As such, travel demand is variable, responding to the particular details of the damaged networks. This model uses a combination of Java code called CT-RAMP (23), and the Citilabs Cube Voyager and Cube Cluster software programs, which are part of a leading commercial software suite for transportation planning (24).

For typical regional planning applications, this model is run in three main iterations, each with two stages. Stage I is the agent-based model called CT-RAMP. An agent represents a person (who is a member of a household) who decides when, where and how he or she wants to travel, if at all, for a typical weekday. Given travel times for different route options and distributions representing travel preferences, travel choices of each agent are simulated. As the travel time increases, for example, it is more likely that an agent may telecommute or forgo a non-mandatory trip. Typically, a sample of the fully simulated population is modeled. At the end of this stage, based on the choices of each agent, a travel demand origin-destination matrix between each of the 1454 travel analysis zone (TAZ) zones is created; the matrix specifies the number of trips between each pair of TAZs and is scaled in proportion to the aforementioned
FIGURE 1. Illustration of the risk framework for one earthquake event including a) ground-motion intensity map, b) damage map, c) map of travel time increase values, and d) map of accessibility values averaged over all market segments per travel analysis zone. (1 mile = 1.6km)
sampling percentage to represent the full population. Thus, the end result per iteration is an estimate of the travel demand between zones in a steady state given this network. Then, in Stage II, the Citilabs software assigns all of these trips to the roads and transit networks. Trips are assigned using a traditional iterative process following the user equilibrium method (25), wherein trip assignment is iteratively adjusted until each agent takes the route with the shortest travel time (within a certain tolerance). In other words, the model might run hundreds of iterations each time Stage II is performed. A key assumption in this method is that the agent has perfect information, so that he or she would indeed take the route with the shortest travel time. On one hand, this may be less realistic after an earthquake, where travelers will not know the new fastest routes. On the other hand, if there is working power and internet, travelers may be more likely to consult navigation tools that would suggest the fastest trip route. The result of Stage II are “loaded networks” that specify the amount of travelers and travel times on each segment of the network; this is then the input to the demand creation of Stage I of the next model iteration.

Note that when running the travel model in this way we are making a host of implicit assumptions about the potential traveler response to earthquakes; these include: 1) large shifts in telecommuting participation do not occur (may overestimate the impact of the seismic event); 2) households and employers do not relocate within the region or move out of the region (may overestimate); and, 3) travelers locate employment opportunities informed by accessibility (may underestimate). Importantly, these assumptions can be easily modified within the MTC travel model; this is an avenue for future research.

While validating the entire model and method would be ideal, this would require measurements of ground-motion intensity, damage states of relevant structures, travel patterns, and travel outcomes after earthquakes. Such data is currently unavailable. For example, in the San Francisco Bay Area, travel pattern data has only been collected since approximately 2006, during which time there have been no major earthquakes. Furthermore, to increase confidence in the validation, one would need to collect these datasets after not just one major earthquake, but rather after a representative set of earthquake events. Thus, until this data is available, only individual parts of the model have been validated, such as the fragility functions. However, this paper presents a method that can be implemented today with available models, and the results will continue to be refined as more sophisticated, validated models come to exist.

We modify the high-fidelity model in two ways for modeling earthquake events: considering a damaged network, and introducing extra iterations to more reliably locate a stable result. Figure 2 illustrates the proposed procedure.
Figure 2. Iterative process to bring high-fidelity model to convergence while in a disrupted state.

**Damaging the Network**

We input a damaged road network and new transit route information into the travel model, with only the functional roads and transit routes for the earthquake event. This is achieved by simulating ground-motion intensity maps and damage maps, which have a realization of the damage state for each component. For each component simulated as causing no functionality, we use look-up tables to identify the network impact. These look-up tables were created by manually inspecting aerial images and free text descriptions of the transportation networks and components to match them. A field investigation also aided this matching, as we explained in (15). For the road network, the tables enable finding the start and end nodes of various road segments that would be impacted (such as roads a bridge carries and crosses).

The next step is to actually damage the network input files. A Python script generates a Citilabs Cube input file that automatically sets the facility type of each affected road segment to a new dummy type. This dummy facility type corresponds to no traffic flow capacity in the travel model. For the transit networks, a script removes transit lines by commenting out the relevant lines or systems in the model input files. The result is that only the functioning transit routes are included in the model.

**Designing for Stability**

We introduce three initial iterations -- each with both stages -- to provide an appropriate warm start for the aforementioned three main iterations, again each with both stages. Since travel times may be significantly different after a catastrophic earthquake, we gradually reduce the capacity on damaged edges from full capacity to none over the first three initial iterations. Otherwise, the agent-based travel demand assignment in Stage I, the trip assignment in Stage II, or both, is not likely to converge.
For Stage I of all six iterations, we randomly sample 1% of the population to be agents. A sample share of 1% balances granularity in demand choices and computation time. For example, for a sample earthquake event, when we increased the sample share to 5% on the last two iterations, the runtime increased from 7 hours to 8.5 hours (21% increase), but the root mean-squared error (RMSE) of the vehicle-miles traveled values between the second to last and last iteration reduced only from 2.7% to 2.5% (7% reduction). Less than 5% is a general rule of thumb used by MTC for the upper bound for these RMSE values for region-wide analysis. In other words, with six iterations, larger population sample sizes were not required to get below this error upper bound. Note that Stage II does not directly consider agents nor a sampled population. On one 2.27 GHz server with 128 GB of RAM, each event simulation with six iterations takes 6 -11 hours, where scenarios with low damage take closer to six hours and scenarios with high damage take longer to converge.

Computing Changes in Accessibility

We use a script in the MTC Travel Model One to compute the accessibility per representative agent at the end of each damage map simulation. The accessibility per agent is aggregated into 12 market segments used by MTC, with all combinations of income class of his/her household (26).

We consider the change in the accessibility values from the base case of no damage to the \( j^{th} \) network damage map. In other words, the accessibility change is defined as:

\[
\frac{\sum_{a=1}^{A} Acc_{a,j'}}{A_{j'}} - \frac{\sum_{a=1}^{A} Acc_{a,base}}{A_{base}}
\]

(2)

where \( Acc_{a,j'} \) refers to the accessibility of the \( a^{th} \) agent for \( a=1,...,A \) in the \( j^{th} \) network damage map; \( Acc_{a,base} \) refers to the accessibility of an agent in the base case; \( A_{j'} \) is the total number of agents in the \( j^{th} \) network damage map; and \( A_{base} \) is the total number of agents in the base case. The number of agents may vary slightly between damage maps, since the model considers a random sampling of agents and households.

DISCUSSION

Example Results

The modeling framework enables comparing the predicted losses in accessibility for each market segment and damage map event. For example, for the M7.45 San Andreas event shown in Figure 1 panel d) shows the losses in accessibility (Equation 2) by TAZ. This figure suggests that people living in South San Jose, in some communities on the San Francisco Peninsula, and along the Bay in the East Bay would be particularly impacted by losses in accessibility, as compared to their fellow Bay Area residents.

With this modeling framework, we can also examine changes in expected travel patterns. In this earthquake event, there is approximately the same number of total trips with interesting changes in travel modes. While the results predict an 18% decrease in trips by transit, there is a 1% increase in trips by foot and a 1% increase in trips by car. For this particular event, the commuter train Caltrain, the light-rail network VTA Light Rail, and 4 of the 14 heavy-rail network BART routes are predicted to be closed one week after this event, due to extensive
FIGURE 3. Example M7.05 San Andreas Fault event in the stochastic event set featuring: a) ground-motion intensity map, b) damage map, c) map of travel time increase values, and d) map of accessibility values averaged over all market segments per travel analysis zone. (1 mile = 1.6km)
FIGURE 4. Example M7.05 Hayward Fault event in the stochastic event set featuring: a) ground-motion intensity map, b) damage map, c) map of travel time increase values, and d) map of accessibility values averaged over all market segments per travel analysis zone. (1 mile = 1.6km)
structural damage that blocks the tracks. Nonetheless, for this particular event, the Muni light rail is expected to be fully operational, and we have modeled ferries and local buses as functioning too. Thus, the slight shift away from transit is not surprising. This result, however, offers a realistic possible event for emergency planners considering possible seismic impacts across agencies and municipalities.

Extending Scenario-Based Assessment to a Stochastic Event Set

Restricting our analysis to a particular event would blind us to the range of possible losses from earthquake hazard for the case study region. Furthermore, we would be unable to estimate the annual rate of exceeding a certain level of losses, which is crucial information for risk-informed decisions and financial models. Thus, we extend the assessment for a single event to a hazard-consistent stochastic event set. We use the stochastic event set described in (15), which contains 40 sets of earthquake scenarios, ground-motion intensity maps, damage maps, and corresponding annual rates of occurrence. Each damage map is modeled using the activity-based model described above. Figures 3 and 4 illustrate two further examples of earthquake events, each with an annual rate of occurrence. The figures show that different events can cause dramatically different patterns of damage and impacts. To automate this process, we have written a batch script in MS-DOS that runs all 40 events in series automatically. Sample code is available for free download at: http://purl.stanford.edu/mh896js1648.

FIGURE 5. Accessibility annual loss exceedance curve with a comparison by 3 case study TAZs as compared to a region-wide average over all TAZ.

By considering all damage map events and the corresponding annual rates of occurrence, we also compute the accessibility-change loss-exceedance curve for each market segment, either for each TAZ or in averaged over all TAZ (e.g., Figure 5). For example, this figure shows that the rate of exceeding losses of 0.04 utils per person per day is approximately 0.015 (67 year return period) for San Franciscans, but at that same annual frequency, people in Pacifica are
expected to suffer losses exceeding 0.27 \textit{utils} per person per day. A \textit{util} is a dimensionless quantity capturing the utility with regards to time and money. In other words, at a 67 year return period, a person in Pacifica (as labeled in Figure 1, panel a) is expected to experience more than six times higher losses in accessibility as compared to his or her compatriot just to the North in San Francisco.

Estimating the accessibility-change loss-exceedance curve has broad applications. The insurance industry has a long history of using these rates to price insurance by assessing how likely certain levels of payout might be. It is also useful to the transportation planner for weighing costs and benefits over a given time span.

\textbf{Applicability to Other Hazards}

In this paper, we have demonstrated our framework with a case study focusing on earthquakes. However, this framework has applicability to other hazards, both natural and man-made. For example, in the case of the combination of sea level rise and storm flooding, each event corresponds to flooding at a certain point in time for a particular event. Then, depending on the flooding, certain road segments and transit lines can be modeled as non-operational, as we have done here. Readers are referred to Heberger (27) for an example of mapping a certain level of flooding, such as 1.4 meters increase in sea level, to the roadways that could be impacted. Then, with the impacted network segments identified and translated to model input files, a transportation planner can go the next step of understanding the impact on accessibility and other sophisticated metrics using the activity-based model, as detailed above. In addition, he or she can extend this scenario-based assessment to a hazard-consistent study using a stochastic event set, as described in the previous section.

Readers are advised of certain limitations of this work and opportunities for future work. In addition to offering the opportunity to consider other hazards, this paper lays the foundation for expanding the analysis to more sources of damage, such as to tunnels or from liquefaction. Another opportunity is adjusting the travel demand properties. In this analysis, destinations are assumed to have a similar desirability before and after an earthquake, but we could imagine that some business districts may suffer heavy damage that could impact demand. Furthermore, one could additionally investigate the social dynamics following catastrophic events and how this might further alter demand.

\textbf{CONCLUSION}

In this paper we have demonstrated a novel application of a practical activity-based model to probabilistic seismic risk assessment of transportation networks. By capturing both travel mode choice and varying travel demand, two major limitations of previous work in seismic risk assessment are overcome. At the same time, this paper contributes a framework for probabilistic risk assessment to transport planning, using a practical model. Moreover, this work is being performed in a manner consistent with how myriad policy analyses are performed in the Bay Area; this model is already used for allocating funds for infrastructure improvements. This paper’s contribution is modifying this model for a new application—assessing probabilistic seismic risk. Future work will further investigate the use of this framework for better targeting seismic risk mitigation efforts. The result will be a continued collaboration between the transport planning and seismic risk assessment communities, in order to reduce the risk of losses after future earthquakes. In addition, the various agencies supporting transportation in the region have
a great opportunity to coordinate risk mitigation, so at least one means of transportation is relatively resilient for each community. Finally, a promising area of future work is applying this framework to other hazards such as the combination of sea level rise and storm flooding.

ACKNOWLEDGMENT

The authors would like to offer heartfelt thanks to Tom Shantz at the California Department of Transportation (Caltrans), Mark Salmon at the Bay Area Rapid Transit (BART), and the Metropolitan Transportation Commission (MTC) for the case study data. The first author also thanks the National Science Foundation Graduate Research Fellowship and the Stanford Graduate Fellowship for generous financial support.

REFERENCES