A Prototype Operational Earthquake Loss Model for California Based on UCERF3-ETAS – A First Look at Valuation

Edward Field, M.EERI, Keith Porter, M.EERI, and Kevin Milner, M.EERI

We present a prototype operational loss model for California based on UCERF3-ETAS, which represents the first earthquake forecast to relax fault segmentation assumptions and to include multi-fault ruptures, elastic-rebound, and spatiotemporal clustering, all of which seem important for generating realistic and useful aftershock statistics. UCERF3-ETAS is nevertheless an approximation of the system, however, so usefulness will vary and potential value needs to be ascertained in the context of each application. We examine this question with respect to statewide loss estimates, exemplifying how risk can be elevated by orders of magnitude due to triggered events following various scenario earthquakes. Two important considerations are the probability gains, relative to loss likelihoods in the absence of main shocks, and the rapid decay of gains with time. We hope this paper inspires similar analyses with respect to other risk metrics to help determine whether operationalization of UCERF3-ETAS would be worth the considerable resources required.

INTRODUCTION

Long-term earthquake forecasts, which are applicable from decades to centuries, represent our first line of defense with respect to mitigating earthquake risk, especially in terms of informing building codes. We also know, however, that aftershocks and otherwise triggered events can be large and damaging, as demonstrated by several recent sequences including that which produced the 2011 M 6.3 Christchurch (e.g., Kaiser et al., 2012) and

a U.S. Geological Survey, 1711 Illinois Street, Golden, CO 80401, USA; field@usgs.gov
b University of Colorado Boulder, 428 UBC, Boulder CO 80309-0438
c University of Southern California, 3651 Trousdale Parkway #169, Los Angeles, California 90089-0742
2016 M 7.8 Kaikoura earthquakes in New Zealand. The ability to forecast earthquakes on shorter time scales, such as days to decades, is known as Operational Earthquake Forecasting (OEF), which also involves the dissemination of authoritative information to inform risk mitigation decisions (Jordan and Jones, 2010; Jordan et al., 2011). The history, motivation, and challenges associated with OEF have been discussed elsewhere (e.g., Jordan and Jones, 2010; Jordan et al., 2011; Peresan et al., 2012; Wang and Rogers, 2014; Jordan et al., 2014, Kossobokov et al., 2015; Goltz, 2015), and significant progress with OEF has been made in both Italy following the 2009 L’Aquila earthquake (Marzocchi et al., 2014) and in New Zealand following the 2010 Darfield/Christchurch events (Gerstenberger et al., 2014). Furthermore, a workshop was recently held to discuss the potential usefulness of OEF, at which there was generally broad support for the capability among a wide range of stakeholders (Field et al., 2016). This view is consistent with the fact that OEF developments are a strategic-action goal of the United States Geological Survey (USGS, http://pubs.usgs.gov/of/2012/1088; page 32).

The USGS has in fact been issuing aftershock forecasts following various earthquakes since the 1980s. These reports are based on the Reasenberg and Jones (1989, 1994) model, which estimates earthquake probabilities from empirical Omori-Utsu and Gutenberg-Richter statistics (Utsu, 1961, 1971; Gutenberg and Richter, 1956). This approach provides aftershock probabilities as a function of time and magnitude, but lacks information on where the triggered events might occur. Gerstenberger et al. (2005) added spatial information to the Reasenberg and Jones algorithm in their Short Term Earthquake Probability (STEP) model, which also provides real-time hazard maps. The USGS produced STEP forecasts for California between 2005 and 2010, but discontinued these postings thereafter due to software maintenance issues.

A significant limitation of these and other candidate OEF models is a lack of information on proximity to known active faults. For instance, the California Earthquake Prediction Evaluation Council (CEPEC), which advises the California Office of Emergency Services (CalOES) on earthquake threats, is known to convene when an M ~5 earthquake occurs near the San Andreas fault, but not when such an event occurs far from any known fault (Jordan and Jones, 2010). In other words, proximity to a fault and its activity level are perceived as important when it comes to large-event triggering. Furthermore, the Omori-Utsu statistics applied in previous OEF models imply that the most likely place for a triggered event is the
location of the main shock, whereas the elastic-rebound theory applied in fault-based forecasts says the opposite – the least likely place for a large triggered event is exactly where the fault just ruptured. The USGS National Seismic Hazard Maps also imply that fault-based sources dominate hazard estimates in most parts of California, so simply ignoring their existence might be a serious limitation with respect to OEF.

Including fault-based information has therefore been a major goal of the candidate OEF model under considered here, which we refer to as UCERF3-ETAS for reasons discussed in the next section. While this model is believed to represent a significant improvement over its predecessors, especially with respect to including faults, it too embodies assumptions and approximations. Given all models are ultimately wrong, potential usefulness can only be ascertained in the context of specific hazard or loss studies (e.g., Field, 2015), as a model may be a suitable approximation in one situation but not another. In other words, we need to add valuation (Jordan et al., 2011) to our verification and validation protocols. This study effectively represents an initial foray into UCERF3-ETAS valuation with respect to statewide loss estimates.

Iervolino et al. (2015) were the first to introduce and demonstrate feasibility with respect to what they termed Operational Earthquake Loss Forecasting (OELF). Specifically, they deployed an experimental system for Italy, called MANTIS-K, which produces real-time risk maps with respect to potential building collapse, displaced residents, injuries, and fatalities. Marzocchi et al. (2015) subsequently used this system to demonstrate that while triggering probabilities may remain low during an aftershocks sequence, the potential consequences, such as individual risk of death, can rise to actionable levels. Similarly, Herrmann et al. (2016) have shown that combining fatality estimates with cost-benefit analyses can lead to reasonable evacuation alarm levels during a foreshock-aftershock sequence (although their study was based on a single, hand-constructed sequence rather than an operational system). The point is that the various components required for OELF appear to be sufficiently mature in terms of providing potentially useful systems.

The potential value added by the UCERF3-ETAS-based OELF system presented here is not only the inclusion of fault-based information, including finite ruptures and elastic-rebound effects, but also loss estimates that are based on synthetic catalog simulations (a.k.a. stochastic event sets). The latter implies that rather than providing only a mean expected loss following a main shock of interest (e.g., 5 billion dollars), we generate a full probability
distribution of potential losses given all possible sequences of events (rather than computing loss from the average event rate implied by all possible sequences). As discussed by Field et al. (2016), the mean loss value will almost never be realized, and a practitioner might actually be more concerned about exceeding some particular value, such as their level of reinsurance coverage.

**UCERF3-ETAS**

UCERF3-ETAS was developed by the Working Group on California Earthquake Probabilities (WGCEP; [http://www.WGCEP.org](http://www.WGCEP.org)), which is responsible for developing authoritative earthquake forecasts for California on behalf of the USGS and the California Geological Survey (CGS). All UCERF3 components were developed in partnership with the Southern California Earthquake Center (SCEC), and the California Earthquake Authority provided financial support. UCERF3-ETAS is actually built upon an elastic-rebound-based, time-dependent forecast (UCERF3-TD; Field et al., 2015, and references therein), which in turn is built upon a long-term time-independent model (UCERF3-TI; Field et al., 2014, and references therein). We give only a brief summary of these forecast models here, as details can be found in the associated main reports.

The time-independent model (UCERF3-TI) provides the long-term rate of all possible earthquakes throughout the region (above a magnitude threshold and at a specified level of discretization), both on and off explicitly modeled faults (Figure 1a). The primary achievements embodied in UCERF3-TI are a relaxation of fault-segmentation assumptions and the inclusion of multi-fault ruptures, both of which were acknowledged limitations of the previous model (UCERF2). The rates of all earthquakes were solved for simultaneously, and from a broader range of data, using a system-level “grand inversion” that is both conceptually simple and extensible. This new approach is more derivative and less prescriptive than that taken previously; for example, rather than assuming a magnitude-frequency distribution (MFD) on most faults, the inversion solves for the MFD that is most consistent with available data. The inverse problem is generally large and underdetermined, so a range of solutions was sampled using an efficient simulated annealing algorithm (Page et al., 2014). An interesting and important finding is that individual faults are generally inconsistent with the Gutenberg-Richter distribution, meaning most faults exhibit a characteristic MFD in terms of
having relatively high rates at higher magnitudes. The model also made more explicit use of geodetic data via three new deformation models, which not only provide alternative fault slip-rate constraints, but also enabled the inclusion of 150 fault sections that were previously excluded due to lack of geologic data. These additions served to fill out and expose the interconnectivity of the fault system, thereby revealing more multi-fault rupture possibilities. The number of fault-based ruptures increased from 10,000 in UCERF2 to more than 250,000 in UCERF3.

Overall, UCERF3-TI has a lower rate of $M_{6.5-7.0}$ earthquakes, reflecting an explicit regional MFD constraint added to avoid a UCERF2 over-prediction at these magnitudes. The rate of larger, multi-fault ruptures generally increased as a consequence, reflecting a tradeoff that was effectively brokered by the grand inversion in satisfying all data constraints. However, the rate of multi-fault ruptures was also effectively minimized by several other constraints, so if anything the model under predicts the frequency of such events (Field et al., 2014). Various epistemic uncertainties are represented in UCERF3-TI via 1,440 logic-tree branches, each of which represents a viable model with an associated probability of being correct. Following an extensive review process, UCERF3-TI was adopted for the 2014 update of the USGS National Seismic Hazard Maps (Petersen et al., 2015; Powers and Field, 2015).

Building on UCERF3-TI, the WGCEP subsequently defined a time-dependent model (UCERF3-TD) that uses renewal models to represent elastic-rebound-implied rupture probabilities (Field et al., 2015, and references therein). A new methodology was developed, which solves applicability issues in the previous UCERF2 approach with respect to un-segmented models. The new algorithm also supports magnitude-dependent aperiodicity in the renewal model, whereby the timing of smaller magnitude events is somewhat less periodic than that of larger ones. It also accounts for the historic open interval on faults that lack a date-of-last-event constraint, thereby allowing time-dependent probabilities to be defined for all fault-based ruptures. Compared to the time-independent model, and as shown in Figure 1b, UCERF3-TD probabilities are relatively low on faults where a large event has recently occurred, and relatively high (by up to a factor of ~2.5) where the time since last event is greater than roughly half the average recurrence interval. In other words, it reflects what we would expect from an elastic-rebound perspective. Epistemic uncertainties are represented
via four different levels of time-dependent predictability, or aperiodicity, leading to a total of 5,760 logic-tree branches for UCERF3-TD.

The third UCERF3 forecast model is UCERF3-ETAS (Field et al., 2017, and references therein), which attempts to model aftershocks and otherwise triggered events by merging UCERF3-TD with a state-of-the-art spatiotemporal clustering component. The latter is the Epidemic Type Aftershock Sequence model (ETAS; Ogata, 1988), which represents a generalization of classic Omori-Utsu aftershock statistics in that every earthquake can trigger its own aftershocks. ETAS thereby implies that many events in a sequence are triggered indirectly (as aftershocks of aftershocks), which produces richer and more realistic triggering behavior, such as the possibility that the 1999 Hector Mine earthquake was ultimately triggered by the 1992 Landers event (Hardebeck, 2013).

UCERF3-ETAS therefore represents the first complete forecast model that combines finite-faults and an elastic-rebound-based renewal model with a spatiotemporal clustering component. Several challenges and surprises were encountered in developing UCERF3-ETAS, including the need for both elastic rebound and characteristic MFDs when including finite faults. That is, excluding elastic rebound causes large aftershocks to simply re-rupture the main shock rupture surface, at least much more often than is seen in nature. Likewise, characteristic MFDs are generally needed in order to satisfy the intuitive expectation that large-event triggering probabilities, given a nearby smaller event, should be relatively high near known active faults (Michael 2012). UCERF3-ETAS is simulation based, meaning it generates one or more synthetic catalogs of $M \geq 2.5$ events for a given timespan, conditioned on the triggering potential of all $M \geq 2.5$ earthquakes known (or hypothesized) to have occurred prior to the forecast start time.

The UCERF3-ETAS report discusses results with respect to both long-term (1,000-year) simulations, as well as for 10-year time periods following a variety of hypothetical scenario main shocks. While the results are both realistic and plausible, they are not always consistent with the simple notion that triggering probabilities should always be greater if a main shock is located near a fault. Important factors include the degree of characteristicness in the MFD of a nearby fault, as well as whether large triggered events can nucleate from within the rupture zone of a large main shock. UCERF3-ETAS also represents various approximations and multiple sources of uncertainty, as will any candidate OEF model, which is why potential usefulness always needs to be considered in the context of specific applications. The purpose
of this paper is, again, to begin exploring UCERF3-ETAS usefulness is the context of statewide loss estimates. To this end, we use only the single branch-averaged model utilized and described in the UCERF3-ETAS report (Field et al., 2017), which is also shown in Figure 1 here.

The one epistemic uncertainty we do account for in this study is whether or not a large fault-based aftershock can be triggered by aftershocks that occur inside the rupture zone of the triggering event, or whether only aftershocks outside this zone can do such triggering, as described in detail by Field et al. (2017). We refer to this branch option as the “triggering uncertainty” in what follows, with the two options being equally weighted in the averages presented below.

**STATEWIDE LOSS CALCULATIONS**

The goal of our OELF system is to provide the distribution of possible losses for any specified timespan, condition on the $M \geq 2.5$ seismicity known to have occurred up to some point in time. This is a very challenging and complicated problem to get exactly right. For example, damage from one large event may change the distribution of assets, their vulnerabilities, and/or their values. Uncertainties also abound throughout the model, a full accounting of which is not presently feasible. Our OELF system is therefore necessarily simplified, as will be any such model. The relevant question is whether the model is potentially useful given the known limitations, a topic we return to after describing the system and some results. At the very least, we can quantify the potential value of such a system assuming the present one is correct.

A key to computational efficiency is pre-computing the mean expected loss for every $M \geq 5$ rupture defined in UCERF3. The loss expected for a given earthquake catalog of events is then simply the sum of the mean loss expected for each event in the catalog (assuming the portfolio is static). If we have many realizations of possible catalogs, representing for example the different ways in which an aftershock sequence might play out, we can construct a probability distribution of potential losses.

The mean loss expected for each rupture in UCERF3 is computed using the methodology and *OpenSHA* tools described by Porter et al. (2012, 2017). We use a portfolio that approximates most of the California building stock, originally constructed using FEMA’s
Hazus-MH 2.1 software, but with some updates to reflect increased asset values and population growth (Porter et al., 2017). Specifically, for each census tract we have a list of asset/building types and the total aggregate replacement cost for each. We also have a vulnerability model for each asset type, which gives a distribution of losses for a given ground motion level or intensity measure (5% damped, 1-second spectral acceleration). Figure 2 shows the net value of assets in each census tract.

For a given rupture, we compute the intensity-measure mean and standard deviation at each census tract location using all five NGA-West2 ground-motion prediction equations (Abrahamson et al., 2014; Boore et al., 2014; Campbell and Bozorgnia, 2014; Chiou and Youngs, 2014; and Idriss, 2014) and two statewide Vs30 models to represent site effects (Wills and Clahan, 2006; Wald and Allen, 2007). The mean loss is then computed using the relevant vulnerability function and total replacement cost for each asset type in the given census tract. For the background/grided seismicity in UCERF3, we approximate the range of possible rupture surfaces using two perpendicular faults where the length is magnitude dependent. Computing the mean loss for every rupture in UCERF3 was time consuming (we used the Stampede supercomputer at the Texas Advance Computing Center), but it only had to be done once.

SIMULATION RESULTS

Our risk metric of interest here is statewide financial losses (cost to repair building damage) in California, and more specifically, anticipated losses conditioned on information about recent events. The potential usefulness of OELF depends on the difference between conditional and unconditional loss estimates, the latter being estimates made in the absence of any information on past seismicity. In other words, there is tangible earthquake risk in California even in the absence of information on recent earthquakes. The relevant question is the extent to which risk fluctuates due spatiotemporal clustering, and whether UCERF3-ETAS predictions of this based on recent earthquakes can constitute useful information to decision makers.

UNCONDITIONAL LOSS ESTIMATES
Figure 3 shows the statewide, unconditional loss exceedance curve derived from 1,000 1,000-year UCERF3-ETAS simulations (that is, based on 1 million yearly samples). Again, unconditional here means the estimate when we lack information on recent events. The dark shaded region in this and other figures represents the range of values implied by the one epistemic uncertainty considered here (the triggering-uncertainty branch discussed above) with the solid line giving the mean based on equal branch weights. The light shaded region in this and other figures represents sampling uncertainties, which are due to the fact that we are inferring statistics from a finite number of simulations (see caption for details). The implied mean annual loss is $4.0 billion (see figure for uncertainty), which is consistent with other estimates as discussed by Porter et al. (2017). More importantly, the curve gives us the probability of exceeding specific loss thresholds, such as a 1.8% chance of exceeding $50 billion. Tables 1 and 2 list the 1-day and 1-year exceedance probabilities, respectively, for other loss thresholds, as well as the uncertainties on these estimates. One might alternatively be interested in the dollar amount that has a certain chance of exceedance. For example, a statewide insurer may have a stipulation that reinsurance coverage shall be based on the 1% annual exceedance probability, which would be $76 billion dollars according to the unconditional losses in Figure 3.

As far as we know, Figure 3 represents the first statewide loss exceedance curve that includes spatiotemporal clustering, or at least the first that includes finite faults and elastic-rebound effects. How important is the spatiotemporal clustering? Figure 4 compares the 1-year unconditional loss exceedance curve with that obtained after randomizing the origin time of each event in each catalog (a uniform probability distribution over the simulation duration, thereby imposing a Poisson distribution). As expected, spatiotemporal clustering widens the tail, giving rise to elevated probabilities above ~$20 billion and lower exceedance likelihoods below, with the latter being a consequence of the two having the exact same mean annual loss.

CONDITIONAL LOSS ESTIMATES

Losses based on actual historical seismicity. The unconditional exceedance curves in Figure 3 are based on long-term (1000-year) simulations, meaning the influence of actual recent seismicity is negligible. We also ran 100,000 1-year simulations with a start year of 2012 to see if estimates based on actual historical seismicity differ, with 2012 representing
the time at which the carefully vetted UCERF3 $M \geq 2.5$ historical catalog ends (Felzer, 2013). No significant differences were found.

**Losses from an $M 7.1$ Hayward-fault scenario.** Conditional losses will obviously be elevated following larger earthquakes because the number of trigged events increases exponentially with magnitude. Based on 200,000 UCERF3-ETAS simulations, Figure 5 shows exceedance curves following an $M 7.1$ earthquake on the Hayward fault, which represents the UCERF3 rupture that most closely matches the USGS “HayWired” planning scenario (https://www2.usgs.gov/natural_hazards/safrr/projects/haywired.asp). The mean expected loss in the year following the scenario is $24$ billion, a factor of $6$ greater than the unconditional value. The unconditional loss exceedance curves (from Figure 3) are shown for comparison, implying significant probability gains following the Hayward scenario. Specifically, the mean gain is a factor of $716$ and $7.09$ in the first day and first year, respectively, at the $50$ billion threshold (gains and associated uncertainties are listed for this and other loss thresholds in Tables 1 and 2). We also note that while the loss estimates are admittedly uncertain, as will be discussed more below, the gains are presumably less so because taking ratios should provide some bias correction. Figure 6 shows how the $50$ billion exceedance probability gain decreases with time following the event, effectively trending back toward unconditional losses as aftershocks decay away. The $1\%$ chance of exceedance value is $138$ billion in the year following the scenario, a factor of $\sim 1.8$ greater than the unconditional value ($76$ billion). To give an indication of which neighboring fault ruptures are contributing to these losses, Figure 1c shows the implied $M \geq 6.7$ probability gain of each fault in the year following the scenario (the fraction of times each fault participated in an aftershock divided by its long-term participation rate). Alternatively, Figure 7 shows the average number of $M \geq 2.5$ aftershocks in a 10-year period following the main shock, compared to an ETAS model that lacks faults, clearly illuminating the influence of finite faults in UCERF3-ETAS.

**Losses from an $M 7.0$ Mojave SAF scenario.** Figure 8 is similar to Figure 5, but for an $M 7.0$ scenario on the Mojave section of the southern San Andreas Fault (SAF), details of which can be found in Field et al. (2017). The mean expected loss in the year following the scenario is $12$ billion, a factor of $3$ greater than the unconditional value. The gain at $50$ billion is a factor of $560$ and $3.85$ in the first day and first year following the scenario, respectively (see Tables 1 and 2 for other threshold values). The $1\%$ chance of exceedance
value is $130 billion in the first year, a factor of ~1.7 greater than the unconditional value. These gains are less than for the Hayward scenario, which makes sense given the relative proximity of these events to populated areas.

**Losses implied by the 2009 M 4.8 Bombay Beach earthquake.** As mentioned in the introduction, CEPEC convenes on behalf of CalOES when earthquakes of triggering concern occur. One such event was the 2009 M 4.8 Bombay Beach earthquake (Jordan and Jones, 2010), located about 4 km from the Coachella Valley section of the SAF, the latter of which is considered to be loaded in an elastic-rebound sense (Field et al., 2015). Do such events have a perceptible influence on statewide loss estimates? Figure 9 attempts to answer this question based on 200,000 UCERF3-ETAS simulations, revealing no significant gains in the year following the event, but up to a factor-of-three gain in the first day (although significance becomes questionable above about $30 billion). If we restrict the loss calculations to only include census tracts that are near the main shock, specifically those within 20 km of the town of Rancho Mirage, CA, 1-year and 1-day gains increase to about a factor of ~2 and ~30, respectively, reflecting expected sensitivity to the geographic spread of the portfolio.

**DISCUSSION AND CONCLUSIONS**

UCERF3-ETAS represents the first candidate OEF model to include fault-based information, and it implies that including elastic-rebound is required to get realistic triggering behavior, both of which have been left out of previous OEF applications. Nevertheless, UCERF3-ETAS is only an approximation of the real system, meaning it is ultimately wrong at some level, as will be true for any future OEF replacement. The question is whether UCERF3-ETAS is right enough to be useful, and useful enough to be worth operationalizing, especially given the latter will require a significant commitment in terms of personnel and technological resources, including on-demand access to high-performance computing.

The question of whether UCERF3-ETAS is potentially useful can only be answered in the context of specific applications; these include public-preparedness activities, emergency management, building inspection and tagging, zoning and building codes, oil and gas regulation, and insurance and capital markets. The net value of UCERF3-ETAS will also depend on the specific application. In fact, UCERF3-ETAS may be overkill for some uses
where, for example, only a magnitude-probability distribution of triggered events is needed. The inclusion of finite faults also adds considerable complexity, and while we believe this makes the model more scientifically credible, it does not necessarily imply that fault-free models will always be less useful or misleading.

This study represents a first attempt at UCERF3-ETAS valuation in the context of statewide loss estimates, particularly with respect to the probability gains implied by aftershocks of various scenarios (relative to long-term or unconditional loss probabilities). Again, many modeling uncertainties have been ignored, including the 5760 UCERF3-TD logic tree branches, alternative ground-motion prediction equations and site-effect models, and the influence of correlated ground motions. We have also ignored uncertainties in the portfolio contents, asset values, and vulnerability functions, as well as how these might change due to ongoing damaging earthquakes. This list of modeling uncertainties could go on (e.g., demand surge, influence of code revisions), and we will indeed need to grapple with the influence of these factors before deeming OEF information useful or valuable. This will take considerable work, however, so it makes sense to pause and explore potential usefulness assuming our existing model is correct; if there is sufficient value under this assumption, or if we can quantify the conditions under which sufficient value would be obtained, then additional effort toward operationalization might be warranted.

The main factor that dictates potential usefulness is the probability gain relative to unconditional values. That is, if short-term OEF estimates do not deviate from long-term averages, then why bother with the former? Rather than reiterating the conditional loss estimates and gains implied by the OELF modeling presented here, we instead turn the questions back to anyone concerned about statewide loss estimates. In fact, we can pose the same set of questions to anyone interested in any other risk or loss metric, such as the collapse probability of San Bernardino city hall (which was closed for two days in 2016 due to another sequence of earthquakes near Bombay Beach, CA; http://www.ci.san-bernardino.ca.us/civica/press/display.asp?layout=1&Entry=23). Specifically, for any given risk or loss metric we ask the following questions:

1. How high would the probability gain need to be, relative to long-term (unconditional) probabilities, in order for OEF information to be actionable to you?

2. What timeframes following significant events are you most interested in?
3. Can you quantify the potential monetary value of this OEF information?

The last question is posed not because the USGS might charge for results, but rather to ascertain whether the benefit-cost ratio justifies operationalization. Although the statewide loss modeling results presented here represent a first step toward valuation, being more specific would require additional information. For example, an insurance company would need to consider things like the cost of changing reinsurance levels, as well as whether triggered events are considered part of the main shock in terms of policy deductibles. Probability gains will also be sensitive to the distribution of assets, as an $M_5$ event may have a negligible influence on statewide loss estimates, but significant gains for more localized assets such as San Bernardino city hall.

Question (2) above is important given the rapid decay of probability gains with time (e.g., Figure 6). While electronic trading might enable taking advantage of the highest gains immediately following a main shock, most other decisions will take more time to make and implement. Our calculations have also assumed zero latency, meaning results can be generated and delivered immediately, which again is a best-case assumption. We estimate that the types of conditional loss estimates presented here should take a matter of minutes to generate provided enough computer processors are available, which should not be a problem given currently available cloud-computing suppliers, although cost may be an issue.

In conclusion, we now have a candidate, fault-based OEF model for California, and we have hereby exemplified its potential usefulness in the context of statewide loss estimates. Determining whether it is worth operationalizing will now depend on potential users quantifying potential value in the context of their particular risk metric(s), and assuming the current model is correct. If these valuation efforts prove promising, then the next step will be to explore the extent to which existing epistemic uncertainties erode this inferred usefulness, including those related to loss modeling. There are also many ways in which UCERF3-ETAS could be improved, so these valuation analyses should help elucidate which modifications would be most worth the effort.

ACKNOWLEDGMENTS

This study was supported by the Southern California Earthquake Center under USGS-SCEC Cooperative Agreement G12AC20038 and NSF-SCEC Cooperative Agreement EAR-
1033462. Calculations were performed at the Texas Advanced Computing Center and the USC Center for High-Performance Computing and Communications. We also thank the USGS Powell Center for Analysis and Synthesis for supporting workshops on UCERF3-ETAS development. Southern California Earthquake Center Contribution #?????. All data reported in this paper are available at http://wgcep.org/ucerf3. Any use of trade, firm, or product names is for descriptive purposes only, and does not imply endorsement by the U.S. Government.

REFERENCES


Campbell, K.W. and Bozorgnia, Y., 2014. NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. Earthquake Spectra 30, 1087-1115.


### Tables

**Table 1.** One-year loss exceedance probabilities for the unconditional case (where no information on prior events is available) and in a year immediately following the various earthquake scenarios evaluated in this study (see text for details). Probability uncertainties (min, max) represent ~95% confidence bounds. Gains for the scenarios are defined relative to unconditional loss estimates, where the uncertainties (min, max) represent ~95% confidence bounds.

<table>
<thead>
<tr>
<th>$\text{Loss (Billions)}$</th>
<th><strong>Unconditional</strong></th>
<th><strong>M 7.1 Hayward</strong></th>
<th><strong>M 7.0 Mojave</strong></th>
<th><strong>M 4.8 Bombay Beach</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Probability (min, max)</td>
<td>Probability (min, max)</td>
<td>Gain (min, max)</td>
<td>Probability (min, max)</td>
</tr>
<tr>
<td>≥5</td>
<td>1.28E-01 (1.2E-01, 1.3E-01)</td>
<td>8.59E-01 (8.4E-01, 8.7E-01)</td>
<td>6.69 (6.58, 6.80)</td>
<td>3.86E-01 (3.0E-01, 4.8E-01)</td>
</tr>
<tr>
<td>≥10</td>
<td>7.63E-02 (7.2E-02, 8.1E-02)</td>
<td>6.49E-01 (6.1E-01, 6.9E-01)</td>
<td>8.50 (8.35, 8.64)</td>
<td>2.20E-01 (1.3E-01, 3.1E-01)</td>
</tr>
<tr>
<td>≥25</td>
<td>3.58E-02 (3.3E-02, 3.9E-02)</td>
<td>2.92E-01 (2.1E-01, 3.8E-01)</td>
<td>8.17 (6.20, 9.90)</td>
<td>1.13E-01 (4.9E-02, 1.8E-01)</td>
</tr>
<tr>
<td>≥50</td>
<td>1.78E-02 (1.6E-02, 2.0E-02)</td>
<td>1.26E-01 (6.0E-02, 1.9E-01)</td>
<td>7.09 (3.74, 9.91)</td>
<td>6.86E-02 (2.1E-02, 1.2E-01)</td>
</tr>
<tr>
<td>≥100</td>
<td>6.02E-03 (5.0E-03, 7.0E-03)</td>
<td>2.84E-02 (9.2E-03, 4.8E-02)</td>
<td>4.72 (1.75, 7.07)</td>
<td>2.06E-02 (6.4E-03, 3.5E-02)</td>
</tr>
<tr>
<td>≥150</td>
<td>2.14E-03 (1.7E-03, 2.6E-03)</td>
<td>7.31E-03 (1.9E-03, 1.3E-02)</td>
<td>3.42 (1.07, 5.22)</td>
<td>6.71E-03 (2.2E-03, 1.1E-02)</td>
</tr>
<tr>
<td>≥200</td>
<td>9.45E-04 (7.3E-04, 1.2E-03)</td>
<td>2.06E-03 (6.7E-04, 3.6E-03)</td>
<td>2.18 (0.83, 3.30)</td>
<td>2.97E-03 (9.3E-04, 5.1E-03)</td>
</tr>
</tbody>
</table>

1. The min and max are based on the extreme 68% confidence bounds, due to sampling errors discussed in Figure 3 caption, from the two logic-tree branches considered separately; because values from each branch are significantly different, and each branch is given an equal weight of 0.5, the 32% chance of being outside the range becomes 16%, leading to a 84% confidence range.

2. Based on the 68% confidence bounds from the scenario and unconditional estimate sampling errors. i.e., the min value is the lower bound of the scenario loss divided by the upper bound for the unconditional estimate, and since each has a 0.16 probability of being lower or greater, respectively, the combined probability is 0.16*0.16 = 0.0256 (~0.025%); likewise for the max value, and thus the 95% confidence range.
Table 2. Same as Table 1, but for one day rather than one year

<table>
<thead>
<tr>
<th>$\text{Loss (Billions)}$</th>
<th>$\text{Unconditional}$</th>
<th>$\text{M 7.1 Hayward}$</th>
<th>$\text{M 7.0 Mojave}$</th>
<th>$\text{M 4.8 Bombay Beach}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\text{Probability (min, max)}$</td>
<td>$\text{Probability (min, max)}$</td>
<td>$\text{Gain (min, max)}$</td>
<td>$\text{Probability (min, max)}$</td>
</tr>
<tr>
<td>$\geq 5$</td>
<td>3.71E-04 (3.4E-04, 4.0E-04)</td>
<td>4.24E-01 (3.9E-01, 4.5E-01)</td>
<td>1142 (1035, 1253)</td>
<td>9.27E-02 (3.9E-02, 1.5E-01)</td>
</tr>
<tr>
<td>$\geq 10$</td>
<td>2.14E-04 (1.9E-04, 2.4E-04)</td>
<td>2.08E-01 (1.6E-01, 2.5E-01)</td>
<td>971 (756, 1194)</td>
<td>5.08E-02 (1.2E-02, 9.0E-02)</td>
</tr>
<tr>
<td>$\geq 25$</td>
<td>8.51E-05 (7.5E-05, 9.7E-05)</td>
<td>8.13E-02 (3.7E-02, 1.3E-01)</td>
<td>955 (399, 1624)</td>
<td>3.09E-02 (4.8E-03, 5.7E-02)</td>
</tr>
<tr>
<td>$\geq 50$</td>
<td>3.27E-05 (2.6E-05, 4.1E-05)</td>
<td>2.34E-02 (8.6E-03, 3.9E-02)</td>
<td>716 (227, 1355)</td>
<td>1.83E-02 (2.4E-03, 3.5E-02)</td>
</tr>
<tr>
<td>$\geq 100$</td>
<td>5.64E-06 (3.1E-06, 9.5E-06)</td>
<td>3.30E-03 (3.0E-04, 6.5E-03)</td>
<td>585 (31.2, 2072)</td>
<td>3.77E-03 (6.4E-04, 7.1E-03)</td>
</tr>
<tr>
<td>$\geq 150$</td>
<td>8.25E-07 (1.6E-07, 3.1E-06)</td>
<td>2.80E-04 (2.2E-05, 6.0E-04)</td>
<td>339 (7.50, 3659)</td>
<td>9.74E-04 (1.8E-04, 1.9E-03)</td>
</tr>
<tr>
<td>$\geq 200$</td>
<td>2.20E-07 (0.0E+00, 2.5E-06)</td>
<td>4.34E-05 (6.9E-06, 1.1E-04)</td>
<td>198 (3.68, *)</td>
<td>3.69E-04 (5.3E-05, 7.4E-04)</td>
</tr>
</tbody>
</table>
**FIGURES**

**Figure 1.** Three-dimensional perspective views of the third Uniform California Earthquake Rupture Forecast (UCERF3), based on the branch averaged model described and utilized by Field et al. (2017). The small black parallelograms represent the 2606 fault subsections utilized in the forecast. (a) The probability that each fault, or 0.1 by 0.1 degree geographic cell, will participate in one or more M≥6.7 earthquakes in a given year according to the long-term, time-independent model (UCERF3-TI). (b) The probability gain in 2012, relative to (a), that each fault will participate in an M≥6.7 according to the elastic-rebound component in the long-term time-dependent model (UCERF3-TD). (c) The 1-year M≥6.7 probability gain implied by an M 7.1 scenario on the Hayward fault (assumed to have occurred at the beginning of 2012), as inferred from 200,000 UCERF3-ETAS simulations. The spotted nature of gains away from the main shock reflect the finite number of simulations, as the result would be smoother if an infinite number were available.
Figure 2. Map of asset-value (replacement-cost) density in each census tract.
Figure 3. Unconditional (i.e., long-term average) statewide loss exceedance curves for various timeframes (as labeled) based on 1,000 1,000-year UCERF3-ETAS simulations. The upper and lower limits for the dark-shaded regions represent results for the two “triggering uncertainty” logic-tree branches considered separately, and the solid lines represent the equally weighted average. The light shaded region represents sampling uncertainty (due to probabilities being inferred from a finite number simulations); specifically, the error represents the 68% confidence interval for the true fraction of exceedances assuming a binomial distribution, and using the "Wilson score interval with continuity correction" developed by Newcombe (1998). The two branches were treated separately, so the light shaded region represents an 84% confidence bound (see note in Table 1 for details). Uncertainties on the mean annual loss labeled here are also 84% confidence bounds for similar reasons.
**Figure 4.** The influence of removing spatiotemporal clustering with respect to unconditional, annual loss exceedance probabilities. The solid line and shaded regions represent results equivalent to those shown in Figure 3, but after randomizing all event times to invoke a Poisson temporal distribution. Shown for comparison are results from Figure 3 (including spatiotemporal clustering), but where the solid line was replaced with dot-dashed and the shaded regions were replaced with dashed and dotted lines to improve overlap visibility.
Figure 5. Loss exceedance probabilities in 1-year (a) and 1-day (b) following the M 7.1 scenario earthquake on the Hayward Fault described in the text. Solid lines and shaded regions represent scenario estimates as described in Figure 3, and dot-dashed, dashed, and dotted lines are the unconditional exceedance curves from Figure 3. The vertical distance between the solid and dot-dashed lines represent the loss exceedance gain (increase in the probability of exceeding losses due to events possibly triggered by the scenario).
Figure 6. Decay of the $50 billion loss exceedance gain with increasing forecast timespans immediately following the M 7.1 Hayward fault earthquake scenario. Gain represents the ratio of the conditional probability (given the main shock) to the unconditional probability (when no information on past events is available). Shaded region represents 95% confidence bounds as described for gains in Table 1.
Figure 7. (a) The average number of M$\geq$2.5 aftershocks (in 2-by-2 km geographic cells) in 10 years following the M 7.1 Hayward fault scenario, as inferred from 200,000 UCERF3-ETAS simulations. (b) Results for a no-fault model, revealing the influence of faults in (a).
**Figure 8.** Same as Figure 5, but for the M 7.0 scenario earthquake on the Mojave section of the San Andreas Fault.

**Figure 9.** Same as Figure 5, but for the 2009 M 4.8 earthquake near Bombay Beach, CA, which is 4 km from the Coachella section of the San Andreas Fault.