The Uniform California Earthquake Rupture Forecast, Version 3 (UCERF3) Project Plan

by the Working Group on California Earthquake Probabilities (WGCEP)

Notes on this version (relative to v52, which was reviewed by SRP in June, 2011):

- *This represents revisions based on the SRP Review (CEA Report #4) and comments from CEA.*
- *The two biggest changes are more specificity of the Deformation Models and Earthquake Rate Models (including anticipated logic-tree branches).*
- *Changes have not been tracked, although one can always compare documents (Ned Field can provide this upon request).*

Table of Contents

Introduction

The development of the Uniform California Earthquake Rupture Forecast, version 3 (UCERF3) is a project of the Working Group on California Earthquake Probabilities (http://www.WGCEP.org), jointly sponsored by the U.S. Geological Survey (USGS), California Geological Survey (CGS), and the Southern California Earthquake Center (SCEC) and partially supported under a SCEC contract with the California Earthquake Authority (CEA).

The UERF3 project will build on the Uniform California Earthquake Rupture Forecast, version 2 (UCERF2), which is a long-term, time-dependent model that was developed by our WGCEP consortium and released to the public on April 14, 2008. The primary goal of the project is improve the UCERF probabilistic framework by incorporating multi-fault ruptures and spatiotemporal clustering. The latter will extend the model into the realm of the short-term probabilities needed for operational earthquake forecasting, and its use for this purpose will thus require robust interoperability with real-time seismicity information.

This document outlines our currently envisioned UCERF3 model, as well as anticipated issues and a research plan for addressing them (in the form of a list of tasks and planned workshops). The appendix here provides more details for those tasks that warrant further discussion at this time (but only if elaboration beyond the main text is in order). This plan is subject to change as the project evolves.

Background

In the UCERF2 project, WGCEP developed a statewide model that uses consistent methodologies, data-handling standards, and treatment of uncertainties across all regions of California. A noteworthy accomplishment was the coordination of UCERF2 development with the National Seismic Hazard Mapping Project (NSHMP); consistency was ensured by using the time-independent version of UCERF2 in the 2008 USGS national seismic hazard maps. The main UCERF2 report as well as its 16 appendices, executive summary, supplemental data, press release, and a fact sheet are publicly available at http://www.SCEC.org/ucerf.

A more extensive analysis of the historical earthquake catalog in the development of UCERF2 revealed that the previous USGS national hazard map model (NSHMP, 2002) significantly over-predicts the rate of earthquakes near magnitude 6.5. This discrepancy was reduced to within the 95% confidence bounds of the observations by adjusting parameters in the UCERF2 model. However, most working-group participants believed that a better solution could be obtained by changing fundamental aspects of the model. For example, the actual cause of the M 6.5 discrepancy may be the assumptions regarding fault segmentation and the lack of fault-to-fault ruptures. If true, then UCERF2 not only over predicts the probability of intermediate-sized events (near M 6.5), but also under predicts the frequency of larger ($M \geq 7$) earthquakes, which could have a significant impact on both hazard and loss estimates.

The working group identified the following problems associated with UCERF2:

- **Interpretation of the "Empirical Model"** – WGCEP (2003) interpreted the apparent recent seismicity lull as a stress shadow cast by the great 1906 event, but the fact that most of the state exhibits an apparent lull calls this interpretation into question. This issue represents the single largest epistemic uncertainty for time-dependent probabilities in UCERF2.
- **Relax Segmentation & Include Fault-to-Fault ruptures** – Fault-to-fault ruptures, like the 2002 Denali earthquake, are not included in UCERF2. As discussed above, their inclusion might solve our remaining M 6.5 over prediction (and a likely M≥7 under prediction).
- **Self-consistent, Elastic-Rebound-Theory Motivated Renewal Models** – Inclusion of multi-segment ruptures, or relaxing segmentation altogether, introduced unresolved conceptual problems in computing conditional time-dependent probabilities.
- **Include Earthquake Triggering and Clustering** – UCERF2 does not include any type of triggering (e.g., as caused by static or dynamic stress changes, or as represented by aftershock statistics). Some believe that these effects are more important than the time dependence presently included in UCERF2, especially if a moderate or large event were to occur.
- **Extent of Earthquake Ruptures with Depth** – Both state-of-the-art earthquake forecast models (like UCERF2) and ground-motion simulations (like SCEC's CyberShake) depend heavily on magnitude-area relationships, and those currently available have big and important differences that must be resolved with respect to the depth extent of large ruptures (e.g., UCERF2 and CyberShake use incompatible models). Closely related to this are the quantification of seismogenic depth, aseismicity, and coupling coefficients (and the magnitude dependence of these).

These problems, which are discussed extensively in the UCERF2 report, represent opportunities for improvement in UCERF3, and each has motivated aspects of the UCERF3 implementation plan presented here.

Much effort in building UCERF2 was put into developing a computational infrastructure that is both modular (object-oriented) and extensible to UCERF3 (click "Model Framework" at http:www.WGCEP.org). We also developed distributed and electronically accessible data resources as well as analysis tools (click "Data" and/or "Tools" at http://www.WGCEP.org) based on the *OpenSHA* software. In short, we have developed a cyberinfrastructure within which we can extend the UCERF framework efficiently and effectively.

Implementation Plan

Epistemic Uncertainties and Logic Trees

Because there will be no single consensus model for UCERF3, it is important that the modeling framework adequately represent "epistemic" uncertainties (which represent our lack of understanding of how nature works, as opposed to "aleatory" uncertainties which represent the inherent randomness assumed in a given model). As with UCERF2, we will represent epistemic uncertainties in UCERF3 using logic-tree branches that account for multiple models constructed under different assumptions and constraints. The logic-tree branches that are currently anticipated for UCERF3 are discussed below. Final decisions, including branch weights, cannot be made until model components are more fully implemented and parameter sensitivity studies have been conducted.

Participants

The WGCEP organizational structure used for UCERF2 development will be maintained for UCERF3; it comprises an Executive Committee (ExCom), a Management Oversight Committee (MOC), and a Scientific Review Panel (SRP). Other WGCEP Contributors include research scientists, resource experts, model advocates, and IT professionals. The current membership of these various groups is given under "Participants" at http://www.WGCEP.org.

The ExCom, chaired by Dr. Field, is responsible for convening experts, reviewing options and making decisions about model components, as well as orchestrating implementation of the model and supporting databases. An important role of the ExCom is to ensure that the components of UCERF3 span the range of model viability. The MOC, chaired by Dr. Jordan, is in charge of allocating resources and approving project plans, budgets, and schedules; it is also responsible for seeing that the models are properly reviewed and delivered. The SRP, chaired by Dr. Ellsworth, is an independent body of experts that will review the development plans and model elements; in particular, they will evaluate whether the WGCEP has considered an adequate range of models to represent epistemic uncertainties. The SRP is participatory in the sense that it was convened at the very beginning of the project and will serve throughout the project period.

It is important to note that the separation of these roles will not always be maintained in an absolute sense. For example, given their expertise or experience, an SRP member may at times play an advocacy role with respect to a given model component. In such circumstances it will be important to identify which "hat" a participant is wearing. In general, the SRP will keep the ExCom in check with respect to any such conflicts of interest, and the MOC will keep the SRP in check.

Consensus Building

Discussion of model options and consensus building will be achieved through a series of community workshops described in the schedule outlined below. These workshops will include participants from the broader community in order to ensure views that go beyond the active

WGCEP participants. Some workshops will focus on the scientific ingredients going into UCERF3, while others will be aimed at informing and getting feedback from user communities.

Decisions with respect to logic-tree branches and weights are the responsibility of the ExCom. The ExCom will also provide the scientific rationale for why the models were selected and how the weights were assigned. The SRP will review the ExCom decisions. Interactions between the ExCom and SRP will be mediated by the MOC.

While the ExCom will need to rely on expert opinion in establishing some logic-tree branch weights, we intend to base these decisions on criteria that are as quantitative, reproducible, and testable as possible. The likelihood of a rupture jumping from one fault to another is an interesting example for the application of these principles. Ideally, we will develop and validate a formula that provides the jumping probability as a function of fault separation, relative geometry, sense of slip, slip rates, hypocenter, etc. Because no such formula has yet been validated, however, we may have to rely on expert judgment from a case-by-case analysis of each neighboring fault combination in California. A standardized formula for fault-jumping probability is clearly more objective than expert opinion, because the former can be formally tested in other active-faulting environments. Nevertheless, expert opinion is a powerful way of assimilating complex and diverse information about how faulting actually works.

Coordination with NSHMP

As with UCERF2, UCERF3 is being developed in full cooperation and coordination with the USGS National Seismic Hazard Mapping Program. It is WGCEP's goal that the time-independent version of UCERF3 be used for the next round of USGS hazard maps for California, which are scheduled for release circa 2013. Coordination will be facilitated by Ned Field's dual role as WGCEP chair and as USGS lead for the California part of the NSHMP forecast model.

Time Dependencies, Operational Aspects, and Potential Users

A particularly ambitious aspect of UCERF3 is to develop an operational earthquake forecast—an authoritative model that can be revised in near real time as significant events unfold. (Here "significant" means events that significantly modify estimates of subsequent earthquake probabilities.) WGCEP's goal is to construct a model that will produce forecasts across a wide range of time scales, from short term (days to weeks), through intermediate term (e.g., annual forecasts), to long term (decades to centuries). Short-term forecasts could be used, for example, to alert emergency officials of the increased hazard due to a moderate-sized earthquake occurring near a fault that is considered close to failure. Yearly forecasts could be used by homeowners to decide whether to buy earthquake insurance for the following year, or by those needing to price insurance premiums or cat bonds. Long-term forecasts are currently used in building codes.

Obtaining a full range of forecasts from a unified model would be an improvement over current practice in which the short-term and long-term forecasts are derived essentially independently. This is because there are significant dependencies between the parameters that control the results at different time scales. For instance, in an Epidemic Type Aftershock Sequence (ETAS) model, the long-term probabilities represented by the background rate of events trade off against the

aftershock productivity parameters that control the short-term probabilities. Also, while aftershock sequences are generally considered to be a short-term phenomenon, it has been demonstrated that they can produce significant probability changes over periods of years to decades. By considering all time dependencies within a single modeling framework, we will be able to develop a consistent set of forecasts.

The utility of UCERF3 will be dictated not only by the interests of user communities, but also by the confidence we have in the forecast given its uncertainties. Therefore, it will be important to have an ongoing dialogue between potential users and model developers throughout the project. This will help to clarify both their needs and our ability to deliver meaningful results. Use in earthquake insurance will certainly be a priority given CEA's involvement in the project. However, there are other potential uses; e.g., as a resource for the California Earthquake Prediction Evaluation Council (CEPEC) and National Earthquake Prediction Evaluation Council (NEPEC), which render advice on earthquake threats following significant events. The USGS currently makes short-term forecasts during earthquake clusters, which are used by the California Emergency Management Agency (CalEMA), other emergency responders, and utilities. UCERF3 will improve the basis for these short-term forecasts by making them consistent with the long-term seismic hazard model. The eventual utility of UCERF3 will be difficult to assess until it is actually deployed and the tools with which to explore loss implications become available. (The USGS ShakeMaps are an instructive example of a product whose utility was not fully anticipated in advance.)

While feedback from potential users and new research results may very well focus our efforts on particular aspects of short-term or long-term earthquake probabilities, there are also good scientific reasons for constructing a "broadband" time-dependent model. As discussed above, there is not a physically meaningful division between short-term and long-term forecasts. Building a system-level model that combines these forecasts will allow us to explore some of the primary scientific issues associated with time-dependent forecasting, such as the problem of distinguishing a multi-fault rupture from a series of separate, but quickly triggered, earthquakes.

Finally, it is important to emphasize that, while we strive to build a model that could be used for operational purposes, the agencies with statutory responsibilities for providing earthquake information have not yet committed to deploying such operational capabilities. The full scope of the technical and societal (e.g., legal) issues associated with the application of operational earthquake forecasting are just now being explored, and the resources for maintaining an operational system have not yet been appropriated.

Contingency Plans

As with any project with this level of ambition, it is likely that not all of our goals will be achieved by the final delivery date. Previous WGCEP efforts have repeatedly demonstrated the truth of two clichés, "it's easier said than done" and "the devil is in the details". In the worst case, WGCEP may conclude that the best available science has not yet validated a representation of multi-fault ruptures and/or spatiotemporal clustering adequate for operational purposes. The project plan has been developed to deal with these uncertainties. In particular, the UCERF3 logic-tree structure will be capable of dealing with de-scoping situations by using model branches developed for UCERF2 as fallbacks. More de-scoping options are given below.

CEA Delivery Schedule

June 30, 2010 - **Methodology Assessment – Issues and Research Plan (Report #1)**

Written report summarizing the status of the model components, a research plan for addressing outstanding questions and issues, and a preliminary implementation plan for the UCERF3 model. Report will provide details broken out by the main model components and/or by task, as deemed appropriate.

December 31, 2010 - **Methodology Assessment – Proposed Solutions to Issues (Report #2)**

Written report summarizing proposed solutions to the questions and issues identified in Report #1, and a revised implementation plan for the UCERF3 model. Report will provide details broken out by the main model components and/or by task, as deemed appropriate. Draft to SRP on Nov 5th (6 days before meeting), and final to CEA by Dec. 20th.

May 31, 2011 - **Proposed UCERF3 Plan (Report #3)**

Written report by WGCEP summarizing the proposed implementation plan for the UCERF3 model. This report will identify the remaining implementation issues requiring short-term, targeted research.

June 30, 2011 - **SRP Review of Proposed UCERF3 Plan (Report #4)**

Written report by the SRP that reviews the proposed UCERF3 implementation plan and recommends modifications.

September 30, 2011 - **Final UCERF3 Plan (Report #5)**

Written report by WGCEP that responds to the SRP review (as well as reviews by NEPEC, CEPEC, and CEA), provides a final implementation plan for the UCERF3 model, and summarizes progress towards implementation.

March 31, 2012 - **Preliminary UCERF3 Model (Report #6)**

Preliminary version of the UCERF3 model by WGCEP, implemented on the OpenSHA computational platform and documented in a written report.

April 30, 2012 - **Review of Preliminary UCERF3 Model (Report #7)**

Written report by the SRP that reviews the preliminary UCERF3 model and documentation and recommends modifications.

June 30, 2012 - **Final UCERF3 Model (Report #8)**

Final version of the UCERF3 model by WGCEP, implemented on the OpenSHA computational platform and documented in a written report. This final report will also include recommendations to CEA on the use of UCERF3, as appropriate, and recommendations on how UCERF3 can be improved by further research and development.

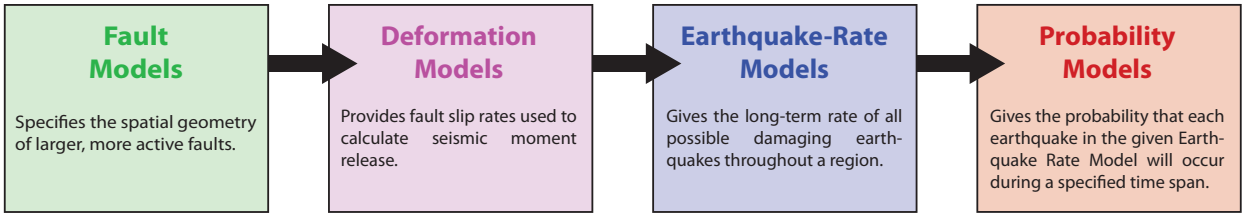
Main Model Components

UCERF3, like its predecessor, UCERF2, will be constructed from the four main model components shown and defined in Figure 1. Dividing any complex interactive system into separate components is to some degree artificial and arbitrary, but at this stage of model development these components are both meaningful and necessary. In the case of UCERF3, the most problematic distinction is between the Earthquake Rate and Probability models. All previous WGCEP and NSHMP forecast models have first defined the long-term rate of each event, which has both physical meaning (in terms of being, in principle, measurable) and practical utility (e.g., in current building codes). However, drawing this distinction can become problematic when constructing a model. For instance, and to reiterate the example given above, how will we differentiate between the rate of a particular multi-fault rupture and the probability that one fault might quickly trigger another as a separate event? Physics-based earthquake simulators, which are discussed below, do not make any modeling distinction between an earthquake rate and a probability component (although one may still need to infer long-term rates in order to apply the results). Therefore, the distinction between the Earthquake Rate and Probability models may become blurred as UCERF3 is developed.

One important reason for separating and encapsulating UCERF3 components is to make construction and maintenance more manageable. Understanding the details of every element of UCERF3 will go beyond the intellectual capacity of any one individual, so compartmentalization is critical. Another reason for modularization is to aid in defining alternative models (logic tree branches), and to have the ability to add new and improved components in the future.

Each of the four model components depicted in Figure 1 are described in more detail in the following four sections of this report. Some of these main components are composed of other sub-modules, which we refer to as "Key Components" of UCERF3. These are also described in the next four sections, and a complete list is given in the "Key Components" section below (some readers might find it helpful to review this section now). This latter is also where deadlines are defined.

Figure 1. Main Components of the UCERFs 2 & 3



Fault Models

Definition

A fault model gives the spatial geometry of the larger, known, and active faults throughout the region, with alternative models representing epistemic uncertainties. By definition, a fault model is composed of a list of fault sections, where each fault sections is in turn composed of:

- Fault Section Name (e.g., "San Andreas (Parkfield)")
- Fault Trace (list of lats, lons, and depths for the upper fault edge)
- Upper and Lower Seismogenic-Depth Estimates
- Average Dip Estimate
- Average Rake Estimate
- Fault Width (to indicate extent to which it's a proxy for a broader, braided zone of faults)

Distinct *Fault Sections* are defined only to the extent that one or more of these attributes vary along strike, so some fault sections can be quite long (e.g., the northern San Andreas Fault has only four sections). A *Fault Section* should not necessarily be considered a rupture segment (at least not in UCERF3). The complete master list of fault sections for California is found in our *Fault Section Database*, which is part of the *California Reference Geologic Fault Parameter Database*. Some fault section entries in this database are mutually exclusive (e.g., representing alternative representations, only one of which can be correct). Again, a fault model is simply a list of fault sections that is intended to be a complete, viable representation of the large, known, and active faults throughout the region.

Note that the "Fault Width" parameter above is new to UCERF3, with the goal being to enable more specificity with request to: 1) specifying what the slip rates actually apply to; and 2) answering questions like whether the 2010 El Mayor-Cucapah earthquake was an event on our Laguna Salada source (as opposed to being a background seismicity event).

Logic-Tree Branches

We anticipate having two alternative fault models: FM 3.1 and FM 3.2 (analogous to versions FM 2.1 and FM 2.2 used in UCERF2). These two models represent alternative representations for several faults. Reducing all possible combinations to just two models implies some correlation between the alternatives for different faults, whereas in reality no such correlation may exist. Trying to honor the uncorrelated nature of these alternatives would lead to an unmanageably large number of logic tree branches (and probably requiring Monte Carlo sampling). The two alternatives proposed here should be adequate in terms of hazard estimates, especially since the vast majority of users are interested in mean hazard.

Current Status

Here we provide a brief summary of the FM 3.1 and 3.2 development, with further details given in the appendix entry for Task F1 listed below. These models have been developed in coordination with the Statewide Community Fault Model (SCFM) project. The majority of

this effort has focused on updating and augmenting the inventory of faults in northern California. Due to the large number of Quaternary-active faults in northern California, a prioritized list was given to the SCFM group to help focus their effort. This list was based on identifying faults that potentially constitute block boundaries (for the deformation models), new faults that provide potential connections between existing faults, and faults with new data that warrant inclusion in the UCERF3 model. Some UCERF2 faults were also revised based on the relocated seismicity catalog of Waldhauser (2009).

These new representations were provided to Tim Dawson for review and integration into the UCERF3 *Fault Parameter Database*. In general, most SCFM additions and revisions were accepted, with a few exceptions where SCFM representations were less detailed than the existing UCERF2 model. An example of the latter is the Great Valley thrust system, where SCFM modeled it with two continuous fault zones whereas the UCERF database has 14 different fault sections (the connective nature of this fault system is still in question, as discussed in the appendix entry for Task F1 listed below).

In April 2011, two regional workshops convened by Tim Dawson, Andreas Plesch, and Ray Weldon, were held in Pomona and Menlo Park to present the fault models and to solicit suggestions from the community for improvements. This activity has so far resulted in the addition of ~60 new fault sections and the revision of 73 fault sections. The appendix entry for Task F1 includes a table listing exactly which faults have been added or revised, along with comments describing the changes.

One key assumption that will be made in UCERF3 is that fault endpoints are reasonably well constrained for purposes of quantifying multi-fault ruptures. Thus, a special effort has been undertaken to examine endpoints of faults in the database. In general, new fault representations from the SCFM are more connected to other faults than in past models. We have also examined and modified the endpoints of some other fault as well. Typically, this involves comparing the fault section trace to a Quaternary fault map, as very little other data is available to go by. This effort is ongoing.

De-scoping Options

We do not anticipate any major problems in developing these faults models, but if something arises, we always have FM 2.1 and 2.2 from UCERF2 to fall back on.

Ongoing Tasks

Table 1. Task List for the Development of Fault Models.
Task leaders are listed in **bold**; USGS Western Region participants are listed in red and USGS Central Region participants are listed in blue (not including Jones or Parsons). People listed as participants have expressed interest, though no contractual obligations are implied. Expanded descriptions of some tasks are given in the appendix, as noted; some may change and others be added as the project evolves. Some tasks overlap with those related to the other three model components.

	Task	Description	Leader & Participants
Fault Model(s)	**F1) Finalize Fault Models**	a) Add a width field to each fault section to indicate the extent to which it represents a narrow surface of deformation versus a broader, braided system of faults. b) Make plots showing the following: additions and modifications relative to FM 2.1 and 2.2 (from UCERF2); fault sections that are common between and unique to FM 3.1 and 3.2; comparison of FM 3.1 and 3.2 to the complete quaternary fault map (to indicate level of coverage, and whether fault sections represent distinct surfaces versus a braided fault zone). c) Finalize FM 3.1 and 3.2, including reevaluation of fault endpoints and indication of whether each fault section represents a well-defined, distinct surface or a broader deformation zone (i.e., a braided fault system). d) Review and publish a report. e) Finalize database implementation and integration with NSHMP. *See appendix for further discussion of this task.*	**Dawson**, Plesch, Weldon, J. Shaw, Haller, Grant, Powers.

Deformation Models

Definition

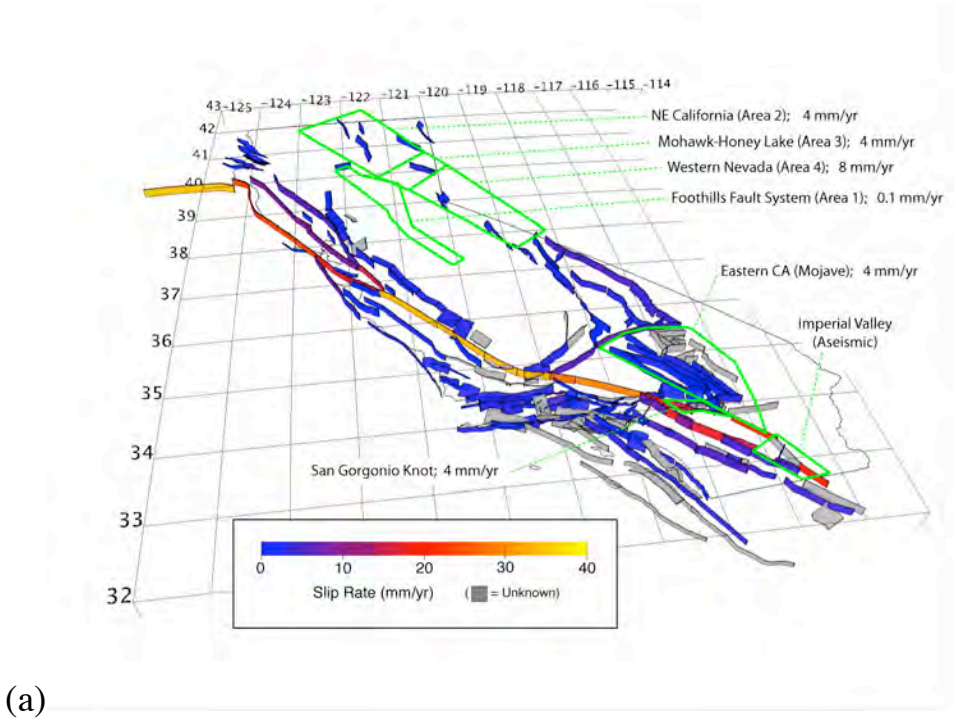
Each deformation model gives slip-rate estimates at various locations on a given fault model, plus deformation rates off the explicitly modeled faults (referred to as "off fault" deformation here, even though this is at least partially occurring on unknown faults). In the UCERF2 deformation models, a single slip-rate estimate was assigned to each fault section in a given fault model, and off-fault deformation was represented with a set of geographic polygons with an associated residual slip rate (e.g., Figure 2a). Another important deformation-model parameter is the aseismicity estimate, defined as the amount of moment between the upper and lower seismogenic depths that is released aseismically.

In UCERF3, we will be replacing Deformation Models 2.x with versions 3.x. UCERF2 slip-rates were assigned based on an expert-opinion evaluation of available data (mostly geologic and geodetic), together with summations across various transects to make sure the total plate tectonic rate was matched. In UCERF3, fault slip rates and off-fault strain rates estimated from inversions of GPS-derived velocities and geologic slip rates with kinematic models will replace the previous expert opinion rates (this was the topic of two workshops: April 1-2, 2010, and June 4, 2011). These models remove the need to introduce path-integral transects because the total plate rate is matched explicitly by the models.

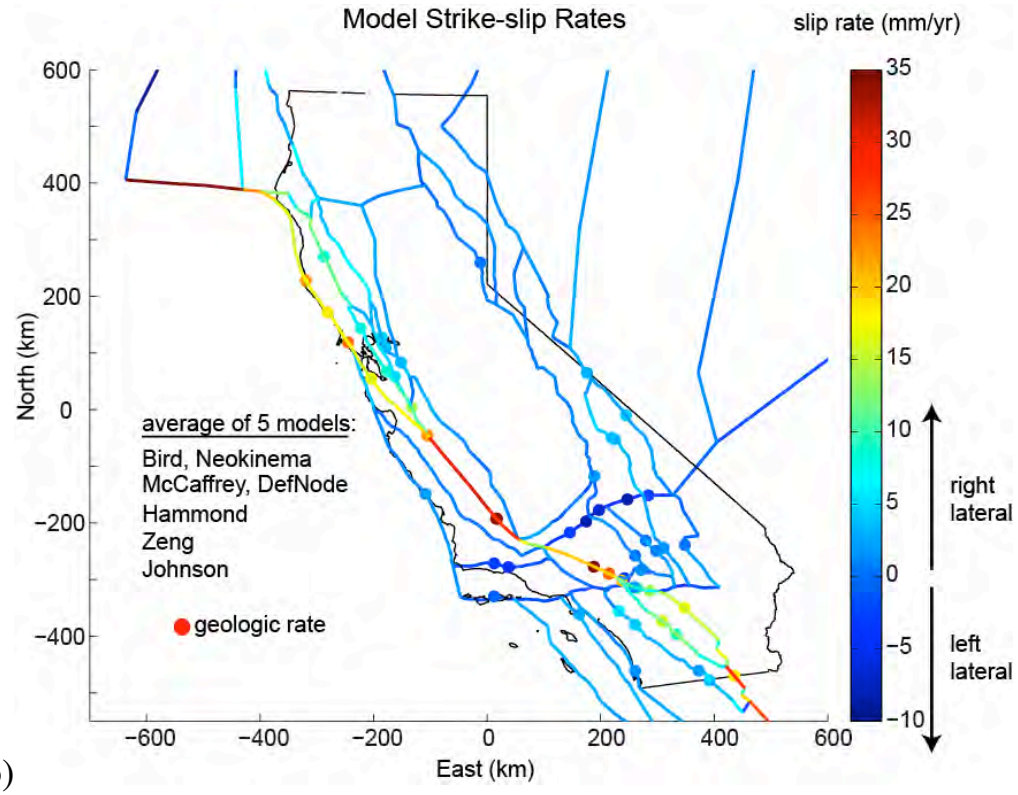
Deformation Models 2.x were constructed to explicitly match the total plate rates (within uncertainties), which basically assumed that all seismic deformation occurs on the modeled faults (i.e., assuming rigid blocks in between). However, there is reason to suspect that a non-negligible portion of the plate motion is accommodated off of the main faults. For example, Bird [2009] suggests that the off-fault deformation accounts for approximately 30% of the plate motion. McCaffrey [2005] showed that the GPS-derived velocity field and geologic slip rates in the western U.S. are best reconciled with models that include non-negligible off-fault strain within crustal blocks.

Deformation Models 2.x incorporated some off-fault deformation in the C-zone polygons illustrated in Figure 2a. Some of these C-zones encompass gray faults in Figure 2a that did not have assigned slip rates. Deformation Models 3.x will not have C-zones and all faults will be assigned slip rates. Rather than C-zones, Deformation Models 3.x will provide the strain rate tensor on a $0.1^\circ \times 0.1^\circ$ grid covering all of California. This grid of strain rates will account for all modeled deformation that is not accommodated on the faults.

The previous ambiguity with respect to whether slip rates represent deformation on the main fault surface, or whether they apply to a zone surrounding the fault, has now been removed by virtue of adding the *Fault Width* attribute to fault sections (described above); slip rates, by definition, apply over these specified widths. Although the kinematic models treat each fault as a narrow, discrete fault surface, the model-derived fault slip rates will be distributed across broader fault zones using well-established relationships between the width of elastic deformation across a fault zone and the depth of fault locking. From elastic dislocation theory, for a very long strike-slip fault, the width, W, of the shear zone surrounding a fault containing a specified fraction, F, of the total interseismic strain across the fault is $W = L*\tan(F*\pi/2)$.



(a)



(b)

Figure 2. **a)** Faults and slip rates for Deformation Model 2.1, plus the polygons (green) representing significant deformation elsewhere in the region. **b)** Average slip rate from 5 block models. Slip rate shown for each segment by color. Colored dots show geologic slip rate constraint (mean value) at the same color scale.

Note that the deformation models developed here will also be of direct use to the physics based earthquake simulators, so we are coordinating accordingly. Development of the final deformation models is the topic of Task D1, which is described below and in the associated appendix entry.

Logic-Tree Branches

We want some number of alternative deformation models that span the range of viability and hazard implications. Due to the computational demands of our anticipated earthquake rate model (outlined below), we want the smallest justifiable number of branches. Within each of these alternative models it will be assumed that the slip-rate uncertainties are uncorrelated (because the alternative models are constructed explicitly to remove such correlations). Specification of these alternatives is part of the ongoing Task D1 described below and in the associated appendix entry. In short, the UCERF3/GPS sub group lead by Wayne Thatcher and Kaj Johnson will identify no more than four alternative deformation models from a suite of candidate deformation models derived from seven different kinematic models. These kinematic models use a variety of forward model assumptions to invert GPS-derived velocities and geologic slip rate data for fault slip rates and off-fault deformation rates.

Current Status

Through frequent emails, teleconferences, and meetings (the most recent being at the Sept. 2011 SCEC Annual in Palm Springs), the deformation model leaders have been making steady progress towards completing their contribution to UCERF3 (see Task D1 appendix entry for details). In short, this group plans to provide not only kinematically consistent slip rates on faults, but also gridded values of off-fault deformation (at 0.1-degree spacing). Specifically, they will provide four such models:

1. Based on Neokinema (Bird, 2009)
2. Based on Yuehua Zeng's model (USGS, golden)
3. A uniform block strain rate model, averaged from five different kinematic models
4. A spatially variable block strain rate model, averaged from five different kinematic models

The two block models will represent a kinematically-consisent average found from a number of different block-modeling approaches (found from a "Great Inversion", as described in the appendix entry for Task D1). The task leaders will suggest logic-tree branch weights for each model.

A number of preliminary models have been constructed over the course of 2011 with subsequent improvements on the geometry and use of data constraints. Each modeler will submit a final model by October 14, 2011. The four candidate deformation models will be ready for carefully study by Thatcher, Johnson, Weldon, and Dawson by October 21, 2011, and the final alternative deformation models will be decided on by early November, 2011. A preliminary average model derived from five kinematic models using a block geometry is shown in Figure 2b. For comparison, the mean geologic slip rate used to constrain the model slip rates are shown at the same color scale. Preliminary models suggest that most model slip rate estimates are in good agreement with geologic rates. The group has identified a plan for placing slip rate estimates on UCERF3 faults that are not in the block model. Because of the inherent difficulty in estimating

slip rates on closely space, low rate faults with the kinematic models, the GPS sub group will be working closely with Ray Weldon and Tim Dawson to assure that estimates are consistent with the expected range of slip rates and style inferred from geology.

De-scoping Options

Our ultimate fallback option here is to revert to the UCERF2 deformation models.

Ongoing Tasks

Table 2. Task List for the Development of Deformation Models.
Notes for Table 1 apply here.

	Task	Description	Leader & Participants
Deformation Model(s)	**D1) Develop New Deformation Models**	Develop a new set of deformation models based on the more sophisticated modeling approaches that have recently emerged (e.g., NeoKinema, Harvard-MIT block model, Shen/Zeng model, Parsons' 3D FE model). A range of models will be needed to represent epistemic uncertainties. The following are some of the other questions that this task will try to address: 1) What is the bulk fraction of "off-fault" deformation? Can each model be very specific about what amount of deformation contributes to slip rates inferred from paleoseismic studies (on main fault surface) versus what amount is occurring in the surrounding zone (& manifested as nearby off-fault earthquakes)? 2) Can we get a more refined spatial distribution of off-fault deformation (e.g., values on a 0.1 by 0.1 degree grid rather than large polygons)? This could be used to constrain the rate and/or maximum magnitude of "off fault" seismicity in our Earthquake Rate Model (as an alternative to the traditional use of smoothed seismicity). 3) Can we constrain slip rates on those faults that have no geologic info (e.g., the gray ones in figure 2a)? 4) Can we differentiate slip rates on closely spaced faults? 5) Can we constrain slip-rate variations along strike at say 5 to 10 km intervals (since how slip tapers at the ends of faults will be very important in terms of multi-fault rupture likelihoods)? If not, will the constraints be at points or on larger fault sections? 6) Can we use GPS to help constrain the distribution of aseismicity and seismogenic depths (the latter presumably being related to locking depths)? 7) What are the long-term after effects of previous large earthquakes (like those in 1857, 1872, and 1906)? 8) How can slip rates determined for simplified block models be mapped back onto the full inventory of faults in our fault models.	**Thatcher**, Zeng, Hearn, Johnson, and Sandwell.

		See appendix for further discussion of this task.	
	D2) Geologic Slip Rate Compilation	Compile geologic slip-rate constraints for points on the faults in our *Fault Section Database*. The goal here is to revert to exactly what is know geologically regarding slip rates, as opposed to using the expert opinion and section-average slip rates applied in UCERF2. These new, relatively pure geologic slip rates will be used to constraint the deformation models being developed in Task D1. *See appendix for further discussion of this task.*	**Dawson,** Weldon
	D3) Line integral tools	f) Implement tools for line-integral testing (Parsons has started this); strain tensor analysis tools for polygons applied to deformation or earthquake-rate models. Vertical components could be an important additional constraint. *See appendix for further discussion of this task.*	**Parsons**, Milner, Powers, Weldon
	D4) Aseismicity Data Compilation	Update the UCERF2 creep data, accompanying maps, and aseismicity parameters applied in the UCERF3 deformation models. There are a number of new studies of creep on California faults, particularly using InSAR. These influence thoughts on how deep creep extends to depth and whether it is releasing potentially seismogenic strain (esp on key portions of the San Andreas system).	**Weldon**, Wisely, and Dawson

Earthquake Rate Models

Definition

The goal of the earthquake-rate model component of UCERF3 is to define the long-term rate of all possible earthquake ruptures, above some magnitude threshold and at some discretization level that is sufficient to capture hazard. These rate models will include aftershocks (which can be removed, if necessary, for the NSHMP maps or in constructing the Earthquake Probability Models). In UCERF3, we aim to relax segmentation and include multi-fault ruptures. We note that relaxing segmentation does not necessarily mean removing it, but rather sampling whatever range of models are consistent with the data, which may or may not exhibit segmentation.

Each earthquake-rate model will be composed of two types of sources: 1) those occurring on explicitly modeled faults; and 2) those modeled as "background" seismicity (the latter being represented with a magnitude-frequency distribution of nucleation rates at each point on a 0.1 by 0.1 degree grid). Rather than building the models for each fault separately and then adding background seismicity later, as was done in UCERF2, we plan to solve for the rates of all events simultaneously. For this purpose, we will apply the inverse approach outlined by Field and Page (2010), which builds on the work of Andrews and Schwerer (2000). We start with a statement of how we intend to solve this problem, and then follow with a discussion of several important issues. Example implementations for northern California are given in the appendix (Task R5).

Methodology

We first consider only those ruptures that occur on the faults defined in our Fault and Deformation Models (addition of background/off-fault seismicity is discussed below). Here we will only model events that have a rupture length greater than or equal to the seismogenic thickness (relegating smaller events to the off-fault seismicity). To relax segmentation, we subdivide each fault section into S equal-length subsections with lengths of about half the seismogenic thickness (e.g., S=1598 for the example shown at the bottom of Figure 3).

Next, we define all "viable" ruptures in the fault system as every set of two or more contiguous fault subsections, where contiguous is defined as those subsections within ~5 km of each other (this distance is adjustable). The minimum number of two subsections ensures that the minimum rupture lengths are approximately equal to the seismogenic thickness, since subsection lengths are about half that. Note in Figure 3 that nearly all the subsections in California are connected to all others without jumping more than 5 km, illustrating high connectivity. We further filter the set of ruptures by requiring that:

1) Strikes cannot vary by more than some amount between neighboring subsections (e.g., 45°)
2) Strikes cannot change by more than some total amount along the rupture length (e.g., 90°)
3) Rakes cannot change by more than some amount between subsections (e.g., 90°)
4) Ruptures cannot include a given subsection more than once (e.g., preventing ruptures from looping back on themselves)

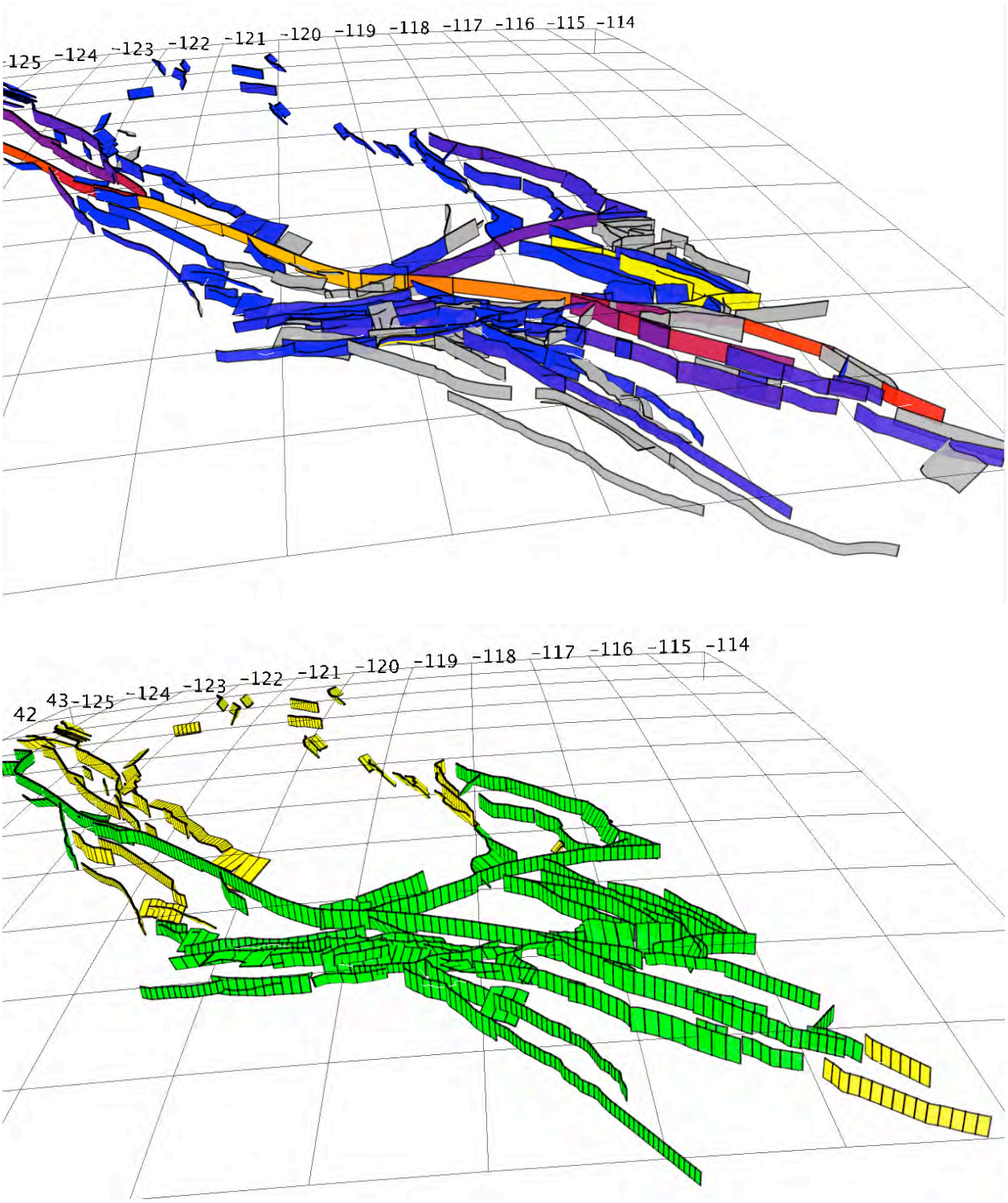


Figure 3. Top: UCERF2 fault sections for Deformation Model 2.1. Bottom: results of dividing sections into an integer number of equal-length subsections (lengths equal to, or just less than, half the section's seismogenic thickness). All subsections shown in green are connected to all others in green without jumping more than 5 km between faults.

By "viable" we mean within the realm of possibility; this does not imply that all ruptures are equally likely, which will be accounted for in the inversion below. We now have the total set of R viable fault ruptures, and we seek to solve for the long-term rate or frequency (f_r) of each r^{th} rupture, which we do so by solving the system of equations described in Box 1.

Box 1. Grand Inversion Equations

A system of equations to solve for the long-term rate or frequency (f_r) of each r^{th} rupture. The constraints below can be applied with varying weights in order to balance the influence of each.

Equation Set		Description
$\sum_{r=1}^{R} D_{sr} f_r = v_s$	(1)	***Slip-Rate Balancing***: v_s is the subsection slip rate (from a deformation model) and D_{sr} is the average slip on the s^{th} subsection in the r^{th} event (by average we mean over multiple occurrences of the rupture, and as measured at mid-seismogenic depth).
$\sum_{r=1}^{R} G_{sr} P_r^{paleo} f_r = f_s^{paleo}$	(2)	***Paleoseismic Event-Rate Matching***: f_s^{paleo} is a paleoseismically inferred event-rate estimate (where known) and P_r^{paleo} is the probability that the r^{th} rupture would be seen in a paleoseismic trench.
$f_r = f_r^{a-priori}$	(3)	***A Priori Constraint***: Constrain the rates of ruptures to any desired values. This can be used on individual ruptures (e.g., make Parkfield occur every ~25 years) or to a complete rupture set in order to obtain a unique solution of interest (e.g., keep the final rates as close as possible to those in UCERF2, but while satisfying other data).
$\lambda_r f_r = 0$	(4)	***Improbability Constraint***: This allows us to force relatively improbable events to have a lower rate (e.g., based on multi-fault rupture likelihoods). A higher value of λ_r adds more misfit for a given rupture rate, forcing the inversion to minimize that rupture rate further.
$\sum_{r=1}^{R} M_{mr}^{g} f_r \le GR_m^g$	(5)	***Regional Gutenberg-Richter (GR) Constraint***: This forces geographic sub-regions (g) to have a magnitude-frequency distribution that is less than or equal to a Gutenberg-Richter rate (to prevent over-prediction "bulges"). GR_m^g represents the GR rate of the m^{th} magnitude bin in the g^{th} sub-region, and matrix M_{mr}^{g} contains the product of whether the r^{th} rupture falls in the m^{th} magnitude bin (either 0 or 1) multiplied by the fraction of that rupture that nucleates within the g^{th} sub-region.

Other equations can be added as deemed useful (see main text for examples).

Conceptually, the approach outlined here is much simpler, more objective, more reproducible, and more unified than that adopted in UCERF2. In particular, we have abandoned the largely artificial distinction between Type-A and Type-B faults, and Type-C zones have been removed (now included with off-fault seismicity). However, it's important to note that we have not removed all expert opinion or subjectivity. Indeed, the weights applied to the various inversion constraints will reflect expert opinion, as will the weights applied to the various final logic-tree branches. Instead, this Grand Inversion should be considered a more general framework for both incorporating expert opinion and more thoroughly exploring the solution space.

Implementation Considerations

The equations in the inversion can be weighted based on uncertainties in the data or the degree to which we wish to impose a constraint. This section elaborates on some of the issues that need to be considered in applying the inversion equations.

Slip-Rate Balancing (Equation Set 1):

This constraint requires knowing the average slip on the s^{th} subsection in the r^{th} rupture (D_{sr}), where "average" means over multiple occurrences, and the slip is interpreted as that at mid-seismogenic depths (because slip may vary with depth). In other words, we are now slip-rate balancing rather than moment balancing in order to avoid depth-of-rupture ambiguities. We first derive the average slip for a given rupture (D_r), and then partition this among the subsections to get D_{sr}.

As in UCER2, we will always compute the magnitude of each rupture from an assumed magnitude-area relationship ($M(A_r)$). However, we will pursue two different ways of getting D_r for each rupture. One way, and again as in UCERF2, will involve converting magnitude to moment and then dividing by the rupture area (A_r) and shear rigidity (μ) to get D_r:

$$D_r = \frac{Mo_r}{\mu A_r} = \frac{10^{1.5*M(A_r)+9.05}}{\mu A_r}$$

where Mo_r is the moment of the r^{th} rupture. In UCERF2 the Hanks and Bakun (2008) and Ellsworth B (WGCEP, 2002) relationships were used with equal weights. For UCERF3 we will also include the Shaw (2009) relationship, as justified in the appendix for Task R2 (listed in Table 3 below). These relationships are hereafter referred to as HB08, EllB, and Shaw09, respectively.

Because A_r is the above equation is based on the depth of microseismicity, D_r might be an over estimate to the extent larger ruptures actually penetrate deeper (a current epistemic uncertainty). To account for this possibility, the second way we will obtain D_r is using one of two viable slip-length scaling relationships derived from surface-slip observations (also described in the appendix for Task R2). The first assumes slip scales as the square root of area, and the second assumes constant stress drop. Examples obtained using these two slip-length models, as well as using the three magnitude-area relationships above, are given in Figure 4. The slip-length models generally give a smaller D_r for a given rupture, which either represents reality (due to slip penetrating below the depth of microseismicity), or a bias from the fact that surface slip measurements are lower than those at mid-seismogenic depths. Including these slip-length

relationships allows us to account for this additional epistemic uncertainty, which was not included in UCERF2. Branch weights will be determined after we have had a chance to explore implications.

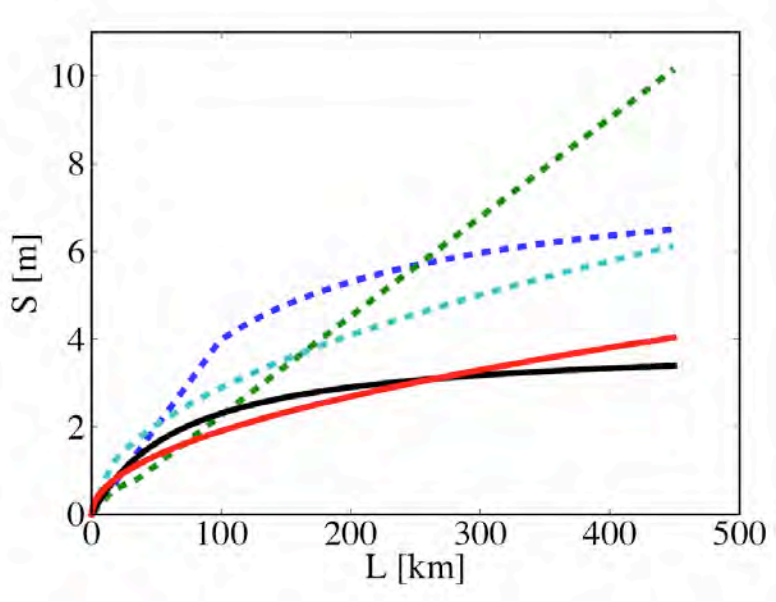


Figure 4. Slip-length scaling relations for large strike-slip events on fault with a seismogenic thickness of W=15 km (from the appendix entry for Task R2, where "S" in this plot is equivalent to D_r). Dashed lines are implied slip-length scaling derived from magnitude-area relations (green for HB08, cyan for EllB, and blue for Shaw09). Solid lines are slip-length scaling relations derived from surface slip observations, where the solid red line represents an $L^{1/2}$ model (S=0.0000491$(LW)^{1/2}$), and the solid black line represents a constant stress-drop model (see Equation (5) of the appendix entry for Task R2).

Once we have D_r, the next step is to partition this among the subsection to get D_{sr}. As in UCERF2, we will apply the "tapered slip" (square-root of sign) model of Weldon et al. (2007), with the further assumption that this is applicable to multi-fault ruptures as well (justified in the appendix entry for Task R1). One could argue that slip might pinch out to some extent at multi-fault-rupture connections (producing multi-rainbow slip distributions). However, applying such a model would also require careful consideration of whether slip rates also pinch out at fault connections, and unfortunately this is beyond the resolving power of our deformation models (which assume no reduction in slip rates at fault connections). This unaccounted for source of epistemic uncertainty may be an important topic to explore in the future, but it is presently beyond the scope of UCERF3.

In addition to the tapered slip model, we also plan to explore the uniform slip model (where D_{sr}=D_r along the entire length) and the WGCEP (2002) model where slip is proportional to the subsection slip rate (both of these alternatives were available in UCERF2, but given zero weight in the final model). While there also appears to be some evidence to also support a characteristic slip model, where the amount of slip on a subsection is the same for all ruptures, applying this would be difficult due to very limited observational constraints (requiring the propagation of large epistemic uncertainties with unknown correlation structure). Our present plan is to check

for consistency between results obtained as outlined above, and any observational constraints for characteristic slip (including slip coefficients of variation as given by Hecker et al. (paper in review, and not available for distribution)).

Because the inversion is solving for the rate of seismogenic-thickness and larger ruptures, the slip rates in Equation Set (1) should really be reduced to account for the moment released in smaller events. This correction may be negligible given overall uncertainties, but we will nonetheless explore the issue.

Paleoseismic Event-Rate Matching (Equation Set 2):

The compilation of paleoseismic event-rate data (f_s^{paleo}) is the topic of Task R3 listed in Table 3 below (with details given in the associated appendix entry). The probability of seeing a given rupture in a trench (P_r^{paleo}) is the topic of Appendix R4 (also listed in Table 3 and described more in the appendix). Ideally the latter would be done on a trench-by-trench basis given the unique depositional environment of each site. However, this is be beyond the scope of UCERF3, so a generic probability model will be developed and applied.

A Priori Constraint (Equation Set 3):

This constraint allows us to obtain unique solutions that are as close as possible to some model of interest. For example, one of the main logic-tree branches discussed more below will represent a solution that is as close as possible to UCERF2 rates, while making whatever minimal adjustments are needed to match the other data constraints better. Another *a priori* model is to keep the solution as close as possible to a smooth Gutenberg-Richter solution (also described more below).

Improbability Constraint (Equation Set 4):

This constraint represents the means by which we can penalize any relatively low probability events (e.g., multi-fault ruptures that involve large jumps). More specifically, Task R6 listed in Table 3 below (and discussed in the associated appendix entry) will provide a numerical algorithm for defining the relative likelihood of every rupture in our model. This will be based on observational fault-jumping statistics (e.g., Wesnousky, 2008), static and dynamic stress modeling, more detailed geologic maps, and relocated seismicity where available (the latter two addressing the actual proximity of our fault-section endpoints). If we write these relative probabilities as $P_r^{multi-fault}$ (ranging from 1.0 for ruptures that have no obvious impediments to 0.0 for impossible ruptures), then the associated equation for each rupture in the inversion will be given a weight of $1/P_r^{multi-fault}$.

Note that the rate of multi-fault ruptures is already limited in the inversion by both slip-rate balancing (larger ruptures consume more slip) and the GR constraint (large events have limited collective rates). Thus, it remains to be seen how much this constraint is actually needed given how common multi-fault ruptures are in nature. Nevertheless, this represents a viable mechanism for controlling the rates of multi-fault ruptures if needed.

Regional Gutenberg-Richter (GR) Constraint (Equation Set 5):

The two main goals of this constraint are: 1) to prevent an over-prediction of M 6.5 to 7.0 events as seen in some regions for UCERF2 (the bulge problem); and 2) to integrate the fault-based rates with off-fault seismicity in a more consistent way than was done in UCERF2. The basic idea is to force the combined models to match an appropriate Gutenberg-Richter (GR) distribution in different sub-regions of the state.

For a given sub-region, defined by a polygon, we can obtain the a-value for the GR distribution from historical and instrumental seismicity (from Tasks R7 and R8 below), or perhaps from the deformation models with some assumptions. Assuming a b-value, we can then define the sub-region's GR distribution out to arbitrarily large magnitude (giving GR_m^g in Equation Set (5)). The constraint in the inversion is simply that the total magnitude frequency distribution (MFD) for all the fault-based ruptures that nucleate in this region must be less than or equal to this GR distribution.

Once the inverse problem has been solved, the MFD for off-fault seismicity is set equal to the target GR distribution minus the fault-based MFD for the sub-region, which ensures that the total MFD for the sub-region is a perfect GR (and lacking a bulge, as illustrated in Figure 5a). However, this algorithm can produce negative rates of off-fault seismicity; e.g., where the fault-based MFD is unable to satisfy Equation Set (5). In this case, we could set the negative rates in the off-fault MFD to zero, and let the consequent bulge remain. However, such a discrepancy would probably motivate us to examine the fault inventory and slip rates for errors, or to ask whether our estimate of the sub-region a-value is biased, perhaps owing to temporal variability.

If the resultant maximum magnitude of the background seismicity is improbably low, as illustrated in Figure 5b, then we can repeat the inversion by imposing minimum rates for the background (by forcing the fault-based MFD to be less than or equal to the regional GR minus these minimum background rates). This can be achieved by replacing GR_m^g in Equation Set (5) with $(GR_m^g - B_m^g)$, where B_m^g is the minimum background-seismicity rate specified for the m^{th} magnitude bin (in Figure 5b this is simply the target rate divided by 100, but something more sophisticated could also be applied). Alternatively, if the resultant maximum magnitude of the background-seismicity is too high (e.g., because there are too few fault-based ruptures in the sub-region), then we simply truncate that distribution at some pre-defined maximum value for that region (illustrated in Figure 5c). This would be consistent with UCERF2 generally applying an upper-magnitude cutoff (M_{max} = 7.0) to the background. Alternatively, we could tune this value to match any off-fault deformation rates provided by Deformation Models 3.x.

The final off-fault MFD for the sub-region will then be partitioned among the 0.1 by 0.1 degree background-seismicity grid cells according to the distribution of a-values determined from spatially smoothed seismicity (the topic of Task R9 in Table 3 below).

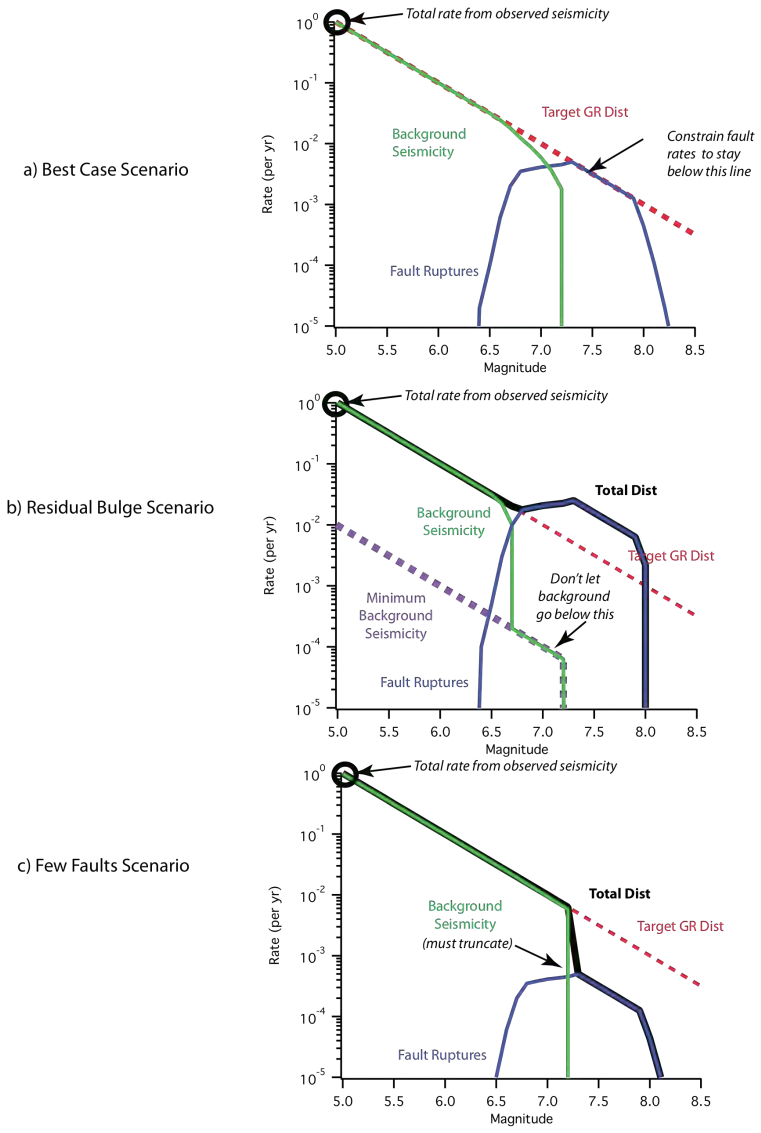


Figure 5. Illustration of the sub-region Gutenberg-Richter (GR) constraint, where the dashed red line is the target GR distribution and the blue, green, and black lines are final results for the faults, background, and total (combined), respectively. a) The case where the target GR can be matched perfectly (total is equal to red dashed line). b) The case where the fault rates cannot be constrained below the target, so a minimum must be applied to the background (purple) to ensure positive rates, and the total distribution has a bulge; c) The case where faults can't match the target so the background is truncated at a chosen M_{max} (arbitrarily set at M7.2 here).

One issue is the choice of spatial scales over which to apply the GR constraint of Equation Set (5). A statewide scale for this constraint is probably too coarse; for example, northern California could then have a strong bulge as long as southern California had a commensurate anti-bulge. Applying the GR constraint at the smallest scale resolved by the model (e.g., on a $0.1^o \times 0.1^o$ grid) would effectively enforce GR behavior everywhere, but this would preclude characteristic-earthquake MFDs on faults, which some experts favor. We presently plan to apply this constraint to $1.0^o \times 1.0^o$ sub-regions (e.g., the grids lines in Figure 3), although other choices will certainly be explored (including zones based on geomorphology).

Other Possible Constraints:

Several other types of constraints can be added to the inversion. For example, the rates of events on specific faults can be constrained by the historically observed rates. We can also impose smoothness constraints on the MFDs for faults or sub-regions to avoid irregular distributions. Likewise, we could constrain the rate of given sized events to be constant along uniform sections of faults (applied with a relatively weak constraint to allow other data to make it variable if needed). Adding such as constraints will increase computational demands, so we will want to explore whether there are alternative approaches for achieving the same thing (e.g., averaging over multiple simulated annealing solutions described below).

Solving the Inverse Problem:

In their exploratory study, Field and Page (2010) computed the non-negative least squares (NNLS) solution to the inverse problem (Lawson and Hanson, 1974). As the NNLS algorithm is not computationally feasible for an inversion using the statewide system of faults, we intend to solve the full problem using a parallelized simulated annealing approach that is being developed by Morgan Page and Kevin Milner (Task R5 in Table 3). Preliminary tests show that this algorithm allows us to obtain significant speedups using a cluster of processors. Furthermore, rather than giving just a particular solution, simulated annealing can provide a range of models that sample the solution space of the under-determined inverse problem. Such sampling may provide a rational basis for constructing logic-tree branches to represent this type of epistemic uncertainty. We think this will allow us to represent a broader range of models than has been included in the past.

Logic-Tree Branches

We envision the following main logic-tree branches for the UCERF3 Earthquake Rate Model:

Branch 1 – As Close As Possible To UCERF2:

For this model we will set the *a priori* rates in Equation Set (3) to be the equivalent UCERF2 rates (most of which will have zero values). By giving this equation set a relatively weak weight, we will obtain a unique solution that is as close as possible to the UCERF2 rates, but where rupture rates are modified by whatever minimal amount is needed to satisfy the other data. Here, we will not utilize the improbability constraint of Equation Set (4), and the regional GR constraint (Equation Set (5)) will either be applied to the entire region, or to northern and southern California regions only.

Furthermore, we will first run this case using the old UCERF2 data constraints in order to determine which changes result from the new inversion methodology. Having answered this, we will then move on to utilizing the new data constraints as well, which will presumably constitute the final model for this logic-tree branch.

This model will represent a minimal, incremental change with respect to UCERF2, which some may find appealing from a policy perspective. It will also reveal which changes result from the new methodology versus which result from new data constraints. A disadvantage of this model is that many of the rupture rates will remain zero, even for ruptures we believe to be quite plausible, simply because non-zero rates are not required to satisfy the data. A possible remedy to this would be to apply a minimal "water level" rate to all ruptures (small enough that it does not effect the fit to data).

Branch 2 – A Smooth GR Model:

For this branch we will set the *a priori* rates in Equation Set (3) according to a regional GR starting model. From the total population of ruptures in the inversion, we can count the number of ruptures that fall into each discrete magnitude-frequency bin. We can then partition the known total rate for each magnitude bin among the associated ruptures, weighted by fault slip rates and penalized by the number of spatially overlapping ruptures (to avoid double counting the latter). This will presumably give a reasonable GR starting model, but it won't exactly match the other data constraints. By applying this a priori constraint with a relatively weak weight, we will obtain a final unique solution that is as close as possible to this starting model, but which is modified in order to satisfy the other data constraints. At first we will exclude the improbability constraint (Equation Set (4)) in order to see if results look reasonable without it. With respect to the regional GR constraint, we will start with 1-by-1 degree regions and adjust as warranted (and perhaps this constraint will not be needed since the *a priori* constraint is GR).

A good thing about this model is that it imposes more GR behavior throughout the region, and thereby constitutes a logic-tree branch that some members of our community have felt has been lacking. This solution may also be equivalent to that obtained using the Monte Carlo approach of Parsons and Geist (2009).

One challenge here is the target magnitude bin rate for the starting model should have off-fault seismicity rates subtracted out, so assumptions will be needed with respect to the contribution of off-fault events.

Branch 3 – Average From Many Simulated Annealing Runs:

Here, no complete set of *a priori* constraints would be applied, leading to an underdetermined or mixed-determined problem, and therefore no unique solution. We will try to use simulated annealing to sample a large number of models from this solution space, each of which could be kept as a viable branch, or we could average all the results to get some mean branch. A serious limitation here will be computation time, so it remains to be seen whether this is a viable approach.

Possible logic-tree sub-branches:

The following are possible logic-tree sub-branches (added to the main branches outlined above):

- Alternative deformation models (influencing slip rates and possibly total off-fault seismicity rates)
- Different magnitude-area or slip-length scaling relationships (influencing D_r)
- Models of along-strike distribution of average slip (influencing D_{sr})
- Different constraints on fault-to-fault rupture probabilities (including uncertainties in distance between faults)
- Uncertainties in paleoseismically determined mean recurrence intervals
- Alternative models for the likelihood of seeing events in a paleoseismic trench
- Uncertainties in observed regional magnitude-frequency distributions, and the spatial distribution of both a-values (Task R9), M_{max} (Task R10), and focal mechanisms (Task R11) for off-fault seismicity.

Due to the large computational demands of the inversion, coupled with the fact that few users actually want more than a mean model, logic-tree branches will have to be chosen with care. Another question is which logic-tree branches would require Monte Carlo sampling, and to what degree this will be feasible. Further specification of branches and their weights therefore awaits sensitivity studies.

Model Evaluation

We plan to look at the following features of each solution (including UCERF2) in order to evaluate the reasonableness of each model:

- Participation rate of each fault section (or the ratio of these between any two solutions)
- Paleoseismically observable rate of each fault section (by applying P_r^{paleo}), plus comparison with observations (f_s^{paleo}) where available)
- Slip rate of each fault section (either the original data constraint, the final implied model value, the ratio of these two, or the ratio of implied slip rates between different model solutions)
- Magnitude-frequency distribution for a region or specified fault section (nucleation or participation rates), or the ratio of these between solutions.
- Implied probability distribution of observed slip for a fault section, or derivative values such a mean and coefficient of variation.
- The rate at which neighboring sub-sections rupture together, where low values would constitute rupture barriers, or segment boundaries, which we would expect to be present a fault-jump locations. Also, the ratio of these between models (e.g., by comparing with UCERF2, we will see what actual fault jumps are being introduced o satisfy the data).

- We can also visualize individual ruptures, where color represents their rate, magnitude, or average slip (D_r), or we can visualize all the ruptures that utilize a chosen subsection.

Most of the above visualization features are already available in the SCEC-VDO 3D visualization platform (http://www.wgcep.org/tools-scec_vdo), and the others can be added in short order. Furthermore, we've implemented these capabilities such that physics-based simulator results can be evaluated as well (when available).

Current Status

The current status is reflected in the appendix entries for the various tasks listed below. Of particular note is Task R5, for which several initial inversion models are presented to demonstrate feasibility. Specifically, the inversion has been applied to northern California (using UCERF2 ingredients), and matches the data better than UCERF2. This took several hours of annealing time on a single processor. Also, a minimally constrained statewide inversion model has been implemented using a parallelized simulated annealing algorithm. This inversion achieves good convergence in approximately 8 hours (real time) on 20 nodes.

De-Scoping Options

The biggest risk here is that the inversion proves to be too computationally demanding, the solution space is too vast and difficult to explore, or the results are not found to be acceptable by the broader community. One option would be to apply the inversion to smaller subset regions (as in Field and Page, 2011), but this would limit the multi-fault ruptures included in the model. In the worst-case scenario, we can fall back onto elements of the UCERF2 earthquake rate model.

Ongoing Tasks

Table 3. Task List for the Development of Earthquake Rate Models.
Notes for Table 1 apply here.

	Task	Description	Leader & Participants
Earthquake Rate Model(s)	**R1) Evaluate Along-Strike Distribution of Average Slip**	What is the average along-strike distribution of slip in an earthquake, especially when multiple faults are involved (e.g., reduce slip at the connections)? What if the style of faulting changes between faults? The big issue here is whether multi-fault ruptures exhibit multi-rainbows (as opposed to one large rainbow), and if so, whether these can by reliably predicted. Currently it appears that we can model multi-fault ruptures as single rainbows. *See appendix for further discussion of this task.*	**Biasi**, Weldon, Dawson, & Wesnousky?.
	R2) Evaluate Magnitude-Scaling Relationships and Depth of Rupture	Resolve discrepancies in existing magnitude scaling relationships. The main question is whether rupture depth for larger events continues to increase with magnitude (or whether it's limited by the depth of microseismicity). We need a range of viable models to represent existing uncertainties. The conclusion here is that in addition to using magnitude area (including a new relationship), we should also include slip-length scaling models as alternative logic-tree branches. *See appendix for further discussion of this task.*	**B. Shaw**
	R3) Paleoseismic Recurrence	Update and/or add to Tom Parsons's compilation of mean recurrence interval (MRI) estimates for paleoseismic sites. We	**Parsons**, Biasi, Weldon, Dawson , & Grant-

	Interval Estimates	also need to consider independence of these from slip rate estimates at the same locations. What level of publication qualifies for this compilation? *See appendix for further discussion of this*.	Ludwig
	R4) Probability of Seeing Events in a Paleo Trench	We need a model giving the probability that a given magnitude event below a site would be seen in a paleoseismic trench. The previously used model of Youngs et al. (2003) only gives the probability that rupture will daylight somewhere along the surface, which is different than the likelihood that it will be seen at a given point along the fault. Tthese models should probably be defined on trench-by-trench basis (to account for the unique attributes of each paleo site), but this appears beyond scope at this time? *See appendix for further discussion of this*.	**Weldon**, Petersen, Biasi, Dawson.
	R5) Solve the "Grand Inversion"	*See main text for general discussion and the appendix entry for the application to N. California*.	**Page**, Field, Milner, Parsons
	R6) Fault-to-Fault Jumping Probabilities	Define fault-to-fault jumping probabilities to use as constraints in the inversion, where the probability is defined relative to the case were there is no fault separation and no change in strike, dip, or rake. These will be based on an expert-opinion analysis of the following: 1) higher resolution fault maps (Dawson and Weldon?) 2) distribution of microseismicity (Peter Powers) 3) statistics from global empirical data (Biasi, Wesnousky?) 4) static stress change analysis (Parsons) 5) dynamic rupture modeling (Oglesby, Harris, Shaw?) One issue is the fact that empirical studies were conducted after given earthquakes, as opposed to being based on the more uncertain information we have before events. Another question is whether we develop generic "tinker toy" rules or consider each case in our fault model individually. *See appendix for further discussion of this*.	**Biasi**, Parsons, Dawson, Weldon, Wesnousky, Oglesby, Harris, Powers, Shaw, Page, Field, & Jackson.
	R7) Reassess Historical Earthquake Catalog	Evaluate whether there may be biased estimates of magnitude and locations from felt reports. For example, treat larger events as lines rather than points. *See appendix for further discussion of this task*.	**Parsons,** Bakun?
	R8) Reevaluate Earthquake Catalog	The primary goal of this task is to provide: 1) an estimate of the total observed magnitude frequency distribution for the entire region (along with uncertainties); 2) a-value estimates for arbitrarily defined regions/polygons (to be used as a constraint in the inversion; and 3) the magnitude-frequency distribution of off-fault seismicity (all events that are not part of our fault-based model, including sub-seismogenic ruptures on those faults). These estimates should include aftershocks, although estimates where aftershocks are removed using the Gardner Knoppoff (1974) methodology may also be needed (since this is what the NSHMP has used in the past). The approach here is to reevaluate the association of events with different faults, and use both historical and instrumental catalogs to determine rates, including the total magnitude-frequency-distribution, using whatever approaches are appropriate (e.g. a range of declustering models, various methods of dealing with parameter tradeoffs in rate determination). One question is how much of an inferred magnitude bias would be needed to remove the UCERF2 discrepancy between predicted and observed rates. Can we pinpoint exactly what data or model components are influencing the discrepancies? *See appendix for further discussion of this task*.	**Michael**, Felzer, Bakun, Parsons, Schorlemmer, Hardebeck
	R9) Smoothed	Reevaluate procedures for smoothing instrumental seismicity in	**Felzer**, Mueller, Biasi,

	Seismicity Model	light of both RELM test results and precarious rocks implying that spatial smoothing should be tighter. *See appendix for further discussion of this task.*	and Grant-Ludwig
	R10) Mmax for off-fault seismicity	Develop more quantitative estimates of maximum magnitude for off-fault seismicity, either by considering the size of those faults left out of the deformation model (for lack of a slip rate) or by what is needed to satisfy the extra deformation in our previously defined Type-C zones (or any new zones defined by the deformation models discussed above). This will allow us to merge C zone sources into the background, which would be good in terms of removing an existing artificial distinction.	**Jackson,** Field, Michael
	R11) Focal mechanisms of off-fault seismicity	Define the probability for different focal mechanisms as a function of space throughout California (for events not on modeled faults). *See appendix for further discussion of this task.*	**Jackson**, Hauksson
	R12) Distribution of Slips in Paleo Trench	Are they more consistent with Characteristic or GR models? Get an update on the Hecker/Abrahamson contention that trenches reveal characteristic slip (slips seem to be the same over multiple events). *See appendix for further discussion of this task.*	**Weldon**, Hecker, Dawson, & Biasi
	R13) Evaluate Physics Based Earthquake Simulators (for rate estimates)	Investigate implications and applicability of physics based simulators for inferring the long-term rate of all possible ruptures (as well as other things). Do this in conjunction with the ongoing SCEC simulator working group being led by Terry Tullis. *See appendix for further discussion of this task.*	**Field**, Michael, Tullis, Dieterich, Richards-Dinger, Ward, Rundle, Pollitz, Beeler
	R14) Reconsider aleatory uncert. in Mag from given Area	In UCERF2 we gave a range of magnitudes for a given fault-rupture area, but this potentially gets double counted in hazard calculations because attenuation-relationship sigmas implicitly include a range of areas for a given magnitude. This also is a very important issue for SCEC's CyberShake project, and may contribute to the precarious-rock problem as well. A joint workshop with the NGA-W2 group to address this and other issues has been scheduled for Oct. 25, 2011.	**Field**, Campbell, Graves, others?
	R15 Cascadia subduction zone	Develop a complete, revised model for Cascadia. This component will be developed separately from the rest of UCERF3 because it is mostly outside California and has a different set of issues and data constraints. There will be significant overlap in participation, however, ensuring that model assumptions and methods are not contradictory. Art Frankel hosted a meeting on Cascadia in Corvallis, OR on Nov 18-19, 2010. *See appendix for further discussion of this task.*	**Frankel**, Weldon, Petersen

Earthquake Probability Models

The earthquake probability models specify the probability that each event in the long-term Earthquake Rate Model will occur over a specified time interval. The main goals for UCERF3 are the following: 1) address the interpretation of the UCERF2 "Empirical Model" (the apparent rate change in seismicity); 2) develop self-consistent elastic rebound models; and 3) apply spatiotemporal clustering models. We are aiming to build these features into a single forecasting model that is applicable across a wide range of time scales.

Resolve Interpretation of Empirical Model

The UCERF2 report demonstrated that the instrumental earthquake catalog for California (1932-present) has a lower average seismicity rate than the historical catalog (1850-1932). This rate decrease can be documented independently for the north coast, the San Francisco Bay Area, and for the central, and southern parts of the state. The rate decrease cannot be documented in the Mojave Desert because the historical data are lacking, but recent high seismicity rates in that region suggest an increase rather than a decrease.

The long-recognized decrease in seismicity rate in the San Francisco Bay Area has traditionally been attributed to the static stress shadow of the 1906 earthquake. The seismicity rates have not recovered in the way that static stress shadow modeling would predict, however, nor can a 1906 stress shadow explain rate decreases in other areas of the state. In the UCERF2 study, these conclusions motivated the weighting of an "Empirical Model", derived by applying the observed seismicity rate changes to the fault-based sources in the long-term Earthquake Rate Model.

Two main issues regarding the empirical model need to be addressed in UCERF3. The first is verifying the size of the seismicity rate change by re-evaluating the historical catalog. Since re-evaluation of all of the historical intensities would be a monumental task, the effort will focus on re-evaluating the errors associated with historical magnitudes and locations. Since earthquakes are more likely to be small than large, re-analysis is likely to lead to a smaller potential contrast between the historical and instrumental seismicity rates. It is very unlikely, however, that the rate contrast will disappear completely.

The second issue is the physical explanation. One hypothesis is that the rate change is due to earthquake clustering, which can be modeled with the stochastic ETAS model. Initial work indicates that ETAS can explain the rate change on a statewide basis. It has also been observed, however, that the rate decrease is focused on the San Andreas fault system; the contrast in seismic activity here between the historical and instrumental catalogs is striking (Felzer and Brodsky, 2005). The ETAS model as it is currently applied, with a constant background rate that comprises 40% of the total seismicity, cannot model the rate changes along some individual segments in this fault system. We will need to explore whether decreases or time-dependent changes in the background rate allow an ETAS-based explanation, and whether such changes are physically reasonable and applicable to modeling future behavior. These issues are explored further under Tasks P1 and P2 in Table 4 below, and in the associated appendix entry. We presently anticipate applying an empirical model as was done in UCERF2, but with more spatial resolution.

Develop Self-Consistent Elastic Rebound Models

Elastic-rebound motivated renewal models have been the foundation of the time-dependent probabilities in all previous WGCEPs. Computing conditional probabilities is simple when a fault is assumed to obey strict segmentation; i.e., where no multi-segment ruptures occur (e.g., WGCEP, 1988, 1990). However, the calculation is not straightforward when multi-segment ruptures are possible, in essence because we are attempting to apply a point-process model to what is clearly a spatially distributed process.

The methodology of WGCEP (2003) was applied by WGCEP (2008) in computing elastic-rebound probabilities for UCERF2. The WGCEP (2003) approach first computes the probability that each segment will rupture (from the long term-rate and date of last event, assuming a BPT distribution) and then partitions these probabilities among all ruptures that could be triggered by the segment. However, and as discussed in Appendix N of the UCERF2 report, this methodology is not self-consistent. One manifestation is that final segment probabilities, when aggregated over all ruptures, are not equal to the segment probabilities as originally computed. Another, revealed by Monte-Carlo simulations, is that the distribution of segment recurrence intervals implied by the model disagrees with the initial assumptions (Figure 6). For example, there is nothing that stops a segment from going by itself one day, and then being triggered by a neighboring segment the next, which leads to shorter than assumed recurrence intervals. The simulated rate of events is therefore biased high relative to the long-term rate (about 3% for the UCERF2 example in Figure 6a).

WGCEP (2008) applied the WGCEP (2003) methodology despite these shortcomings, because 1) an alternative was lacking; 2) the effects were minor since UCERF2 generally had only a few segments per fault; and 3) the methodology captured the overall intent of pushing probabilities in a direction consistent with elastic rebound, making the final values acceptable from the Bayesian perspective of probability as a statement of "the degree of belief that an event will occur" (D'Agostini, 2003)).

Unfortunately these problems worsen as the fault is divided into more and more segments, and especially if we relax segmentation altogether. Figure 6b shows the results for a simple "unsegmented" example, where the final distribution of recurrence intervals looks nothing like that assumed, and there is a non-negligible bias (~20%) in the total rate of events and overall moment rate. We therefore need an alternative approach for UCERF3.

In developing and evaluating other approaches, we shall make use of results from physics-based simulators (Ward, 2000; Rundle et el. 2006; and Dieterich and Richards-Dinger, 2010), which adhere to elastic rebound and make no assumptions regarding segmentation. Of course, no simulator correctly represents natural earthquake processes, so it will be important to evaluate any inferred statistical behavior for robustness against the range of simulator results, as well as against actual observations.

If a fault does not obey segmentation, then it is not possible for all points on that fault to honor a renewal-model distribution such as BPT or log-normal, especially where the tails of neighboring ruptures overlap. This is exemplified in Figure 7, which shows the distribution of recurrence

intervals at a point on the northern SAF from a simulation by Dieterich and Richards-Dinger (2010), which models the entire northern California fault system. This plot does not display any of the usual renewal-model distributions.

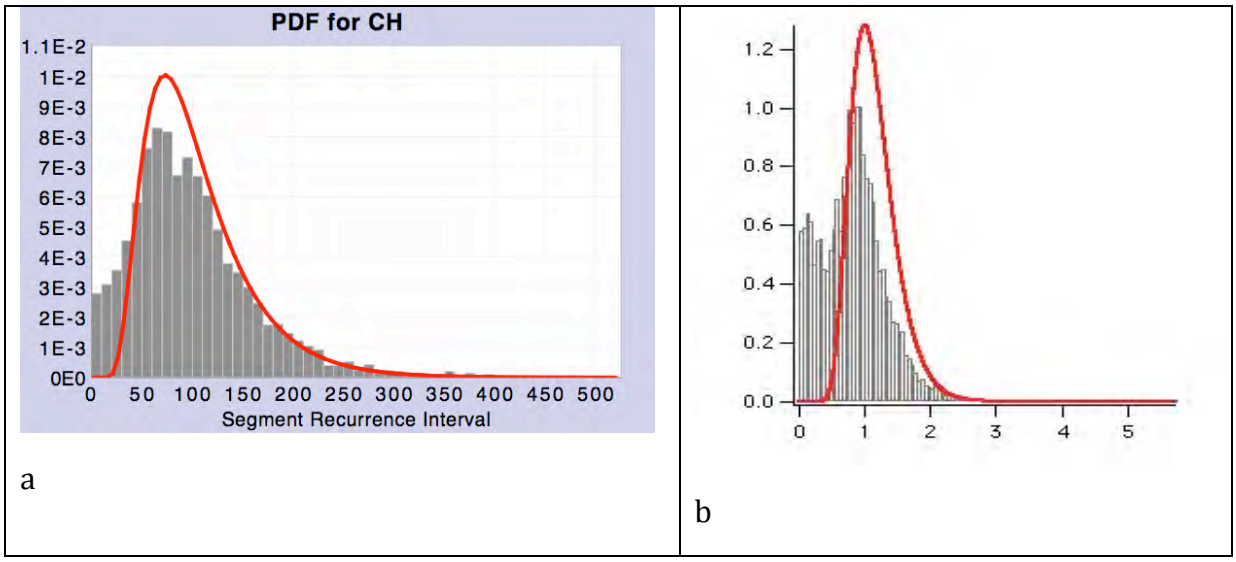


Figure 6. The distribution of recurrence intervals for the WGCEP (2003) methodology of computing time-dependent probabilities. Those assumed are shown in red (BPT with a COV of 0.5), and those implied by Monte Carlo simulations are shown as gray bins. a) An example for the Cholame segment of the southern SAF as modeled for UCERF2. b) an example for an 80-km fault with 5-km segments (essentially un-segmented) and a Gutenberg-Richter distribution of events. Both examples are taken from Appendix N of the UCERF2 report.

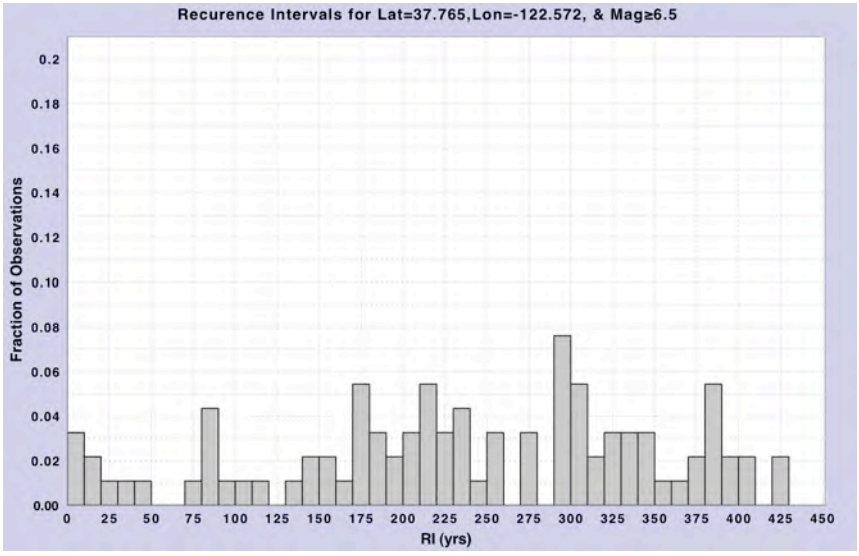


Figure 7. The distribution of M≥6.5 recurrence intervals at one location on the northern San Andreas Fault from the Dieterich and Richards-Dinger (2010) simulator. The model used in this simulation is the so-called "norcal1" fault system for northern California that has been implemented as part of the SCEC Simulators Working Group.

Stated another way, even if we had perfect knowledge of the recurrence-interval distribution at one or more points on a fault (as in Figure 7), it is not clear how to turn this information into rupture probabilities for an un-segmented model. Again, the problem seems to arise from the fact that we are attempting to apply a point-process model to a problem that, at least as heretofore posed, is not a point process.

We have formulated a promising alternative procedure. Consider the situation where we know exactly *where* the next big earthquake will occur, and are simply left with the task of predicting *when* it will occur. A sensible approach would be to apply an average time-predictable model:

$$T_r^{pred} = \frac{\sum_{s=1}^{S}\left(\frac{D_s^{last}}{v_s} + T_s^{last}\right)}{S} = \frac{\sum_{s=1}^{N}\frac{D_s^{last}}{v_s}}{S} + \frac{\sum_{s=1}^{N} T_s^{last}}{S} = \Delta T_r^{pred} + \overline{T}_r^{last}$$

This equation states that the predicted time of this r^{th} rupture (T_r^{pred}) is the average time at which the slip rate (v_s) on each subsection has recovered the amount of slip (D_s^{last}) that occurred in the last event at time T_s^{last} on each subsection. The average is taken over the total number of subsections (S) involved in the given event. The fact that T_s^{last} can vary along the rupture reflects the un-segmented nature of the model, and thus represents a straightforward generalization of the "time-predictable" model introduced by Bufe et al. (1977) and Shimazaki and Nakata (1980). The equation can be rewritten as

$$T_r^{pred} = \Delta T_r^{pred} + \overline{T}_r^{last}$$

where

$$\Delta T_r^{pred} = \frac{\sum_{s=1}^{N}\frac{D_s^{last}}{v_s}}{S} \approx \frac{\overline{D}_r^{last}}{\overline{v}_r}$$

where $\overline{D}_r^{last}$ and $\overline{v}_r$ are the slip-in-last-event and slip rate, respectively, averaged over the sub-sections involved in the r^{th} rupture, and

$$\overline{T}_r^{last} = \frac{\sum_{s=1}^{N} T_s^{last}}{S}.$$

Data are not available to directly test the agreement between the predicted intervals, $\Delta T_r^{pred} = T_r^{pred} - \overline{T}_r^{last}$, and the observed intervals, $\Delta T_r^{obs} = T_r^{obs} - \overline{T}_r^{last}$, where T_r^{obs} is the occurrence time of an event. However, from synthetic catalogs produced by physics-based simulators, we can examine the distribution of the ratio of the "observed" (i.e., simulated) to predicted intervals ($\Delta T_r^{obs} / \Delta T_r^{pred}$).

To evaluate the consistency of this formulation with simulator data, we need to establish a size threshold for events that "reset the clock" to T_s^{last}. For example, if a fault really does exhibit a

Gutenberg Richter distribution of earthquakes down to low magnitude, do the smallest events reset the clock? This could be a problem because the low amounts of slip associated with these little earthquakes would imply short recurrence intervals. To avoid this problem, at least for now until we can investigate the issue further, we limit our attention to earthquakes that rupture the full seismogenic thickness, and we use an M 6.5 threshold as a proxy for such ruptures. This is a sensible choice, because the fault-based events in our long-term Earthquake Rate Model are restricted to those that rupture the full seismogenic thickness.

Figure 8a shows the distribution of $\Delta T_r^{obs} / \Delta T_r^{pred}$ obtained from the Dieterich and Richards-Dinger (2010) simulator for all M≥6.5 events that occurred in a 22,000-year synthetic catalog. Compared to Figure 7, the results in Figure 8a are much more consistent with a BPT or log-normal distribution (with COV=0.3) now that we have changed the "probability of what" question from a point on the fault (Figure 7) to the time of the next event given knowledge that it will be the next one to rupture. For comparison, Figure 9 shows the same result as in Figure 8a, but where event times in the simulation have been randomized, which generates a Poisson-like distribution of recurrence intervals.

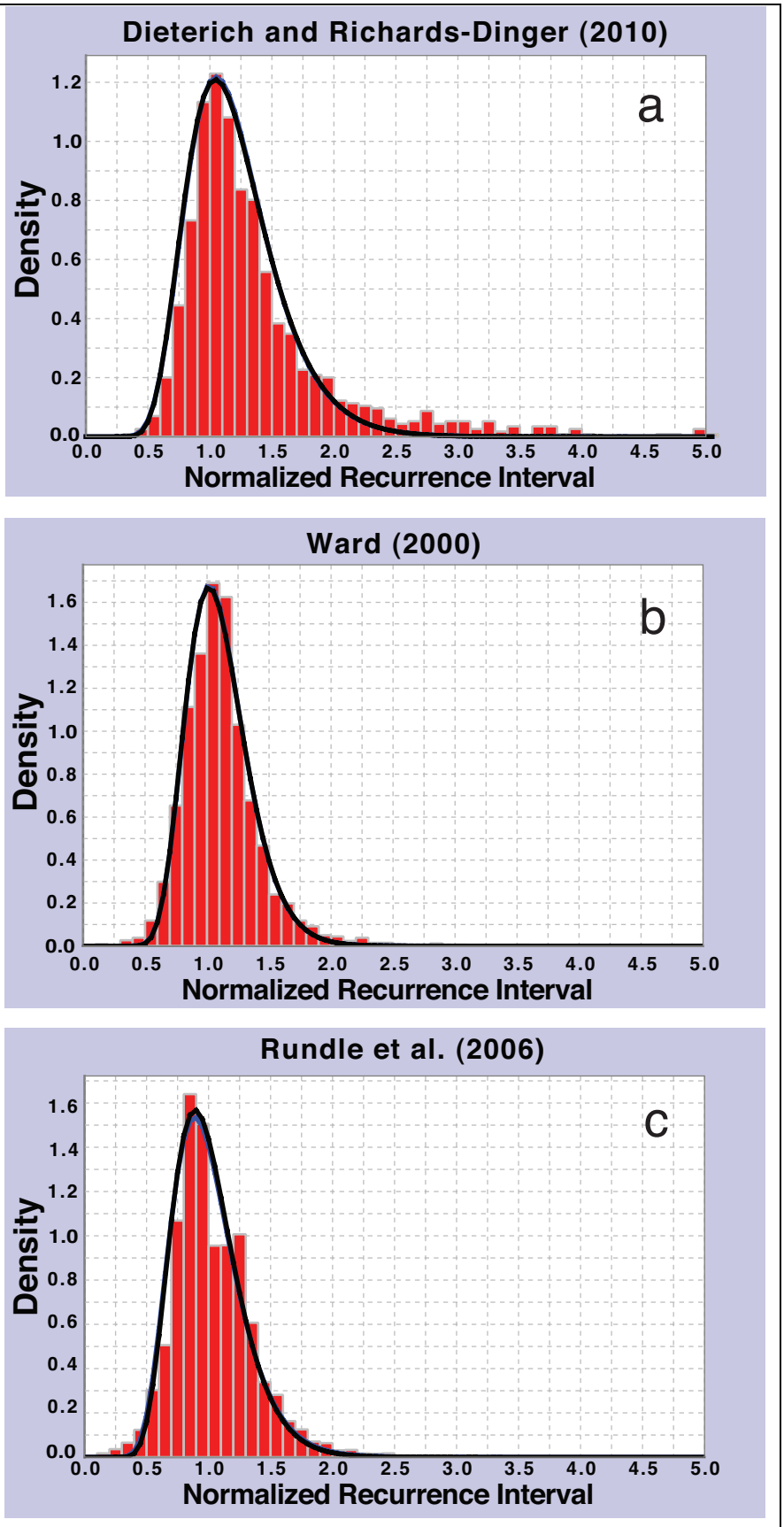
Figures 8a and 8c show the same results as in Figure 8a, but for the Ward (2000) and Rundle et al. (2006) simulators, respectively. The agreement between simulators in Figures 8 is encouraging, as all seem consistent with a BPT or log-normal distribution with a COV between 0.23 and 0.3. In particular, there are no short recurrence intervals in Figure 8, implying that the simulators do not generate M≥6.5 "aftershocks" or triggered events within the rupture surface of larger main shocks.

The preceding discussion assumes we know exactly *where* the next rupture will occur, leaving only the question of *when*. Because we do not know which of the many ruptures in our long-term model will be the next to occur, we must consider all possibilities. We propose using the following formula:

$$P_r \approx P_r^{Pois}(\Delta T^{longTerm}) \frac{P_r^{BPT}(\Delta T_r^{pred}, \overline{T}_r^{last}, COV)}{P_r^{Pois}(\Delta T^{pred})}$$

where $P_r^{Pois}(\Delta T^{longTerm})$ is the Poisson probability given the *long-term* rate of the r^{th} rupture, and $P_r^{BPT}(\Delta T_r^{pred}, \overline{T}_r^{last}, COV)$ and $P_r^{Pois}(\Delta T_r^{pred})$ are the probabilities computed from the BPT and Poission models, respectively, assuming the r^{th} rupture is the next to go. Note that the long-term Poisson probability of a rupture drops to zero as the fault is represented with an increasing number of smaller subsections. The ratio in the last term of the equation acts as a probability gain or reduction factor for the r^{th} rupture. Probabilities of events will be correlated to the extent they overlap spatially (i.e., share a large number of subsections). This model gives lower probabilities for events in areas that have recently ruptured and higher probabilities where they have not, as required to fill seismic gaps. It also allows some spatial overlap of events that occur in close temporal proximity (as observed for cascading sequences, such as the one on North Anatolian fault).

Figure 8. Distribution of of $\Delta T_r^{obs} / \Delta T_r^{pred}$ obtained from three physics-based simulators. The input model used in these simulations is the same as in Figure 7. Shown with a black and blue lines are best-fit BPT and log-normal distributions, respectively, with parameters as follows: a) mean=1.2 and COV=0.30; b) mean=1.1 and COV=0.23; and c) mean=1 and COV=0.27. The blue line is generally hidden below the black line.



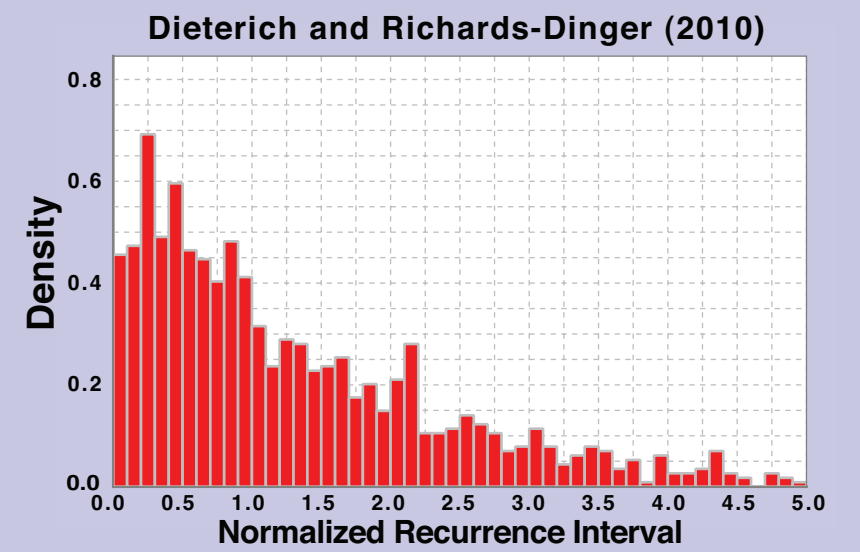


Figure 9. Same as Figure 8a, but where the event times were randomized uniformly over the simulation duration before computing normalized recurrence intervals.

Monte Carlo simulations with the model shown in Figure 6b demonstrate that this methodology is relatively unbiased in terms of event rates and moment rates. This consistency will be reconfirmed once the methodology has been implemented for the full UCERF3 model.

A seemingly reasonable alternative implementation would be the equivalent average *slip-predictable* model, where D_s^{last} above is replace with D_s^{next}, the slip in the next event. However, Monte Carlo simulations reveal a significant bias in this approach; the procedure preferentially chooses smaller events earlier in the cycle, thereby skewing both the overall rates and the magnitude-frequency distribution relative to the long-term model.

One drawback of this approach is that it requires knowledge of the amount of slip in the last event on each subsection D_s^{last}. Where this is unknown, we may be able to quantify the associated epistemic uncertainties for D_s^{last} from the long-term model, the assumed time dependence, and the observed open interval. Whether such uncertainty bounds add any value compared to a Poission model remains to be seen.

More work will be required over the remaining duration of the project to further justify this approach. The physics-based simulators used to evaluate the methodology will themselves have to be evaluated. To this end, we have been working with the SCEC Simulator Working Group, led by Terry Tullis, including a workshop on June 9, 2011 to obtain consensus on what simulators might be able to provide (see the appendix for Tasks R13 and P9 for more).

It is not surprising that the simulators imply elastic-rebound predictability, because they are based on the physics of elastic rebound. However, the results in Figure 8 suggest that elastic interactions within a complex fault system do not mask this type of predictability. Further simulations will be conducted to see how robust this conclusion is with respect to model tuning.

Other issues to be explored with this methodology include:

- What conditions most effectively reset the clock on each subsection (e.g., magnitude threshold, down-dip width of rupture)?
- What's the magnitude dependence of this predictability (e.g., does the COV decrease with increasing magnitude)?
- Are there differences in the predictability among faults or fault sections?
- How should tests of this methodology against either real or simulated observations be formalized?
- What are the implications of recent evidence that micro repeaters and laboratory earthquakes are neither time nor slip predictable (Justin Rubinstein's USGS Meno Park talk on Oct 27, 2010)?

The point of this section has not been to argue that elastic-rebound predictability indeed exists. Nor have we sought to find some perfect physics-based representation of this behavior. Rather, we've attempted to present a simple, relatively self-consistent, and probabilistic rule-based approach for modifying Poisson probabilities to be more consistent with elastic-rebound theory. The approach presented here seems as defensible as anything applied by previous WGCEPs, and the support from physics-based simulators is value added. If there exists a desire for elastic rebound to be represented in UCERF3, then so far this is the most promising approach we know of.

Apply Spatiotemporal Clustering Models

One goal for UCERF3 is to include spatiotemporal clustering to account for the fact that triggered earthquakes can be large and damaging. A good example of such clustering is the Joshua Tree, Landers, Big Bear, and Hector Mine sequence that occurred in southeastern California in the 1990s. A more recent example is the M 7.1 2010 Darfield earthquake in New Zealand, which produced a very damaging M 6.3 aftershock in Christchurch 5 months later. Even the great M 9.0 2011 Tohoku earthquake in Japan can be considered an aftershock of an M 7.2 earthquake that occurred just two days before. According to UCERF2, such chains of events are pure coincidence. The weight of opinion represented by the scientific literature, however, points to some kind of triggering phenomenon. If we accept this interpretation, then the next relevant question is whether such triggering is important for the policy decisions represented in building codes, earthquake insurance, and other forms of risk reduction. For example, would the California Earthquake Authority still be solvent had the Mohave sequence occurred in the LA basin? No one will know the answers to such questions until we have a model that can be used to explore implications.

Because the physical process responsible for earthquake triggering remain controversial (e.g., Felzer and Brodsky, 2006; Richards-Dinger et al., 2010), we plan to base UCERF3 on empirical, statistical clustering models (e.g., Ogata, 1988; Reasenberg and Jones, 1989, 1994). The Short Term Earthquake Probability (STEP) methodology of Gerstenberger et al. (2005) is one available approach (http://earthquake.usgs.gov/earthquakes/step/), which applies aftershock statistics to revise earthquake probabilities in real time for M≥3 events that occur throughout an

aftershock zone. The model we propose for UCERF3 builds on the STEP methodology. In so doing, we intend to address the following issues with STEP:

1. STEP requires that each observed event be associated with a single main shock, which becomes problematic where aftershock zones overlap, especially because these zones evolve with time as more data are collected.

2. In STEP, triggered events are sampled from a Gutenberg-Richter distribution between M5 and M8 everywhere in the region, which is inconsistent with the underlying long-term model (which constrains, for example, M8 events to occur on only a few faults such as the San Andreas).

3. There is nothing in the STEP formulation to prevent an M 8 event from immediately triggering itself. In fact, the likelihood of any particular event in STEP is greatest the moment after it actually occurs, which is inconsistent with elastic rebound.

4. Only one aftershock sequence influences probabilities at a given point in space (whichever sequence has the highest rate change).

5. STEP over-predicts both the total rate and moment rate of large earthquakes due to an inconsistency between the de-clustering applied in the long-term model (Gardner and Knopoff, 1974) and the aftershocks statistics used for spatiotemporal clustering (Reasenberg and Jones, 1989, 1994).

6. STEP combines different models based on a sophisticated analysis of generic, sequence-specific, and spatially variable parameters. This may improve predictability, but it introduces significant complexity, and it makes the tracking aleatory versus epistemic uncertainties a challenge, especially because the combination of models is spatially variable.

Here we outline the proposed methodology for UCERF3 and discuss how it addresses these issues. Recall that the Earthquake Rate Model provides the long-term rate of all possible events throughout the region (at some level of discretization and above some magnitude threshold). Assuming a uniform distribution of nucleation points on each earthquake surface, rupture rates can be translated into nucleation rates as a function of space (e.g., within each 0.1-by-0.1 degree bin). Likewise, an occurrence of a magnitude M event in a given bin can be mapped into one of the viable ruptures in the long-term Earthquake Rate Model (it's simply a matter of bookkeeping). With that background, the steps involved for the anticipated UCERF3 spatiotemporal clustering model include:

a) For a given start time and forecast duration, we collect all previously observed M≥2.5 events, plus randomly sampled spontaneous (non-triggered) events from our long-term Earthquake Rate Model (and including any empirical model and/or elastic-rebound modifications as described in the previous sections). We now have all candidate main shocks.

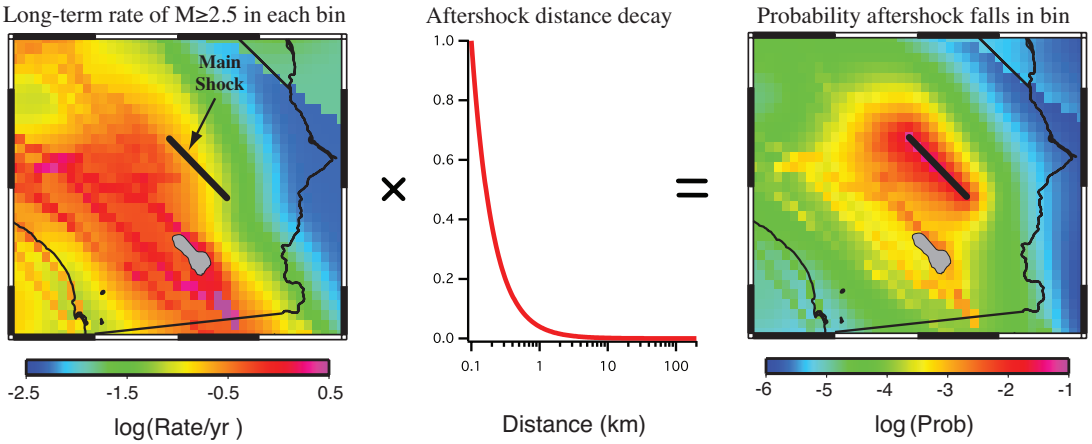
b) For each main shock in (a), we randomly sample times of occurrence of *primary* aftershocks from the ETAS formulation of the modified Omori law (e.g., Felzer, 2009):

$$n(t) = k10^{(M_{main} - M_{\min})}(c + t)^{-p}$$

where, lacking sequence-specific values, we use the following generic California parameters from by Hardebeck et al. (2008): k=0.008, p=2.34, c=0.095 days, and M_{min}=2.5.

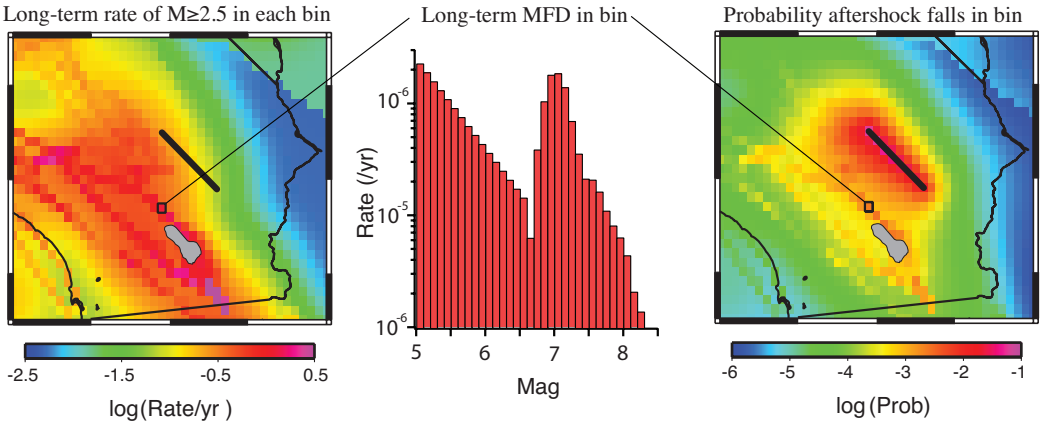
We next need to decide where each of these *primary* aftershocks occurs.

c) Using the long-term nucleation rate of M≥2.5 events throughout the region (from the Earthquake Rate Model as exemplified to the left below), multiplied by a spatial decay of $(R + R_{\min})^{-n}$, where R is the distance from the main shock fault surface (exemplified in the middle below), we randomly sample a nucleation grid-cell for the primary aftershock (from the distribution in the image on the right):



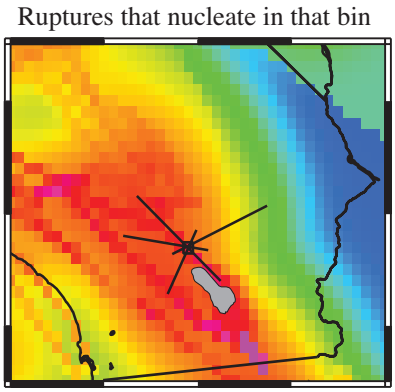
We next need to decide the magnitude of the primary aftershock.

d) Using the nucleation magnitude-frequency distribution for the grid cell chosen in step (c), which may or may not be Gutenberg-Richter, randomly sample a magnitude according to the relative rate of each magnitude:



We next need to decide which specific rupture (from the long-term Earthquake Rate Model) the primary aftershock represents.

e) Randomly sample a rupture from the long-term Earthquake Rate model according to the relative rate that each viable rupture (of that magnitude) nucleates in that grid cell:



We now need to collect secondary aftershocks from the primary aftershocks.

f) Repeat steps (b) thru (e) to get secondary aftershocks from all primary aftershocks, then likewise for tertiary events, and so forth until no more events are generated.

We now have a complete synthetic catalog for the chosen time span.

g) Repeat (a) through (f) to generate whatever number of alternative synthetic catalogs are needed to get statistically meaningful hazard or loss estimates.

This algorithm avoids having to assign each observed event to a main shock, and effectively allows multiple events to influence triggering probabilities at a given location. It also samples aftershocks directly from the long-term model, avoiding the inconsistency noted in item (2) above. This means a main shock is more likely to trigger an M8 earthquake if it occurs near a fault capable of generating such an event (e.g., the Bombay Beach scenario). Furthermore, by using long-term rates in steps (c) through (e) that have been corrected for elastic rebound influences as discussed in the previous section, we can prevent large, fault-based main shocks from sampling themselves as aftershocks. Furthermore, updating the model based on ongoing M≥2.5 seismicity will delineate any "blue" versus "red" lobes that would be present if static stress changes are important. Including events down to M 2.5 will allow sequences of smaller events to trigger larger, more distant earthquakes via secondary and subsequent triggering as demonstrated by Felzer et al. (2002) for the Landers and Hector Mine sequence. The smaller events can also connect together large earthquakes over long periods of time. For instance, the M≥2.5 seismicity following the 1971 San Fernando shows a decaying aftershock sequence that were still above background rates at the time of the Northridge earthquake 23 years later in 1994. Because the fraction of events that are triggered is magnitude independent, we simply reduce the rates of events in the background model by a common multiplicative factor so that the total simulated rate of events equals that observed (thereby avoiding the double counting problem with STEP at large magnitudes).

This algorithm generates suites of synthetic catalogs, each of which represents a viable sequence of triggered events. This is an advantage, because loss modeling is generally conducted using synthetic catalogs (referred to as "stochastic event sets") in order to account for the spatial

correlation of ground motions across a portfolio of sites. New for loss modeling will be UCERF3 event sets that include spatiotemporal clustering, not just samples from a Poisson process.

One advantage of aftershocks being sample from the long-term model is that losses for every event in that model can be pre-computed and stored (assuming the portfolio does not change with time). Then the losses for each synthetic catalog from UCERF3 can be easily (and quickly) aggregated, and statistics can be compiled over the different viable synthetic catalogs. This efficiency will facilitate operational earthquake loss forecasting.

The creation of synthetic catalogs differs from the current STEP implementation, which is not Monte-Carlo based, but rather gives the rates of events averaged over all viable sequences. The two approaches should be equivalent, all other things being equal, as long as a sufficient number of synthetic catalogs are sampled and averaged in the Monte Carlo approach. If there exists a need for a single, STEP-like forecast representing the average over all possible sequences, and the Monte Carlo approach is inefficient, then we can explore alternative formulations for achieving this.

Another contrast with STEP is that, as outlined here, we do not solve for and apply sequence-specific parameters (other than how ongoing seismicity changes subsequent forecasts). Our philosophy is to see how well our model does in simplified form before adding such sophistication. The CSEP testing center will be useful in terms of deciding what further complexities are warranted.

Example Implementation Using UCERF2:

We have implemented the software components needed for this spatiotemporal model using the *OpenSHA* platform. Here, we present example results when applied to the UCERF2 long-term model. Numerical implementation details, such as finite discretizations and Monte Carlo sampling of probability distributions, can be found in the *OpenSHA* documentation. For this report, we rely on the results themselves to provide evidence of code verification.

Figure 10 shows a simulated aftershock sequence for an M 7.25 Landers earthquake as represented in UCERF2, where the distance decay and minimum distance for this simulation are n=1.7 and R_{min}=0.3, respectively. According to Figure 10d, the expected number of M≥6.1 aftershocks is ~1.0 (consistent with Bath's law), the expected number of M≥6.5 aftershocks is ~0.5, and the expected number of M≥7.25 aftershocks (the main shock magnitude) is 0.06.

One important aspect of the implementation shown in Figure 10 is that the Landers main shock was prohibited from sampling itself as a direct or indirect aftershock. Figure 11 shows the expected number of aftershocks in an alternative simulation where Landers events are allowed to occur as aftershocks. This plot, which can be compared to Figure 10d, shows that the expected number of M≥6.1, M≥6.5, and M>=7.25 aftershocks is ~3, ~2 and 0.5, respectively. Furthermore, and perhaps more importantly, if an M>=6.5 event is indeed triggered, it has a 64% chance of being a re-rupture of Landers itself. If an M>=7.25 event is triggered, it has an 86% chance of being another Landers event. These probabilities are clearly inconsistent with global observations of large triggered events, where we just don't see such a high fraction of large

aftershocks occurring on same rupture surface as a large main shock (and by "large" we mean full-seismogenic thickness ruptures).

Ironically, these results imply that aftershocks statistics might indeed constitute the greatest evidence for elastic rebound. That is, given the distance decay of aftershocks, without elastic rebound the most likely event for any earthquake to trigger is itself. Of course this assertion depends not only on the ETAS-parameter values (especially n and R_{min}), but also on how the long-term models is constructed (e.g., how background seismicity in UCERF2 is modeled relative to fault-based sources like Landers). We are therefore currently conducting further tests, and if the conclusions hold, we intent to publish a paper outlining the evidence that aftershocks statistics may provide the strongest observational evidence for elastic rebound. If elastic rebound is indeed required by ETAS, then for UCERF3 we intend to apply the methodology outlined in the previous section.

Figure 12 shows a simulated aftershock sequence for an M 6.75 Northridge earthquake as represented in UCERF2. The same distance decay and minimum distance have been used for this simulation (n=1.7 and R_{min}=0.3, respectively) and Northridge events are prohibited from being sampled as aftershocks in this example.

According to Figure 12d, the expected number of M>=6.7 aftershocks is ~0.5, implying a 50% chance of triggering something larger than itself. This high probability is a consequence of the very characteristic (or non Gutenberg Richter) magnitude-frequency distribution for sampled aftershocks in this area (Figure 12c). In fact, the likelihood of sampling an M 6.7 aftershock from this model is about the same as sampling an M 5.6 earthquake. This is a direct reflection of very characteristic distribution implied by UCERF2 in this area, as the magnitude frequency distribution of sample aftershocks is simply that of the long-term model weighted by distance from the main shock. Not only is this M>=6.7 triggering probability dubiously high, but these Northridge ETAS simulations often run away and never converge (because large triggered events can keep on triggering other large events). Note that these issues become worse if the Northridge main shock is allowed to re-rupture itself as an aftershock.

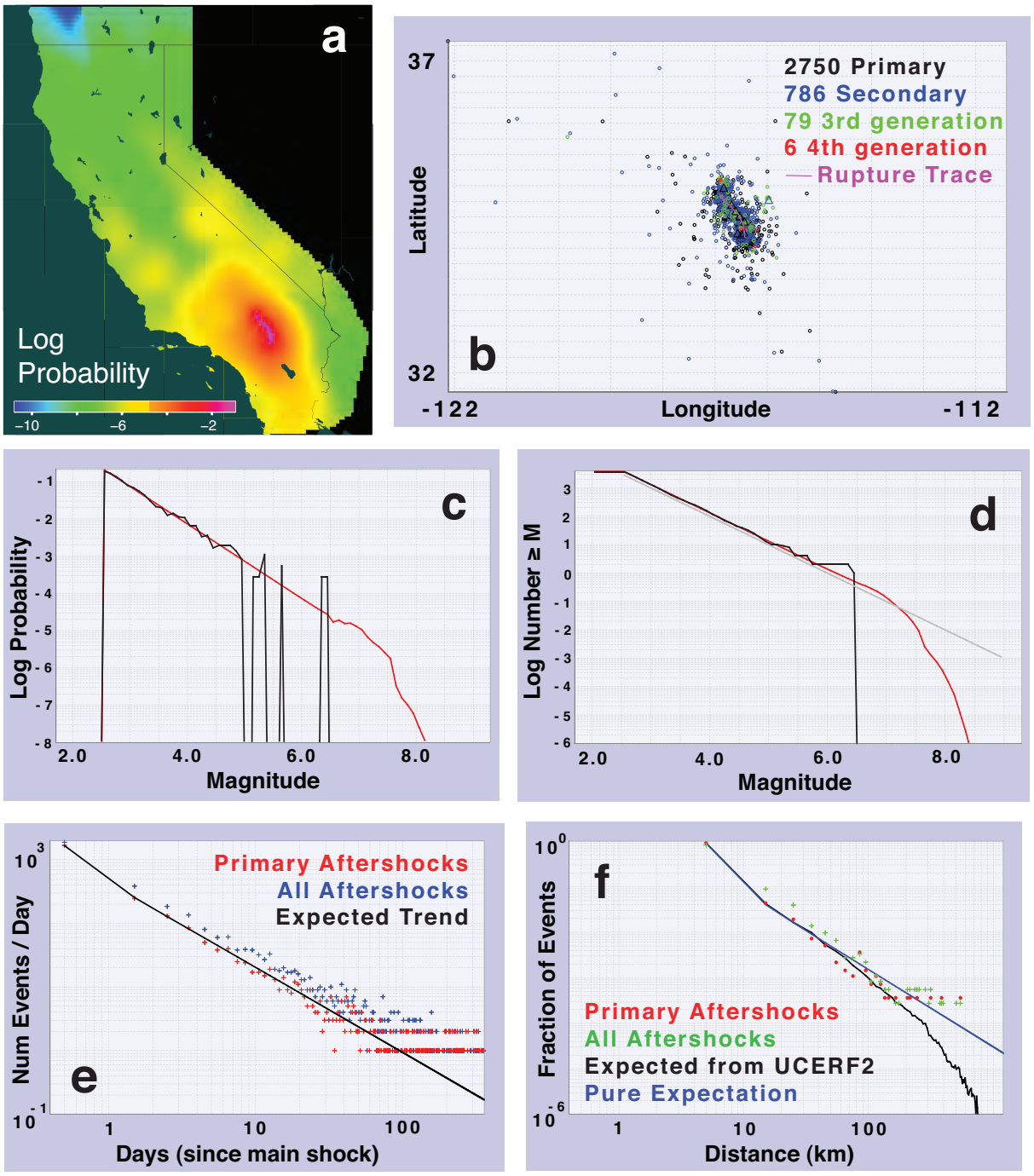


Figure 10. Results of an ETAS-simulated aftershock sequence for 360 days following the occurrence of an M 7.25 Landers earthquake as represented in UCERF2, and as described in the text. **a)** Spatial probability distribution for primary event epicenters (in 0.1 by 0.1 degree lat-lon bins). **b)** Map of simulated aftershock hypocenters, where the number of events in each generation is indicated at the upper right. **c)** Expected magnitude probability distribution for all aftershocks in 0.1 magnitude bins (red), together with that sampled in the simulation (black). **d)** Expected number of aftershocks greater than each magnitude (red), the numbers sampled in the simulation (black), and a GR distribution with b=1 for comparison. **e)** Temporal decay for the sequence (in 0.5-day bins). f) Spatial decay for the sequence (10 km bins), where the difference between "Expected from UCERF2" and "Pure Expectation" is due to a lack events outside California.

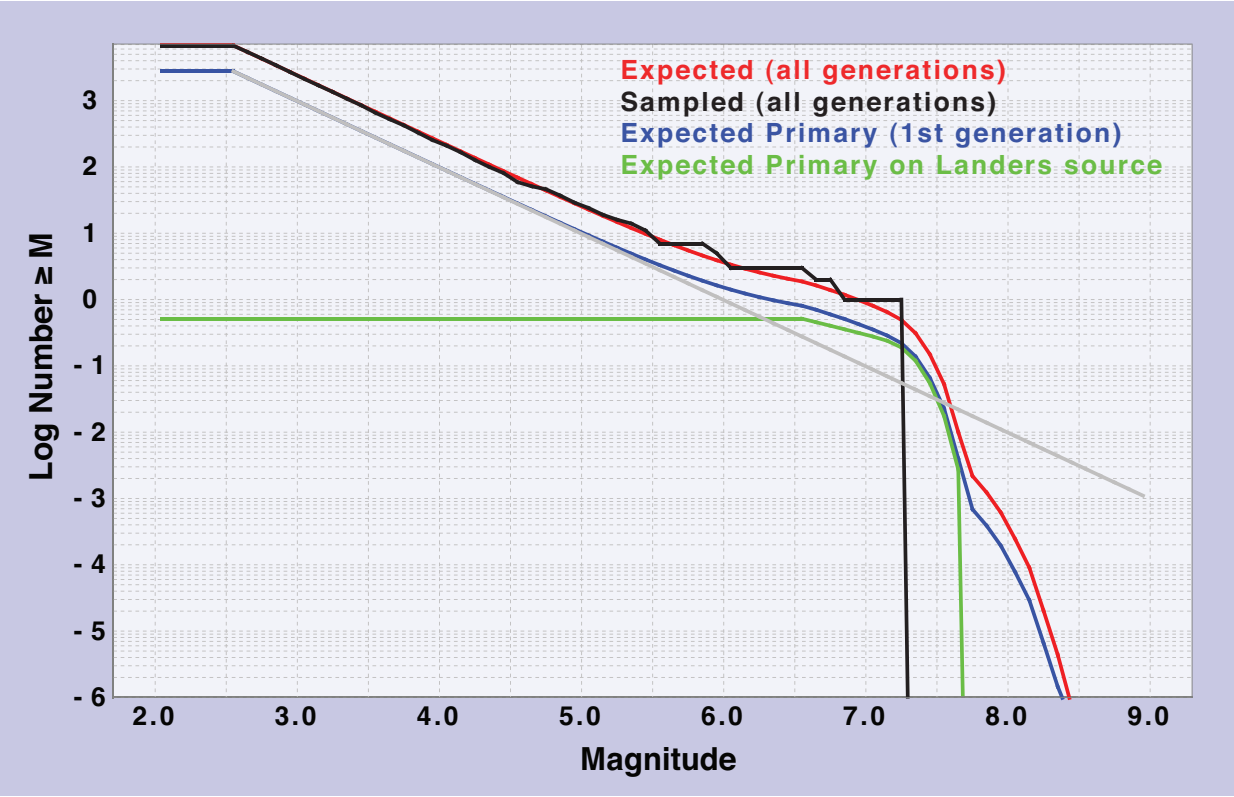


Figure 11. Same as Figure 10d, but where the Landers main shock is allowed to trigger itself as a primary or subsequent-generation aftershock. Also shown here are the expect number of primary events only (blue), as well as the number of primary events expected to be triggered on the Landers source (green). The point here is that a large fraction of triggered events are on the Landers source itself, with the fraction generally increasing with magnitude.

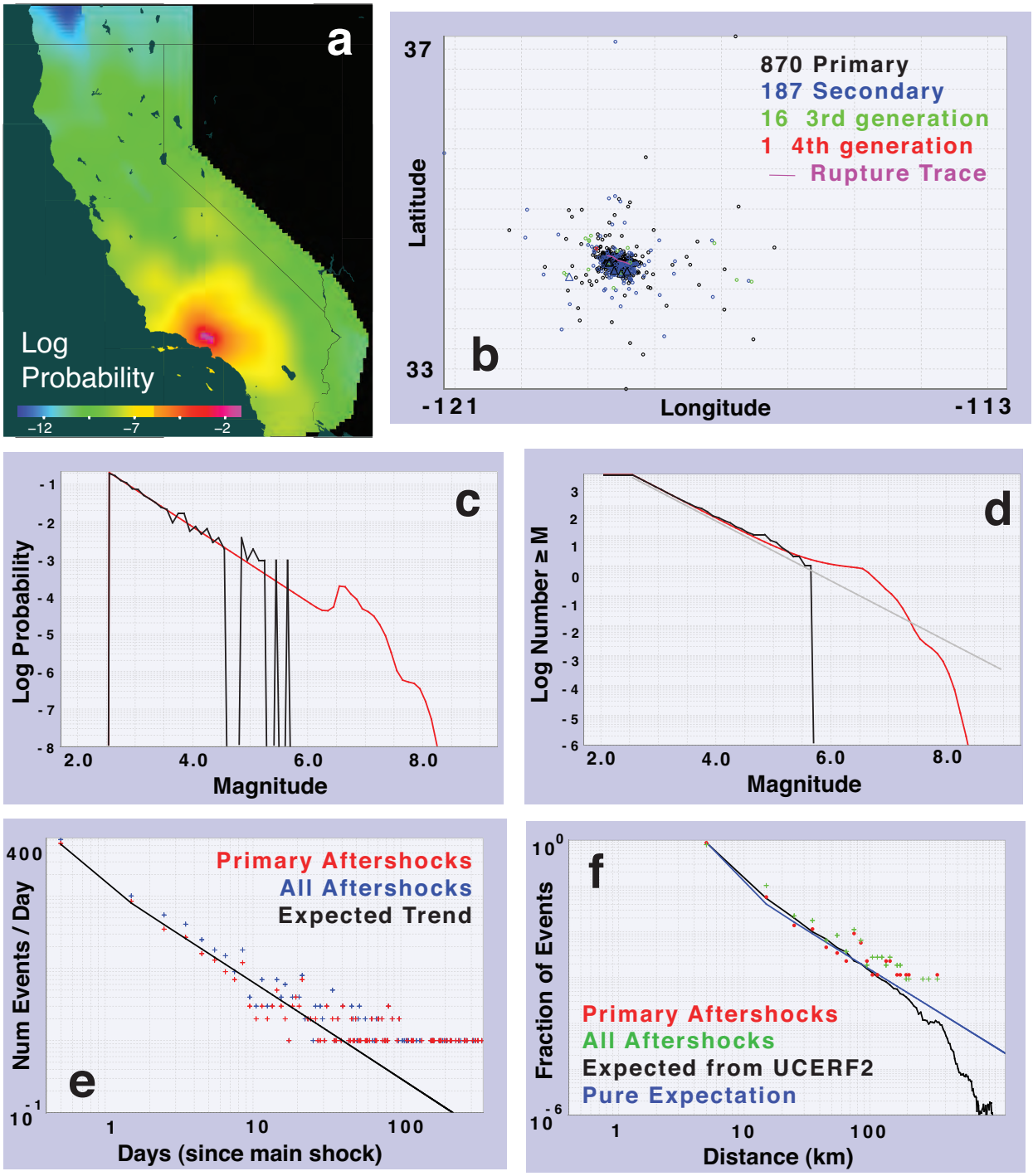


Figure 12. Same as Figure 10, but for an M 6,75 Northridge main shock as represented in UCERF2.

This raises the issue of just how non Gutenberg Richter the long-term model can be over sub-regions without running into problems with the methodology proposed here. Figure 13 shows the magnitude frequency distribution from the UCERF2 long-term model in the vicinity of Northridge, revealing again a very non-Gutenberg Richter distribution. While this region represents one of the more extreme examples from UCERF2, and may itself stretch scientific credibility, many other regions exhibit this behavior to one degree or another (e.g., see Figure 19 in the UCERF2 Report). Recall that the inversion approach to constructing the UCERF3 earthquake rate model explicitly includes a regional Gutenberg Richter constraint, so we have a mechanism for invoking this to the extent it is needed. But the question is, again, how much do we need to invoke this constraint, and will doing so create other problems?

If problems remain with respect to non-Gutenberg-Richter regions in some branches of the Earthquake Rate Model, one option will be to sample the aftershocks from a pure Gutenberg-Richter distribution while still sampling from those events available in the long-term model (according to their relative nucleation rates). Thus, aftershocks will still come from, and therefore be consistent with, the long-term model (an improvement over STEP), but we will have lost the feature where long-term ETAS simulations have rates that match our Earthquake Rate Model (so we won't have solved that problem with STEP).

Near term work also includes comparing aftershocks sequences simulated using this methodology against those actually observed in California for as many sequences as possible (Task P5 below). This will not only indicate the extent to which sequence-specific parameters are important, but also the extent to which the model fails to reproduce features of real sequences. An example of the latter would be a higher rate of aftershocks occurring near large, mature faults (e.g., Powers and Jordan, 2010), and where this behavior is not already manifested in background seismicity rates.

Other issues that will need consideration include:

- For an operational system, how far back in time do we need to go in collecting the main shocks that will be carried forward as such in the simulations? Is this magnitude dependent? Is there a tradeoff here with declustering assumptions in developing the background seismicity model? This issue is addressed further in the appendix entry for "Task R9 - Smoothed Seismicity Model".

- Influence of spatial smoothing in the development of the background seismicity model on the spatial distribution of aftershock sequences (e.g., does tighter smoothing account for the effect identified by Powers and Jordan (2010)).

- Computation time.

- The spatiotemporal clustering model is not completely independent from the elastic rebound model (e.g., the COV in the latter must be somewhat influenced by aftershock statistics in the former).

- What is the physical difference between a multi-fault rupture (added to the earthquake rate model above) and a quickly triggered separate event (as modeled here)? Could slip-length scaling distinguish these populations?

We don't anticipate the implementation articulated here to be the final model, but rather an appropriate starting point for exploring issues and justifying further complexities. The object-oriented, extensible implementation will lend itself to such future improvements.

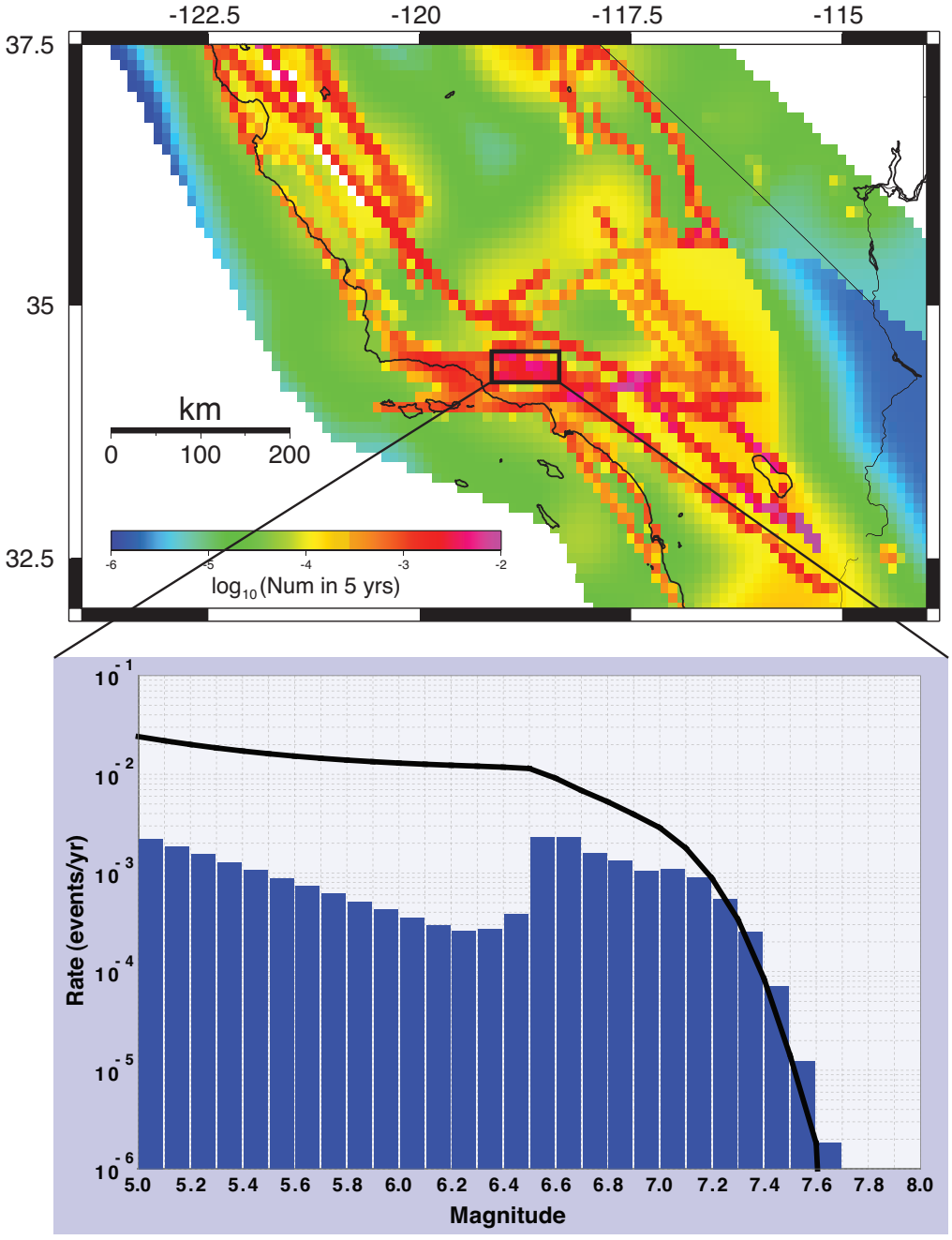


Figure 13. Map above shows the expected number of M≥6.5 hypocenters in 0:1° × 0:1° bins in a 5 yr period predicted by UCERF2, and the plot below shows the incremental (blue) and cumulative (black) magnitude-frequency distributions for events that nucleate inside the black box show in the map. UCERF2 sources that nucleate inside this box include: *"Holser, alt 1", "Northridge", "Oak Ridge (Onshore)", "Oak Ridge Connected", "Pitas Point Connected", "San Cayetano", "San Gabriel", "Santa Susana, alt 1", "Santa Ynez (East)", "Santa "Ynez Connected", "Sierra Madre (San Fernando)", "Sierra Madre Connected", "Simi-Santa Rosa", "Ventura-Pitas Point","Verdugo"*, and background gridded seismicity. The box is defined by latititudes 34.25 and 34.55 and longitudes -119.15 and -118.35.

Logic-Tree Branches

Possible logic-tree branches here include the following:

- Application (or not) of the revised Empirical Model
- COV used in the average time-predictable, elastic-rebound calculations
- Uncertainty in amount of slip in most recent along each fault
- Alternative ETAS parameters and perhaps a sequence-specific option
- Alternative ways of defining which observed events will be carried forward as active sequences in the operational forecast

These options will be explored more fully as we continue with the implementation.

Current Status

The current status of each time-dependent component has been discussed in the respective, dedicated sections above.

De-Scoping Options

De-scoping options are limited here because the elastic-rebound methodology used in UCERF2 no longer seems viable. There are also no previous ETAS implementations to fall back on, so de-scoping of this would probably mean removing it all together.

Ongoing Tasks

Table 4. Task List for the Development of Earthquake Probability Models.
Notes for Table 1 apply here.

	Task	Description	Participants
Earthquake Probability Model(s)	**P1) Address "Empirical" model**	Examine robustness of apparent rate changes given reevaluation of historical catalog (task above) and for different time and space slices. We anticipate indeed having an empirical model with higher spatial resolution than in UCERF2. *See appendix for further discussion of this task.*	**Felzer**, Parsons
	P2) ETAS explains Empirical Model?	Investigate whether ETAS is sufficient to explain the observed rate changes in the empirical model. *See appendix for further discussion of this task.*	**Felzer**, Page, Michael?
	P3) Coulomb Stress explains Empirical Model?	Investigate whether static coulomb stress changes can explain the observed rate changes. *See appendix for further discussion of this task.*	**Parsons**, Powers?, Pollitz
	P4) Develop self-consistent renewal models	Develop self-consistent, elastic-rebound-motivated renewal models, which are currently lacking for anything but strictly segmented models. *This issue is already described above in this report and also in Appendix N of the UCERF2 Report.* Remaining tasks involve reevaluation of the approach when simulator results based on the UCERF3 deformation models are available.	**Field** & Page
	P5) Evaluate ETAS implementation	This task is aimed at evaluating the ETAS implementation by, for example, comparing simulated events to those that have occurred in the instrumental record. This will help define generic parameters and determine the value added by sequence-specific parameters. We are in the process of obtaining finite-fault representations of the larger historical events in California, which will be needed for this analysis. *See appendix for more discussion.*	**Felzer & Michael**, Page, Field, Powers
	P6) Evaluate Agnew and Jones	Does the Agnew and Jones (1991) approach constitute a unique and implementable model? *See appendix for more.*	**Michael**
	P7) Evaluate Static Stress Change Models	Do static-stress change models constitute unique and implementable models (from an operational perspective)? *See discussion of Task P3 in appendix for further info.*	**Parsons**, Powers
	P8) Evaluate other time dependencies	Are there important rate variations at other time scales (e.g., implied by empirical model, or by the mode switching identified by Rockwell and Dolan in paleo data). How do we model these? *See appendix for further discussion of this task.*	**Hardebeck,** Dolan?
	P9) Evaluate Physics-based simulators (for probabilities)	Investigate implications and applicability of physics based simulators for inferring elastic-rebound probabilities and clustering effects. Do this in conjunction with the ongoing SCEC simulator working group being led by Terry Tullis. *See appendix entry fro related task Task R13 for discussion.*	**Field**, Michael, Tullis, Dieterich, Richards-Dinger, Ward, Rundle, Pollitz, Beeler
	P10) Compile Slip in Last Event Data	Slip in last event data are needed for the elastic-rebound probability calculations proposed here. This task involves collecting published historic, paleoseismic and microgeomorphic data to define this parameter along as many faults as possible in the model. *See appendix for further discussion of this task.*	**Weldon**, Biasi, Arrowsmith, Hudnut, Dawson, and Madden

Model Implementation

This sections lists tasks related to model implementation.

Remaining Tasks:

Table 5. Task related to model implementation.

Notes for Table 1 apply here.

	Task	Description	Participants
Implementation Issues	**I1) Documentation and access to input data and results**	UCERF2 created issues with respect to the delivery of data and model results, especially with respect to how the NSHMP provides this information. Relaxation of segmentation and allowing fault-to-fault ruptures will only compound these issues, so we are actively working on resolving them. We do not anticipate any big issues.	**Field**, Haller, Petersen, Powers, Wills, Dawson, Jordan, Milner, Husband
	I2) Loss Modeling Tools	Develop loss-modeling tools to help quantify what model uncertainties are important (a "tree-trimming" tool). Such tools would also allow us to quantify the practical implications of UCERF3 enhancements (e.g., spatiotemporal clustering). Note that this activity is not part of the CEA-sponsored scope of work. Funding for this activity remains in question.	**Porter**, Field, Luco
	I3) Address potential issues for the user community	User-community issues that will be raised by UCERF3 include 1) how they will deal with much larger event sets (due to relaxing segmentation and allowing fault-to-fault ruptures); 2) changes in the definition of "aftershocks" and how or if they're removed from the complete UCERF3 model (this is important because building codes currently have aftershocks removed, and CEA's earthquake insurance policies have specific and important wording with respect to the definition and treatment of aftershocks); and 3) how hazard and loss calculations can most efficiently be conducted from an operational earthquake forecast (where probabilities may be changing in real time). We have had several one-on-one meetings with the various loss modeling companies, and they don't seem too concerned. A workshop is nevertheless scheduled for Oct 24th and 25th, 2011 to address these issues.	**Field**, Luco, Petersen, Porter, Campbell
	I4) Address IT issues in deploying an operational earthquake forecast	UCERF3 will involve real-time interoperability with seismic network information in order to update probabilities immediately following significant events. A robust implementation will be very important, including how the model interfaces to user communities. What exactly are we promising, and do we have the support for long-term operations given this will require dedicated resources that don't currently exist? These issues can only be addressed as we obtain a better idea of exactly what we intend to implement.	**Powers**, Field, Milner, Jordan, Gerstenberger, Jones, Earle, Petersen, Buland, Michael

	I5) Model testing	As soon as the project plan reaches a more advanced state of refinement (e.g., based on prototyping efforts), we will then compile a list of key assumptions being made in UCERF3, and then host a workshop to discuss how the various assumptions might be formally tested (listed below under workshops). This activity is being coordinated with the CSEP effort. Note: we had intended to have such a workshop in May 2011, but this was postponed under the recommendation of the MOC Chair.	**Schorlemmer,** Jackson, Jordan, Field, Felzer, Page, Michael, Weldon

Other Possible Tasks:

- Formalize rules for gray literature data inclusion to avoid some double standards that were applied in the UCERF2 development.

Planned Workshops & Review Meetings

The following two tables list SPR review meetings and workshops, respectively. Workshops, which will include participants from the broader community, are aimed at addressing one or more topical issues. The review meetings will involve formal evaluations by the SRP and possibly by members of CEPEC, NEPEC, and the CEA science evaluation team. The topics and dates are subject to change as plans evolve, and it is possible that some of the review meetings and workshops will be combined for efficiency. Not listed here are the many anticipated meetings among WGCEP participants, as well as those that might be convened by the USGS National Seismic Hazard Mapping Program to satisfy their programmatic requirements. For example, the USGS has already held a workshop on Cascadia forecasting that is not listed below.

Planned SRP Review Meetings

Review Meetings	Description	Date
1) Methodology Assessment	An overview of both Report #1 (Issues and Research Plan) and Report #2 (Proposed Solutions to Issues)	November, 2010 (about a month before Report #2 is due)
2) Proposed UCERF3 Plan	A comprehensive overview of the UCERF3 implementation plan (Report #3) ; 1st day open to broad community.	June 13-14, 2011 (~2 weeks before SRP Report (#4) is due)
3) Preliminary UCERF3 Model	A comprehensive overview of the preliminary UCERF3 model (Report #6).	Mid April, 2012 (~2 weeks before SRP Report (#7) is due)

Planned Workshops

Title	Description	Date
Past		
UCERF3 Planning Meeting	This workshop, which had broad community participation, was to discuss the goals and anticipated issues with building UCERF3.	Feb. 17-18, 2010
Incorporating Geodetic Data into UCERF3	This workshop began a comprehensive scientific discussion of how to incorporate GPS constraints on strain rates and fault slip rates into UCERF3.	April 1-2, 2010
Statewide Fault-Model & Paleoseismic Data	This workshop addressed what changes are in order for the statewide fault model, with particular emphasis on our understanding of fault endpoints and potential biases in slip-rate estimates for the lesser faults. This workshop also addressed paleoseismic trench data and its interpretation.	April 6 & 8 in S. & N. CA, resp., 2011
UCERF3 Deformation Models	This workshop will present new deformation models based on a more sophisticated analysis and treatment of GPS data, as well as present the vision for making further progress in the future. 2nd day for core only.	June 4-5, 2011
Time-Dependent Models	This workshop will address what represents both "best-available science" and implementable models with respect to time-dependent probabilities in	June 8, 2011

	our UCERF3 operational forecast. Of particular emphasis here will be how to apply elastic rebound in un-segmented fault models, how to represent spatiotemporal clustering (e.g., ETAS), and the interpretation of the empirical model (apparent seismicity rate changes).	
Use of Physics-based Simulators	This workshop, which will be co-convened with the SCEC earthquake-simulators working group, will address what physics-based simulations can provide with respect to defining both long-term earthquake rates and shorter-term probabilities. As earthquake simulators hold promise for addressing many of our current goals and challenges, this workshop will be critical for gauging the maturity of these models.	June 9, 2011
Instrumental & Historical Seismicity	This workshop will review issues and proposed solutions with respect to the historical and instrumental earthquake catalogs, with particular emphasis on how this influences: a) the association of events to specific faults; b) inferred temporal variations in earthquake rates; and c) regional magnitude-frequency distribution estimates. This workshop will also address best practices for estimating the spatial distributions of a-values, maximum magnitudes, and focal mechanisms for background seismicity (events off our explicitly modeled faults).	June 10, 2011
Distribution of Slip in Large Earthquakes	This workshop will address the following: a) slip distribution along strike, especially when multiple faults are involved; b) slip distribution down dip and whether larger events penetrate deeper (important for resolving current mag-area discrepancies); and c) theoretical and observational constraints on the propensity for ruptures to jump from one fault to another.	June 11, 2011
Future		
UCERF3 Plan Overview (Emphasizing the Grand Inversion) for Users	This workshop will constitute a complete overview of the anticipated UCERF3 model to PSHA experts, emphasizing the Grand Inversion. The timing of this workshop is to enable feedback to influence the final product.	Oct 24, 2011
Joint UCERF3 NGA-W2 Workshop on Common Issues	This workshop will address the following issues of joint interest and concern between UCERF3 and NGA-W2: 1) possible double counting of aleatory mag-area variability between GMPEs and ERFs; 2) computation of NGA prediction variables for multi-fault ruptures (possible ambiguities?); 3) definition of aftershocks, and whether they produce different ground motions; 4) development of a joint Glossary of terms?; 5) Issues associated with NSHMP implementation of NGAs (e.g., treatment of gridded seismicity).	Oct 25, 2011
Overview of Preliminary UCERF3 Model for User Communities	This workshop among key stakeholders and general users will present the preliminary UCERF3 model. Particular emphasis will be given to dealing with the significantly increased number of events, given the relaxation of segmentation and inclusion of multi-fault ruptures, as well as challenges associated with using a real-time, operational forecast.	April, 2012?
Assumptions & Model Testing	This workshop will review likely UCERF3 assumptions and discuss how these might be formally tested.	Date to be determined
Overview of Final UCERF3 Model for User Communities	This will present UCERF3 to key stakeholders and user communities with the goal of facilitating use of the model. Note this date follows our final delivery to CEA.	Sept., 2012

Key Components

This table lists key components needed for UCERF3 (elements needed for the actual implementation).

Component Name	Description	Outstanding Issues	Related Task(s); Person in Charge	Final Delivery
Fault Models	Alternative fault models, where each is composed of a list of fault sections (where each section is defined with a trace, ave dip, ave upper- and lower-seismogenic depth, ave rake, and width).	1) How many alt models?	F1; Dawson	10/01/11
Geol. Slip Rates	Geol. slip-rate estimates for points on faults, which are needed by the deformation modelers	1) How extensive can this compilation be? 2) Stored in Paleo Sites Database?	D2; Dawson	10/01/11
Deformation Model Slip Rates & Off-fault Deformation	Slip-rate constraints for points on the above fault model(s), plus spatially gridded values representing residual off-fault deformation (i.e., not including the explicitly modeled faults).	1) How to define alternative models given infinite tradeoffs (to span hazard implications? to span off-fault deformation possibilities?) 2) How densely along strike can these be sampled? 3) How block model results are mapped back on the faults?	D1; Thatcher & Johnson	12/01/11
Aseismiciy estimates	Assign an aseismicity parameter to each fault section in each fault model.	1) Interpretation of this parameter (related to upper seis depth defined above? applies as a reduction of area?)	D4; Weldon	12/01/11
D_{r}	Average slip for the r^{th} rupture (both the mean over multiple occurrences and the intrinsic variability).	This work is ready for review.	R2; Shaw	12/01/11
D_{sr}	Slip on the s^{th} subsection for the r^{th} rupture.	1) Are the current slip-along-length options sufficient (sqrt -sine, boxcar, slip-rate proportional), or are multi-rainbows needed? 3) What are the uncertainties and intrinsic variabilities of Dsr? 4) How could we apply char slip?	R1; Biasi	12/01/11
f_s^{paleo}	Paleoseismic recurrence interval estimates at points on faults	?	R3; Parsons	12/01/11
P_r^{paleo}	Probability of seeing a given rupture in a Paleo	?	R4; Weldon	12/01/11

	trench			
Multi-fault rupture Constraints	Expert-opinion assessment of multi-fault rupture likelihoods based on detailed geol maps, micro seis., global jumping stats, coulomb calcs, and dynamic modeling.	1) Global stats are based on obs taken after large events, rather than on the more limited info available before a rupture. 2) Tinker-toy rules or case-by-case analysis?	R6; Biasi	12/01/11
Additional inversion constraints	What Weldon has been advocating: 1) ave paleo slip per event (and COV?) and correlation of timing of paleo events	1) How do we get these constraints and how do we apply them in the inversion	R12; Weldon	12/01/11
Revised earthquake Catalog	Update the earthquake catalog with recent/revised data, plus go down to M 2.5 and include finite fault models for large ruptures. Also quantify catalog completeness.	1) Can/should catalog completeness be quantified as a probability of missed events as a function of space, time and magnitude?	R7 & R8; Michael & Felzer	12/01/11
Obs mag-freq-dists and/or GR parameters	Obs MFD for arbitrarily defined regions (or at least for all CA, N. CA, and S. CA) including uncertainties. Also provide GR params (at least a-value) for arbitrary sub-regions (polygons) to use as inversion constraint. These should include aftershocks.	1) Can we distinguish between rates on vs off the modeled faults? 2) Can a-value estimates include uncertainties associated with possible temporal variations? 3) How to decide what to include here versus what's carried forward by ETAS?	R8; Michael & Felzer	12/01/11
Smoothed Seismicity Models	For spatial dist of GR a-values	1) How to ensure this is consistent with the component directly above (including what's carried forward in ETAS)? 2) Do we have different versions depending on whether Empirical model has been applies?	R9; Felzer	12/01/11
Inversion Implementation	Implement efficient simulated annealing inversion methodology for solving for rupture rates	1) Computationally feasible for everything we want to include in the inversion?	R5; Page	12/01/11
Spatial distribution of focal mechs.	To use in off-fault seismicity forecasts	1) How (or if) to separate the off-fault component from their compilation?	R11; Jackson	12/01/11
Aleatory mag-area var.	Determine whether we're double counting with GMPEs in PSHA calcs.		R14; Field	12/01/11
Cascadia	Develop the complete model		R15; Frankel	05/01/12
Empirical Model	Provide updated and more refined (than in UCERF2) empirical model.	1) Are we closer to a physical explanation?	P1, P2; Felzer	12/01/11

Renewal Models	Develop methodology to be applied	1) Methodology still supported by more recent simulation results?	P4; Field	12/01/11
Slip in last event(s)	Compile data need for renewal model	?	P10 & R12; Weldon	12/01/11
ETAS Parameters	Determine those to be applied by analyzing our methodology against as many CA sequences as possible.	1) Event specific params? 2) Does ETAS really adequately model the 1990s Mojave sequence? 3) How do we decide what obs events get carried forward in the simulation?	P5; Michael, Felzer, or Field	12/01/11
ETAS Implementation	Implement the methodology in the overall framework	1) Do we need more GR in long-term model to prevent runaway sequences? 2) Is elastic rebound needed to make this work, and what exactly would this get applied to (what is "self")? 3) Physical diff between multi-fault rupture and quickly triggered event?	Field	12/01/11

Outstanding Issues

This section lists some of the potentially vexing, deeper issues we face.

- The results of the inversion for the earthquake rate model depend critically on how average slip varies along strike, how slip rates transition from one fault to another, and the proximity of neighboring faults (plus other features that influence the likelihood of mult-fault ruptures). Given the uncertainties in all of these, are the results anything more than an educated guess?
- The importance of including elastic rebound in the ETAS simulations raises the question of what exactly do the elastic rebound corrections get applied to. Will the width field being added to faults in the fault-section database (as a proxy for whether each represents a narrow surface or broader, braided fault zone) be adequate for this?
- What is the difference between a multi-fault rupture and a separate event that happens to be triggered quickly? Is the distinction important, and can it be made in a meaningful way in UCERF3 (e.g., how can we avoid double counting)?
- How aggressively will we have to apply the Gutenberg-Richter constraint to sub-regions in order for the ETAS model to work properly?
- Given the difficulty in declustering catalogs, coupled with likely temporal variability in seismicity rates, is smoothed seismicity a reliable basis for setting rates for "off-fault" events in our long-term model (question raised in the appendix entry for Task R9)? Is the "off-fault" deformation to be provided by the deformation modelers a better constraint?

While we acknowledge that these are indeed vexing issues, it's important to point out that they apply equally to, and in no way provide additional support for, previous models such as UCERF2.

References

Agnew, D. C., and L. M. Jones (1991), Prediction probabilities from foreshocks, *J. Geophys. Res.* ***96*****(7)**, 11,959-911,971.

Andrews, D. J., and E. Schwerer (2000). Probability of rupture of multiple fault segments, *Bull. of the Seism. Soc. Am.* **90**, 1498-1506.

Bakun, W. H. , B. Aagaard, B. Dost, W. L. Ellsworth, J. L. Hardebeck, R. A. Harris, C. Ji, M. J. S. Johnston, J. Langbein, J. J. Lienkaemper, A. J. Michael, J. R. Murray, R. M. Nadeau, P. A. Reasenberg, M. S. Reichle, E. A. Roeloffs, A. Shakal, R. W. Simpson and F. Waldhauser (2005). Implications for prediction and hazard assessment from the 2004 Parkfield earthquake, *Nature* **437**, 969-974, doi:10.1038/nature04067.

Bird, P. (2009). Long-term fault slip rates, distributed deformation rates, and forecast of seismicity in the western United States from joint fitting of community geologic, geodetic, and stress direction data sets, J. Geophys. Res., **114**, B11403, doi:10.1029/2009JB006317.

Bufe, C. G., P. W. Harsh, and R. O. Burford (1977). Steady-state seismic slip--a precise recurrence model, *Geophys. Res. Letters* **4**, 91-94.

Dieterich, J.H. and K.B. Richards-Dinger (2010). Earthquake recurrence in simulated fault systems, *Pure And Applied Geophysics*, in press.

Felzer, K. R., T. W. Becker, et al. (2002). Triggering of the 1999 Mw 7.1 Hector Mine earthquake by aftershocks of the 1992 Mw 7.3 Landers earthquake, *J. of Geophys. Res.* 107 (B9).

Felzer, K. R. & E. E. Brodsky (2005). Testing the stress shadow hypothesis, *J. Geophys. Res.*, **110**, doi:10.1029/2004JB003277, 2005.

Felzer, K. R. & E. E. Brodsky (2006). Decay of aftershock density with distance indicates triggering by dynamic stress. *Nature* **441**.

Felzer, K. (2009). Simulated Aftershock Sequences for an M 7.8 Earthquake on the Southern San Andreas Fault, *Seism. Res. Lett.* **80**, 21-15.

Field, E. H., and M. T. Page (2010). Estimating earthquake-rupture rates on a fault or fault system, *Bull. Seismol. Soc. Am.*, In Press.

Gerstenberger, M., et al. (2005), Real-time forecasts of tomorrow's earthquakes in California, *Nature*, *435*, 328-331.

Hanks, T. C., and W. H. Bakun (2008). M-log A observations of recent large earthquakes, *Bull. Seismol. Soc. Am.*, ***98***, 490.

Hardebeck, J. L., K. R. Felzer, and A. J. Michael (2008). Improved test results reveal that the accelerating moment release hypothesis is statistically insignificant. *J. of Geophys. Res.* **113**, B08310; doi:10.1029/2007JB005410.

King, G. C. P. and S. G. Wesnousky (2007). Scaling of Fault Parameters for Continental Strike-Slip Earthquakes, *Bull. Seism. Soc. Am,* **97**, 1833-1840, doi: 10.1785/0120070048.

Lawson, C. L., and D. J. Hanson (1974). Solving Least Squares Problems. *Prentice-Hall*, Englewood Cliffs, New Jersey.

Ogata, Y. (1988), Statistical models of point occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.*, *83*, 9-27.

Parsons, T., and E. L. Geist (2009). Is there basis for preferring characteristic earthquakes over Gutenberg-Richter distributions on individual faults in probabilistic earthquake forecasting?, *Bull. Seismol. Soc. Am.* **99**, 2012-2019; DOI: 10.1785/0120080069.

Reasenberg, P., and L. Jones (1994), Earthquake aftershocks: update, *Science*, *265*, 1251-1252.

Reasenberg, P. A., and L. M. Jones (1989), Earthquake hazard after a mainshock in California, *Science*, *243*(4895), 1173-1176.

Richards-Dinger, K., R. S. Stein, & S. Toda (2010). Decay of aftershock density with distance does not indicate triggering by dynamic stress, *Nature*, **467**, 583-586.

Rundle, P. B., J. B. Rundle, K. F. Tiampo, A. Donnellan, and D. L. Turcotte (2006). Virtual California: fault model, frictional parameters, applications, Pure Appl. Geophys. **163**, no. 9, 1819–1846.

Shaw, B. E. (2009). Constant stress drop from small to great earthquakes in magnitude-area scaling, *Bull. Seismol. Soc. Am.*, ***99***, 871.

Shimazaki, K. and T. Nakata (1980). Time-predictable recurrence model for large earthquakes: *Geophys. Res. Lett.* **7**, 279-282.

Ward, S. N. (2000). San Francisco Bay area earthquake simulations: A step toward a standard physical earthquake model, *Bull. Seismol. Soc. Am.* **90**, 370–386.

Wesnousky, S. G. (2008), Displacement and geometrical characteristics of earthquake surface ruptures: Issues and implications for seismic-hazard analysis and the process of earthquake rupture, *Bull. Seismol. Soc. Am.*, ***98***, 1609.

Weldon, R. J., II, G. P. Biasi, C. J.Wills, and T. E. Dawson (2007). Overview of the southern San Andreas fault model; Appendix E in The Uniform California Earthquake Rupture Forecast, version 2 (UCERF 2), *U.S. Geol. Surv. Open-File Rept. 2007-1437-E*, and *California Geol. Surv. Special Rept. 203-E*.

Working Group on California Earthquake Probabilities (1988). Probabilities of large earthquakes occurring in California on the San Andreas fault, *U.S. Geological Survey Open-File Report*, p. 62.

Working Group on California Earthquake Probabilities (1990). Probabilities of large earthquakes in the San Francisco Bay Region, California, *U.S. Geological Survey Circular*, p. 51.

Working Group on California Earthquake Probabilities (1995). Seismic hazards in southern California: Probable earthquakes, 1994-2024, *Bull. Seismol. Soc. Am.* **85**, 379-439.

Working Group on California Earthquake Probabilities (2003). Earthquake probabilities in the San Francisco Bay Region: 2002–2031, *USGS Open-File Report 2003-214*.

Working Group on California Earthquake Probabilities (2007). The Uniform California Earthquake Rupture Forecast, Version 2 (UCERF 2), *USGS Open-File Report 2007-*

1437, and *California Geological Survey Special Report 203*.

Youngs, R. R., W. J. Arabasz, R. E. Anderson, A. R. Ramelli, J. P. Ake, D. B. Slemmons, J. P. McCalpin, D. I. Doser, C. J. Fridrich, F. H. Swan III, A. M. Rogers, J. C. Yount, L. W. Anderson, K. D. Smith, R. L. Bruhn, P. L. K. Knuepfer, R. B. Smith, C. M. dePolo, D. W. O'Leary, K. J. Coppersmith, S. K. Pezzopane, D. P. Schwartz, J. W. Whitney, S. S. Olig, and G. R. Toro (2003). A methodology for probabilistic fault displacement hazard analysis (PFDHA), *Earthquake Spectra* **19**, 191-219.

Appendix – Detailed Task Descriptions

Task F1 - Finalize Fault Models

Task Leader: Tim Dawson (CGS)

A key component of the UCERF3 study is the *Fault Section Database*, part of the *California Reference Geologic Fault Parameter Database*, which specifies the spatial geometry of known active faults and provides the basis for building a fault-based earthquake rupture forecast. This task will develop a revised California fault model and focus on a re-evaluation of the faults included in the UCERF2 fault model, as well as identifying faults from recent studies that should be considered for inclusion in the UCERF3 fault model.

Background: UCERF2 relied heavily on two primary sources for defining the fault geometry used in the fault model. The Community Fault Model (CFM) developed for southern California (Plesch and others, 2007) provided much of the geometry for the major active faults in southern California, while the 2002 National Seismic Hazard Map (NSHM) fault model (Frankel and others, 2002) provided the fault model for the remainder of the State. The UCERF2 fault model also included additional revisions by WGCEP 2007 although, for the most part, the revisions to the fault geometries of the CFM and NSHM were minor. Recent studies either published or in progress, are leading to a better understanding of fault locations, geometries, and rates of deformation throughout California. Integrating these new data into the UCERF3 fault model is the primary objective of this Task, and lead to an improved representation of the known active faults included in the fault model.

This task includes:

- *Integration of new faults and revision of existing faults from recent studies.*
- *Integration of the Statewide Community Fault Model into the UCERF3 fault model.*
- *Development of a Geologically-based block model for geodesy-based deformation models.*
- *Reevaluation of fault endpoints.*
- *Development of a fault zone width parameter.*
- *Integration of the UCERF3 fault parameter database with the National Seismic Hazard Map (NSHM) database.*

Task Status: For UCERF3, Fault Models 3.1 and 3.2 have been developed in coordination with the efforts of the Statewide Community Fault Model (SCFM). The majority of this effort has been focused on updating and augmenting the inventory of faults in northern California. Because of the large number of Quaternary-active faults in the northern California region, a prioritized list of faults was provided to SCFM, so that fault representations could be developed (Figure F1-1). The prioritized list was based on identifying faults that potentially make up block boundaries (for the geodetically-based deformation models), faults that provide

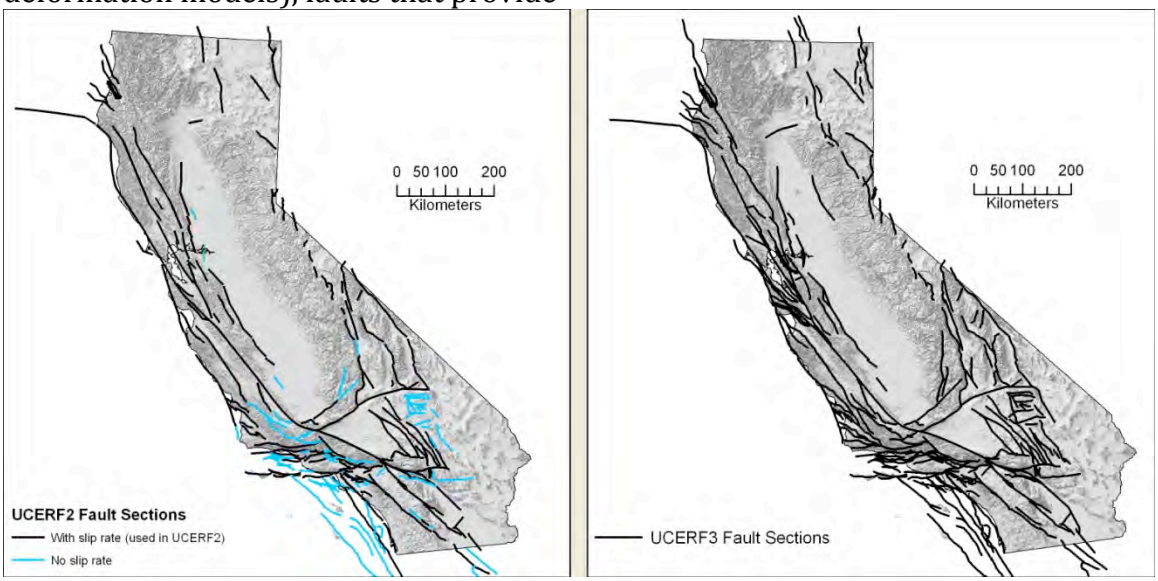


Figure F1- 1. Comparison of Fault Sections in UCERF2 to Fault Sections developed for UCERF3.

potential connections between faults included in the UCERF2/3 model, and faults with new data that warrant inclusion in the UCERF3 model. Faults from the UCERF2 model were also revised by SCFM, using available data and a revised model of depth of seismicity based on the relocated seismicity catalog of Waldhauser (2009). New representations were developed and provided to Tim Dawson for review and integration into the UCERF3 Fault Parameter Database. In general, most SCFM additions and revisions were accepted, with a few exceptions, such as where SCFM representations were less detailed than the existing UCERF2 model. An example of this would be the Great Valley thrust system, which SCFM modeled as two continuous fault zones, in contrast to the 14 fault sections UCERF has in its database. Other exceptions include dips for normal faults in eastern California, where SCFM generally assigned steeper dips than the UCERF2 model assigned. In this case, the UCERF2 default dip value of 50 degrees was retained (maintaining the fault area used in the UCERF2 model). As a result of this exercise, approximately 82 fault sections were added to the UCERF3 fault model and the geometry of 88 fault sections were revised. Table F1 - 1 is a detailed list of which faults were added and revised, along with comments for additional documentation.

In April 2011, two regional workshops, convened by Tim Dawson, Andreas Plesch, and Ray Weldon, were held in Pomona and Menlo Park to present the fault model and solicit

suggestions from the community for improvements to the fault model. Based on presentations and discussions at the southern California workshop, significant issues were brought up regarding fault geometries for Inner Borderlands faults offshore, due to new mapping by the USGS. Based on these presentations and addition discussions with Holly Ryan and Jamie Conrad (USGS), alternative fault models were developed based on this new data (some of it still preliminary). The alternative fault models can be viewed in SCEC-VDO and a description of the modifications is available in Table F1 - 1. Revisions were also made to selected faults in the Transverse Ranges based on a presentation and discussion with Scott Lindvall (WLA-Fugro). These revisions are also noted on Table F1 - 1.

At the northern California workshop, Jeff Unruh (WLA-Fugro) made a presentation regarding segmentation of the Great Valley thrust system, with the main conclusion being that the system of faults is more segmented than the representation UCERF2/3 currently uses. With a few exceptions, very little data is available to help better define the geometry of these faults (at least without embarking on a major research project). An alternative, more segmented, geometry was developed, but ultimately rejected, as there is no data to either support or refute this alternative model. As currently modeled, the majority of Great Valley fault sections are either adjacent, or in very close proximity to another and at this time there does not appear to be a way to rationally make this system more segmented (larger gaps or steps) without additional geological data that would help define the actual locations of these structures.

The development of the UCERF3 Block Model (Figure F1-2) was done in consultation with the Geodetic Modeling group, led by Wayne Thatcher and Kaj Johnson. The block geometry was developed by defining blocks bounded by significant faults and zones of deformation. Block boundaries were defined in a GIS platform and the block boundaries follow the UCERF3 fault sections as much as possible so that block boundary rates can easily be mapped onto UCERF3 fault sections. Because some boundaries follow zones of faulting and other boundaries exist simply to connect blocks, each boundary is attributed in the GIS as one of three categories: 1.) fault, 2.) fault zone, or 3.) connector. These designations will be used to guide the slip rate assignments to the fault sections once the deformation model results are available.

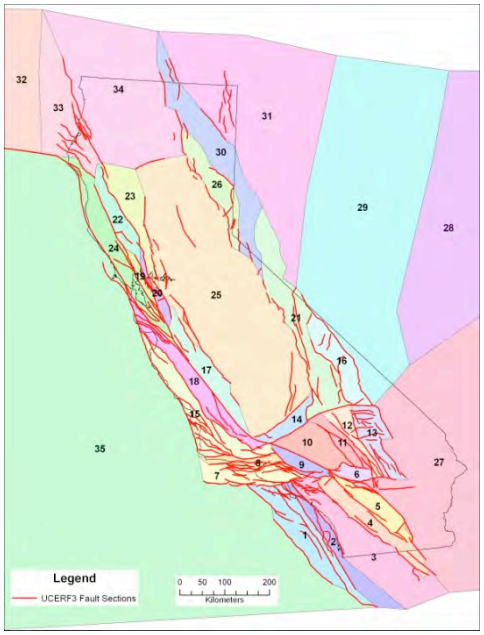


Figure F1- 2. UCERF3 Block Model (colored polygons) and UCERF3 Fault Sections.

Fault endpoints: One key assumption of UCERF3 is that fault endpoints are reasonably well-constrained for purposes of quantifying multi-fault ruptures, thus an effort to examine the fault endpoints for faults in the fault model is a defined task for UCERF3. In general, new fault representations generated by SCFM are more "connected" to other faults than in past models. We have also made an effort to examine other fault endpoints as well and modify them as necessary. Typically, this involves examining the fault model trace and comparing it to the mapped trace as depicted on the inventory of Quaternary-active faults in the USGS Q-faults database and 2010 Fault Activity Map of California, as very little other data is available, or provides significant insight without a case-by-case detailed examination. This effort is essentially complete and, in general, UCERF3 fault sections are more connected and better follow the traces of mapped Quaternary-active faults in the USGS/CGS Quaternary Fault and Fold Database.

Fault Zone Width: A new concept in the UCERF3 Fault Model is a specified fault zone width for each of the fault sections. The motivation behind this is two-fold: 1.) One reason is to specify what faults the deformation model slip rates actually apply to. In the fault model, some fault sections represent discrete individual faults, while others represent fault zones, consisting of multiple individual faults. Once the deformation modelers provide block boundary slip rates, these rates will need to be assigned to the fault sections, and in some cases, may need to be partitioned across a zone of faults; and 2) answering questions like whether the 2010 El Mayor-Cucapah earthquake was an event on the Laguna Salada source (as opposed to being a background seismicity event). The current plan is to assign each fault section a fault zone width, which will be defined by examining the distribution of Quaternary-active faults surrounding each UCERF3 fault section. Because this width will likely vary along strike, it may be necessary to subsection the fault sections to allow for

changes of fault zone width along strike. The fault sections (with the width of the zone defined) will represent a proxy for every seismogenic-thickness and larger event that occurs in that zone with a similar strike, dip and rake (say within 15 degrees of those of the fault section). We will remove these events from the background grid sources in order to avoid both double counting and ambiguity of what future events get associated with. This effort is in its formative stages and a methodology to do this in a rational and efficient way is being developed.

Database integration: Because the results of UCERF3 will be integrated with the National Seismic Hazards Map (NSHM), coordination between the two projects is necessary to ensure integration between the databases used by the individual projects. This was an issue with UCERF2 and led to issues such as certain alternative fault models being inadvertently left out of the NSHM model and difficulty importing the UCERF2 database into the NSHM database. In March 2011, Tim Dawson met with Kathy Haller (USGS) in Golden to discuss these issues. As a result of this discussion, it was decided that the best way to facilitate database integration between the two projects is to develop a single GIS fault parameter geodatabase that both databases can use. This geodatabase is currently being used in parallel with the UCERF3 fault section database, and will eventually be provided to the NSHM once the deformation models are completed and it can be populated with fault slip rates.

Table F1-1. Revisions and additions to UCERF Fault Model.

Fault Section	Addition/Revised	Comments	
Airport Lake	Addition	Added based on FAM trace.	FAM - Fault Activity Map (CGS)
Almanor 2011 CFM	Addition	Addition from CFM	CFM - Community Fault Model
Antelope Valley	Addition	Was in UCERF2, but not in fault section database. Used 2008 parameters.	
Ash Hill	Addition	Addition from FAM.	
Baker	Addition	Added based on FAM traces and as a boundary in block model.	
Bartlett Springs 2011 CFM	Revised	Replaced 2008 geometry with CFM	
Battle Creek 2011 CFM	Revised	Replaced 2008 geometry with CFM	
Bear River	Addition	Added based on FAM mapped trace. Extension of the Garberville-Briceland.	
Bennett Valley 2011 CFM	Addition	Addition from CFM	
Big Lagoon - Bald Mountain	Revised	Revisions based on CFM	
Birch Creek	Removed	Combined with Independence fault based on mis-assigned slip rate and on strike continuity with Independence fault.	
Blackwater	Revised	Trace modified to better match FAM traces.	
Bradford	Addition	Added as southern extension of Quien Sabe fault.	
Breckenridge 2011	Addition	Addition based on new info. Part on Kern Canyon zone.	
Bullion Mountains	Addition	Added based on FAM traces.	
Butano 2011	Addition	Addition from CFM	

Cady	Revised	Trace modified to better match FAM traces.
Calaveras (Central) 2011 CFM	Revised	Replaced 2008 geometry with CFM
Calaveras (No) 2011 CFM	Revised	Replaced 2008 geometry with CFM
Calaveras (So) 2011 CFM	Revised	Replaced 2008 geometry with CFM
Calaveras (So) Paicines extension	Addition	Addition based on CFM, part of Paicines fault. Aseismicity and slip rate assignments may need to be vetted by others.
Camp Rock	Addition/Revised	2008 Landers dissassembled into individual faults
Carson Range (Genoa)	Addition	Was in UCERF2, but not in fault section database. Used 2008 parameters.
Casmalia 2011 CFM	Revised	Replaced 2008 geometry with CFM
Cerro Prieto	Needs revision?	Mike Oskin thinks the fault extends too far north. Checking with Andreas Plesch to see what CFM based the representation on.
Clayton	Addition	Added based on FAM trace, extension of Greenville fault.
Cleghorn	Revised	Modified trace for consistency with North Frontal fault geometry.
Cleghorn Lake	Addition	Added based on FAM traces.
Cleghorn Pass	Addition	Added based on FAM traces.
Collayami 2011 CFM	Revised	Replaced 2008 geometry with CFM
Concord 2011 CFM	Revised	Replaced 2008 geometry with CFM
Contra Costa Shear Zone Briones 2011	Addition	Addition from CFM
Contra Costa Shear Zone connector 2011	Addition	Addition from CFM
Contra Costa Shear Zone Dillon Pt 2011	Addition	Addition from CFM
Contra Costa Shear	Addition	Addition from CFM

Zone Lafayette 2011		
Contra Costa Shear Zone Lake Chabot 2011	Addition	Addition from CFM
Contra Costa Shear Zone Larkey 2011	Addition	Addition from CFM
Contra Costa Shear Zone Ozal Columbus 2011	Addition	Addition from CFM
Contra Costa Shear Zone Reliz 2011	Addition	Addition from CFM
Contra Costa Shear Zone Southhampton 2011	Addition	Addition from CFM
Contra Costa Shear Zone Vallejo 2011	Addition	Addition from CFM
Coronado Bank alt1	Revised/Addition	Revised (truncated) alternative of Coronado Bank fault based on new mapping by USGS (as presented by Ryan at UCERF3 Fault Model Workshop). Because the Palos Verdes fault does not connect, the implication is the slip rate is unknown in this model alternative (per Ryan).
Coyote Canyon	Revised	Trace modified to better match mapped traces (Dave Miller, USGS).
Coyote Lake	Revised	Trace modified to better match mapped traces (Dave Miller, USGS).
Coyote Lake	Revised	Trace modified to better match mapped traces (Dave Miller, USGS).
Davis Creek	Addition	Added based on FAM.
Death Valley	Revised	Trace modified to better match FAM traces.
Death Valley (north)	Revised	Extended to the south to better match mapped trace
Dog Valley	Addition	Added based on FAM, may form connection between Polaris and

		Last Chance faults.
Dry Mountain	Addition	Added based on FAM, major range front fault.
East Huasna 2011 CFM	Addition	Addition from CFM
Eaton Roughs 2011 CFM	Addition	Addition from CFM
Emerson-Copper Mountain	Revised	2008 Landers disassembled into individual faults
Fickle Hill (alt1)	Revised	Now an alternative model for Mad River fault zone
Fish Slough 2011 CFM	Revised	Modified trace from CFM. Note, dip is opposite than 2008 model
Fitzhugh Creek	Addition	Added based on FAM.
Franklin 2011 CFM	Addition	Addition from CFM
Garberville - Briceland 2011 CFM	Addition/Revised	Replaces Garberville portion of Maacama-Garberville in 2008 model. Assigned 2008 parameters for slip rate and aseismicity.
Gillem-Big Crack	Revised	Replaced 2008 geometry with CFM
Goldstone Lake	Revised	Trace modified to better match FAM traces.
Goose Lake 2011 CFM	Addition/Revised	In 2008 model, but not in 2008 database (did NSHM handle this?)
Green Valley 2011 CFM	Revised	Green Valley (no) and (so) combined by CFM
Greenville (no) 2011 CFM	Revised	Replaced 2008 geometry with CFM
Hartley Springs 2011 CFM	Revised	Replaced 2008 geometry with CFM.
Hat Creek 2011 (CFM)	Not used	Keep 2008 geometry for Hat Creek-Mayfield-McArthur combined zone, rather than individual CFM faults.
Hayward (no) 2011 CFM	Revised	Replaced 2008 geometry with CFM
Hayward (so) 2011	Revised	Replaced 2008 geometry with CFM

CFM		
Hayward (so) extension 2011 CFM	Addition	Addition from CFM
Hilton Creek 2011 CFM	Revised	Replaced 2008 geometry with CFM
Holser (alt1)	Revised	Trace modified to better match FAM traces.
Homestead Valley 2011	Revised	2008 Landers disassembled into individual faults
Honey Lake 2011 CFM	Revised	Replaced 2008 geometry with CFM
Hunter Mountain - Saline Valley	Revised	Extended to the south to better match mapped trace
Hunting Creek - Bartlett Springs Connector 2011 CFM	Addition	Addition from CFM
Imperial	Revised	Trace extended south to better match 1941 rupture trace as depicted on Jennings (1994).
Incline Village 2011 CFM	Revised	Replaced 2008 geometry with CFM
Independence 2011	Revised	Combined with Birch Creek fault
Jess Valley	Addition	Added based on FAM.
Johnson Valley 2011	Revised	2008 Landers disassembled into individual faults
Keddie Ridge 2011 CFM	Addition	Addition from CFM
Kern Canyon (Lake Isabella) 2011	Addition	Addition based on new info. Part of Kern Canyon zone.
Kern Canyon (N. Kern) 2011	Addition	Addition based on new info. Part of Kern Canyon zone.
Kern Canyon (S. Kern) 2011	Addition	Addition based on new info. Part of Kern Canyon zone.
King Range 2011 CFM	Addition	Addition from CFM
La Panza 2011 CFM	Addition	Addition from CFM
Landers	Revised	2008 Landers disassembled into individual faults

Last Chance	Addition	Added based on FAM, may form connection between Dog Valley and Honey Lake zones.
Likely 2011 CFM	Revised	Replaced 2008 geometry with CFM
Lions Head 2011 CFM	Revised	Replaced 2008 geometry with CFM
Little Lake	Revised	Trace modified to better match FAM traces.
Los Alamos 2011 CFM	Revised	Replaced 2008 geometry with CFM
Los Alamos connector	Addition	Added as a potential connector fault between the Los Alamos and Santa Ynez faults.
Los Medanos - Roe Island 2011 CFM	Addition	Addition from CFM
Los Osos	Revised	Modified to extend further south along mapped trace.
Los Osos 2011 CFM	Revised	Replaced 2008 geometry with CFM
Los Positas 2011 CFM	Addition	Addition from CFM
Ludlow	Revised	Trace modified to better match FAM traces.
Maacama 2011 CFM	Revised	Replaced 2008 geometry with CFM
Mad River - Trinidad alt2 2011	Addition	Alternative geometry developed to combine individual faults into single seismic source. Note: Depths in CFM are much greater than in 2008 model and may need to be considered
Manix	Revised	Trace modified to better match mapped traces (Dave Miller, USGS).
Mission (connected) 2011 CFM	Addition	Addition from CFM
Mission Hills	Addition	Trace added to connect with Verdugo, based on source characterization developed by WLA-Fugro for Van Norman Dam Complex. Presented at 2011 UCERF3 Fault Model Workshop by Scott Lindvall. Fugro also has slip rates calculated from balanced cross sections

Mohawk Valley 2011 CFM	Addition	Addition from CFM
Mono Lake 2011 CFM	Revised	Replaced 2008 geometry with CFM
Monte Vista-Shannon 2011 CFM	Revised	Replaced 2008 geometry with CFM
Mt Diablo Thrust N (lower)	Addition/Revised	Addition from CFM, reflect upper and lower parts of fault
Mt Diablo Thrust N (upper)	Addition/Revised	Addition from CFM, reflect upper and lower parts of fault
Mt Diablo Thrust S (lower)	Addition/Revised	Addition from CFM, reflect upper and lower parts of fault
Mt Diablo Thrust S (upper)	Addition/Revised	Addition from CFM, reflect upper and lower parts of fault
Naciemento	Removed	Not classified as Quaternary active by the USGS.
Nelson Lake	Revised	Trace modified to better match mapped traces (Dave Miller, USGS).
North Channel (alt2)	Removed	Removed from fault model because both this fault and Pitas Point (upper) in FM2.2 are essentially the same, only North Channel has a more shallow dip. In 3D view, it doesn't look reasonable to have both faults in the same fault model. Also, UCERF2 assigned the North Channel fault a bottom depth of 4.2 km, which seems too shallow. Finally, new work by the USGS suggests structures in the Santa Barbara Channel are more steeply dipping, and support the "thick skinned" model (per Sam Johnson, personal communication), rather than shallow dipping faults.
North Tahoe 2011 CFM	Revised	Replaced 2008 geometry with CFM
Oceanic	Revised	CFM trace modified to better match mapped traces.
Oceanic 2011 CFM	Addition	Addition from CFM

Oceanside Blind Thrust alt1	Revised/Addition	Truncated alternative model developed for the Oceanside thrust based on new mapping by USGS. Truncation based on presentation by Ryan, with the conclusion "Oceanside Blind Thrust does not continue south of La Jolla Canyon and may not be active south of Carlsbad.
Ortigilita	Revised	Trace modified to better match FAM traces.
Owens Valley	Revised	Extended to the south to better match mapped trace
Owens Valley (Keough Hot Springs)	Addition	Added based on FAM, may form connection between Owens Valley and Fish Slough faults.
Ozena	Addition	Added because it connects with S. Cuyama. Boundary in block model.
Palos Verdes	Revised	Trace extended ~8 km to reflect new mapping by USGS (presented by Holly Ryan at UCERF3 Fault Model Workshop)
Paradise	Revised	Trace modified to better match mapped traces (Dave Miller, USGS).
Pilarcitos 2011 CFM	Addition	Addition from CFM
Pinto Mountain	Revised	Trace modified to better match FAM traces.
Pisgah	Revised	Southern-most point modified to reflect mapped trace.
Pittsburg-Kirby Hills lower 2011 CFM	Addition	Addition from CFM
Pittsburg-Kirby Hills upper 2011 CFM	Addition	Addition from CFM
Pittville 2011 CFM	Addition	Addition from CFM
Polaris 2011	Addition	Addition from CFM/CGS
Pt Reyes 2011 CFM	Revised	Replaced 2008 geometry with CFM

Pt Reyes connector	Addition	Added based on FAM traces southeast of Point Reyes fault. Traces appear to splay off of San Andreas.
Quien Sabe 2011 CFM	Revised	Replaced 2008 geometry with CFM
Reliz 2011 CFM	Revised	Replaced 2008 geometry for Rinconada with CFM
Rinconada 2011 CFM	Revised	Replaced 2008 geometry for Rinconada with CFM
Rocky Ledge 2011 CFM	Addition	Addition from CFM
Rodgers Creek - Healdsburg 2011 CFM	Revised	Replaced 2008 geometry with CFM. Trace modified to better match mapped faults on FAM.
Rose Canyon	Revised	Extended to the north to better match mapped trace.
Russ 2011	Addition	Addition from CFM
SAF Creeping 2011 CFM	Revised	Replaced 2008 geometry with CFM
SAF North Coast 2011 CFM	Revised	Replaced 2008 geometry with CFM
SAF Offshore 2011 CFM	Revised	Replaced 2008 geometry with CFM
SAF Peninsula 2011 CFM	Revised	Replaced 2008 geometry with CFM
SAF Santa Cruz Mts 2011 CFM	Revised	Replaced 2008 geometry with CFM. CFM dip was not used due to Loma Prieta seismicity making the CFM representation dip more than the 2008 model.
San Andreas (Banning)	Revised	Trace modified to better match FAM traces.
San Diego Trough alt1	Revised/Addition	Revised fault trace based on new work by Jamie Conrad and Holly Ryan (USGS) that suggests San Diego Trough fault zone and San Pedro basin fault are connected. PRELIMINARY slip rate assigned based on presentation at UCERF3 workshop.
San Gregorio (North) 2011 CFM	Revised	Replaced 2008 geometry with CFM

San Gregorio (South) 2011 CFM	Revised	Replaced 2008 geometry with CFM
San Jacinto (Glen Helen)	Addition	Added to follow FAM mapped traces, provides direct connection to Cucamonga in map view.
San Jacinto (San Bernardino)	Revised	Northern part modified to better match FAM mapped traces.
San Jose (Bay Area) 2011 CFM	Addition	Addition from CFM
San Juan	Revised	Extended to the north to better match mapped trace
San Luis Bay 2011 CFM	Addition	Addition from CFM
San Luis Oceano 2011 CFM	Addition	Addition from CFM
San Luis Pecho 2011 CFM	Addition	Addition from CFM
San Luis Range (south margin)	Revised	Modified to extend further south along mapped trace.
San Luis Range 2011 CFM	Revised	Replaced 2008 geometry with CFM
Santa Cruz Catalina Ridge alt2	Revised/Addition	Alternative geometry to reflect new work by USGS, as presented by Holly Ryan at UCERF3 fault model workshop. No active faults on south margin of Santa Catalina Island.
Santa Susana East connector	Addition	Added to connect with San Fernando, based on source characterization developed by WLA-Fugro for Van Norman Dam Complex. Presented at 2011 UCERF3 Fault Model Workshop by Scott Lindvall
Santa Ynez (West)	Revised	Extended to the west to better match mapped trace
Santa Ynez River	Addition	Added as a potential block boundary.
Sargent 2011 CFM	Addition	Addition using 1996 fault parameters

Sheephole	Addition	Coincident with boundary in block model.
Shoreline 2011	Addition	Addition using 2011 PG&E fault parameters
Sierra Madre (San Fernando)	Revised	Trace modified to better follow location of 1971 surface rupture.
Silver Creek 2011 CFM	Addition	Addition from CFM
Skinner Flat 2011 CFM	Addition	Addition from CFM
South Cuyama (north extension)	Addition	Added from FAM, connects with Rinconada. Boundary in block model.
South Klamath Lake East 2011 CFM	Addition	Addition from CFM
South Klamath Lake West 2011 CFM	Addition	Addition from CFM
Surprise Valley 2011 CFM	Revised	Replaced 2008 geometry with CFM
Swain Ravine - Spencville 2011 CFM	Addition	Addition from CFM
Tank Canyon	Revised	Revision based on FAM and Walker and others (2005).
Tin Mountain	Addition	Added based on FAM, major range front fault.
Tolay 2011	Addition	Addition from CGS
Towne Pass	Addition	Added based on FAM, major range front fault.
Ventura - Pitas Point	Revised	Trace modified to better match FAM traces.
Verdugo	Revised	Trace modified to better match FAM traces.
Walker Spring 2011 CFM	Addition	Addition from CFM
West Napa 2011 CFM	Revised	Replaced 2008 geometry with CFM

Wight Way 2011 CFM	Addition	Addition from CFM
Zayante - Vergelis 2011 CFM (lower)	Revised	Replaced 2008 geometry with CFM
Zayante - Vergelis 2011 CFM (upper)	Revised	Replaced 2008 geometry with CFM

Task D1 – Evaluate New Deformation Models

Task Leaders: Wayne Thatcher (USGS), Kaj Johnson (Indiana University)

In UCERF3, we will be replacing Deformation Models 2.x with versions 3.x. UCERF2 slip-rates were assigned based on an expert-opinion evaluation of available data (mostly geologic and geodetic), together with summations across various transects to make sure the total plate tectonic rate was matched. In UCERF3, fault slip rates and off-fault strain rates estimated from inversions of GPS-derived velocities and geologic slip rates with kinematic models will replace the expert opinion rates (this was the topic of two workshops: April 1-2, 2010, and June 4, 2011). These models remove the need to introduce path-integral transects because the total plate rate is matched explicitly by the models. Deformation Models 3.x will not have C-zones and all faults will be assigned slip rates. Rather than C-zones, Deformation Models 3.x will provide the strain rate tensor on a $0.1^\circ \times 0.1^\circ$ grid covering all of California. This grid of strain rates will account for all modeled deformation that is not accommodated on the faults.

The previous ambiguity with respect to whether slip rates represent deformation on the main fault surface, or whether they apply to a zone surrounding the fault, has now been removed by virtue of adding the *Fault Width* attribute to fault sections (described above); slip rates, by definition, apply over these specified widths. Although the kinematic models treat each fault as a narrow, discrete fault surface, the model-derived fault slip rates will be distributed across broader fault zones using well-established relationships between the width of elastic deformation across a fault zone and the depth of fault locking. From elastic dislocation theory, for a very long strike-slip fault, the width, W, of the shear zone surrounding a fault containing a specified fraction, F, of the total interseismic strain across the fault is $W = L*\tan(F*\pi/2)$.

Kinematic Models

Wayne Thatcher and Kaj Johnson are heading up a GPS sub-group to provide several alternative deformation models derived from a wide range of kinematic models. Seven different kinematic models are being produced by Johnson, Rob McCaffrey, Peter Bird, Bill Hammond, and Yuehua Zeng as illustrated in Figure 1. The kinematic models that we will use for the inversions and the rationale are described below:

1. **DefNode** (Rob McCaffrey, formerly RPI)

Assumptions: Elastic blocks with long-term motions parameterized as rotations around Euler poles. Interseismic elastic distortion at block boundaries introduced with back slip.

Capabilities: Jointly solve for Euler pole rotation rates (slip rates), locking/creeping distribution, and uniform internal block strain rates.

Rationale for use: It is well-established and has been used in published research for at least 10 years.

2. **Zeng buried fault model** (Yuehua Zeng, USGS/Golden)

Assumptions: Interseismic deformation field is generated with imposed slip on buried dislocations that extend from locking depth to infinite depth. Model assumes uniform creep below a specified locking depth. Continuity in slip rates is assumed across segment intersections.

Capabilities: Jointly solve for deep creep rates (long-term fault slip rates) and locking depth.

Rational for use: The model assumptions are quite different from the conventional elastic block model (DefNode). The conventional block model assumes no long-term block distortion (or solved-for uniform strain) while this buried slip model necessarily requires some unspecified amount of elastic distortion of intervening crust to accommodate imposed buried creep. Comparison between this model and conventional block model would be informative.

3. **Johnson elastic quasi-block model** (Kaj Johnson, Indiana University)

Assumptions: Blocks in an elastic plate overlying layered Maxwell viscoelastic half-space (only high viscosity case (i.e., essentially elastic) will be considered for this study to avoid the need to parameterize timing of earthquakes). Long-term block motions parameterized as rigid-body rotations around Euler poles plus non-rigid deformation parameterized with cubic polynomials. Only fault-parallel velocity discontinuities are allowed. Interseismic elastic distortion at block boundaries introduced with steady back slip on faults and periodic forward slip to represent earthquakes. Creeping is assumed to occur at zero stressing rate.

Capabilities: Jointly solve for deep long-term fault slip rates, locking depths, and distributed deformation within blocks.

Rational for use: Internal block strain is parameterized slightly more generally than in DefNode. Incorporates strains due to non-planar fault geometries and dipping faults.

4. **NeoKinema** (Peter Bird, UCLA)

Assumptions: Elastic distortion near faults is removed from velocity field using solutions for dislocations in an elastic half-space. Remaining velocitiy field is representing with a non-parametric fitted surface that allows for discontinuities at fault traces.

Capabilities: Jointly solve for long-term fault slip rates and surface velocity field.

Rational for use: Model formulation is quite different from above models. Probably is the most general formulation of kinematics; does not parameterize surface motions using blocks. Allows for solved-for strain in intervening regions between faults and therefore can be compared with Johnson's solutions.

5. Elastic block model of Bill Hammond (University of Nevada, Reno).

Assumptions: Similar to DefNode. Elastic blocks with long-term motions parameterized as rotations around Euler poles. Interseismic elastic distortion at block boundaries introduced with back slip.

Capabilities: Jointly solve for Euler pole rotation rates (slip rates), locking/creeping distribution, and uniform internal block strain rates.

Rationale for use: Formulation is very similar to DefNode and it will be good for comparison purposes.

All models are constrained by the same GPS velocity field (Figure 2B) provided by Tom Herring, the same geologic slip rate data set (Figure 3), and for the block models, the same block geometry (Figure 2A). Five of the models are classified as 'block models' because they use the block geometry illustrated in Figure 2A. The other two models are classified as 'fault-based' models because the fault elements are not the block boundaries but actual UCERF3 fault traces. A preliminary fault-based model by Peter Bird is illustrated in Figure 5.

The five models across the top of Figure 1, highlighted in green, are all models that use the block geometry illustrated in Figure 2A. Two of these are traditional elastic block models (McCaffrey, Hammond), two are quasi-block models (Zeng, Johnson) in which the assumption of long-term rigid-block motion is relaxed somewhat, and one is not a traditional block model formulation (Bird). Results from the five block models will be averaged in a kinematically consistent way. This averaging is done by using the model slip rate estimates from all five inversions as 'data' in a block model inversion with uniform or spatially variable block strain rates. McCaffrey will compute the uniform strain rate model average and Johnson will compute the spatially variable strain rate average. A preliminary example of average slip rates for the block models is shown in Figure 3, along with the geologic slip rate constraints used in the inversions.

The final step in the block modeling is to compute slip rates on faults that are not on block boundaries (the red traces in Figure 2A). This will be done by integrating block strain rates across the smaller faults. A preliminary example of the slip rates obtained from this integration is shown in Figure 4 for a model computed by Johnson. The majority of the faults display slip rates of 0.5 mm/yr or less, but there are a few areas where the strike slip rates exceed one mm/yr.

The two averaged block models will be candidates for alternative deformation models for UCERF3. The other two candidate deformation models will come from the fault-based models of Zeng and Bird. Candidate deformation models are shown in red boxes in Figure 1.

After the four candidate deformation models are completed, Thatcher and Johnson will work with Ray Weldon and Tim Dawson to check for consistency with geologic information about the

slip rate and style (sense of slip) on the smaller, off-block-boundary faults. This group will also work to assign fault slip rates on discrete model faults to broader shear zones where necessary. The number of candidate deformation models that are adopted for UCERF3 will be unknown until the models are completed and scrutinized.

References:

Bird, P. [2009] Long-term fault slip rates, distributed deformation rates, and forecast of seismicity in the western United States from joint fitting of community geologic, geodetic, and stress direction data sets, *J. Geophys. Res., 114*(B11), B11403, doi:10.1029/2009JB006317

McCaffrey, R. (2005), Block kinematics of the Pacific–North America plate boundary in the southwestern United States
from inversion of GPS, seismological, and geologic data, J. Geophys. Res., 110, B07401, doi:10.1029/2004JB003307.

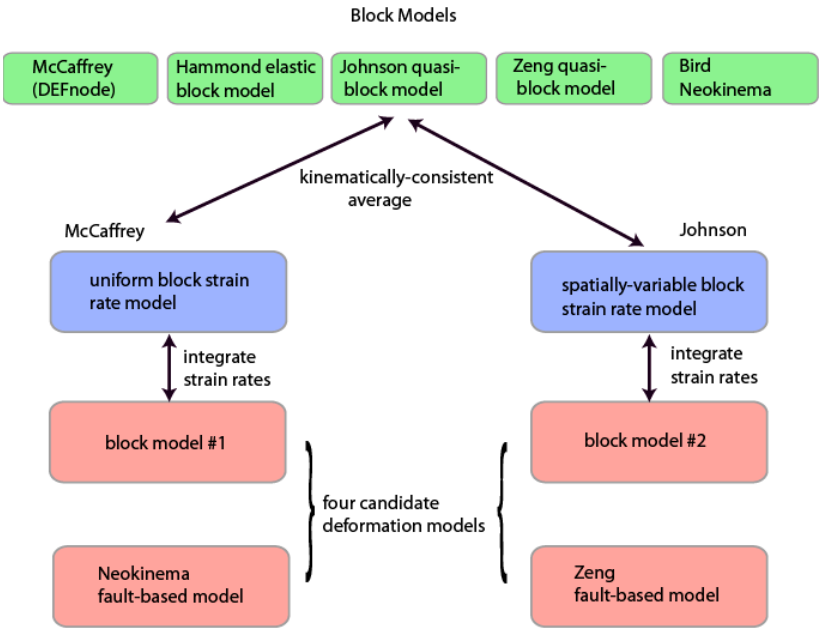


Figure 1. Construction of the four candidate deformation models for UCERF3. Two of the models will be derived from kinematically-consistent averages of five block models. The other two models are fault-based models that do not use the block construction.

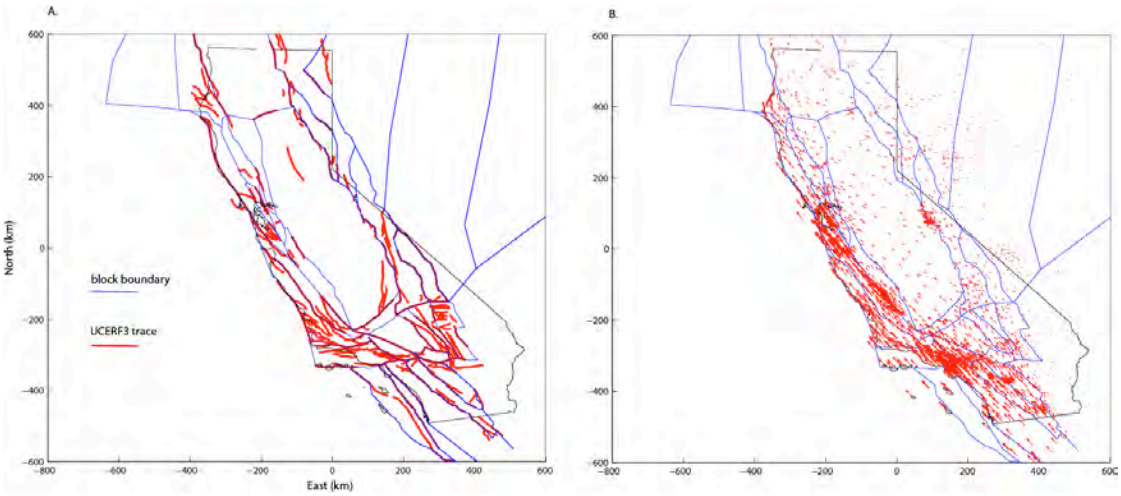


Figure 2. A. Comparison of block model for UCERF3 and UCERF3 fault traces. B. GPS velocity field used in all models (Tom Herring, MIT). Velocities on the Pacific Plate are ~50 mm/yr relative to stable North America.

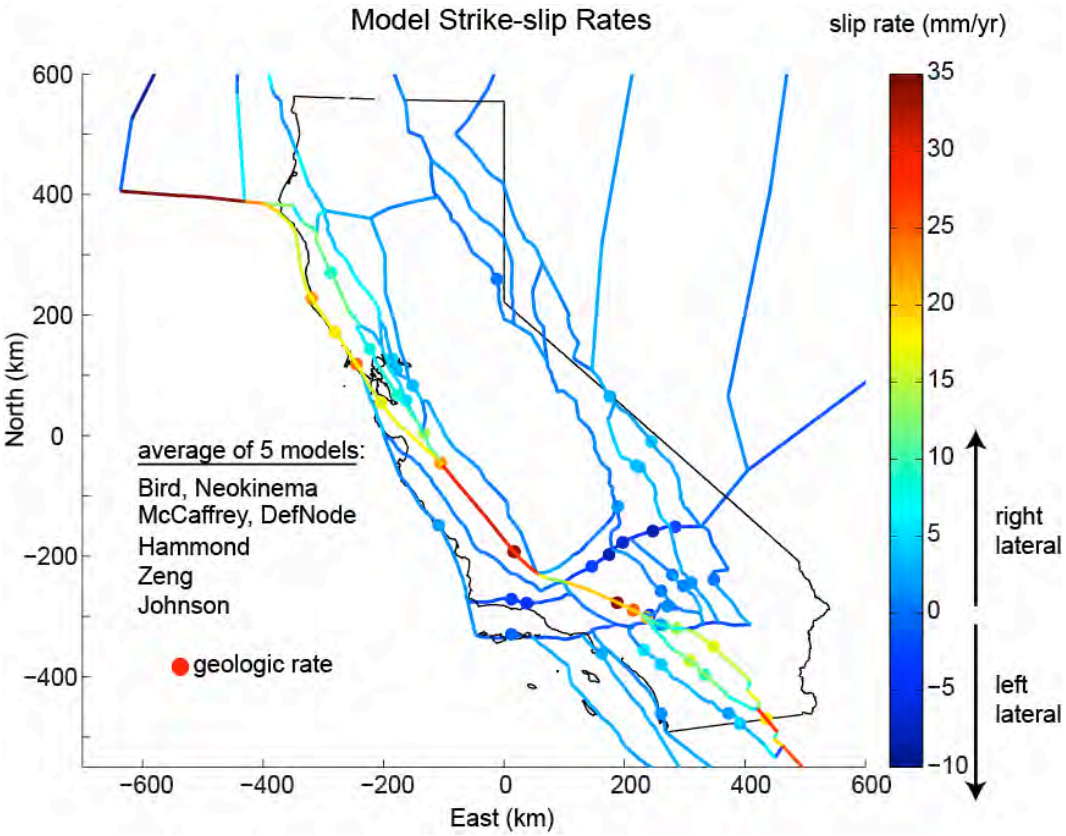


Figure 3. Average slip rate from 5 block models. Slip rate shown for each segment by color. Colored dots show geologic slip rate constraint (mean value) at the same color scale.

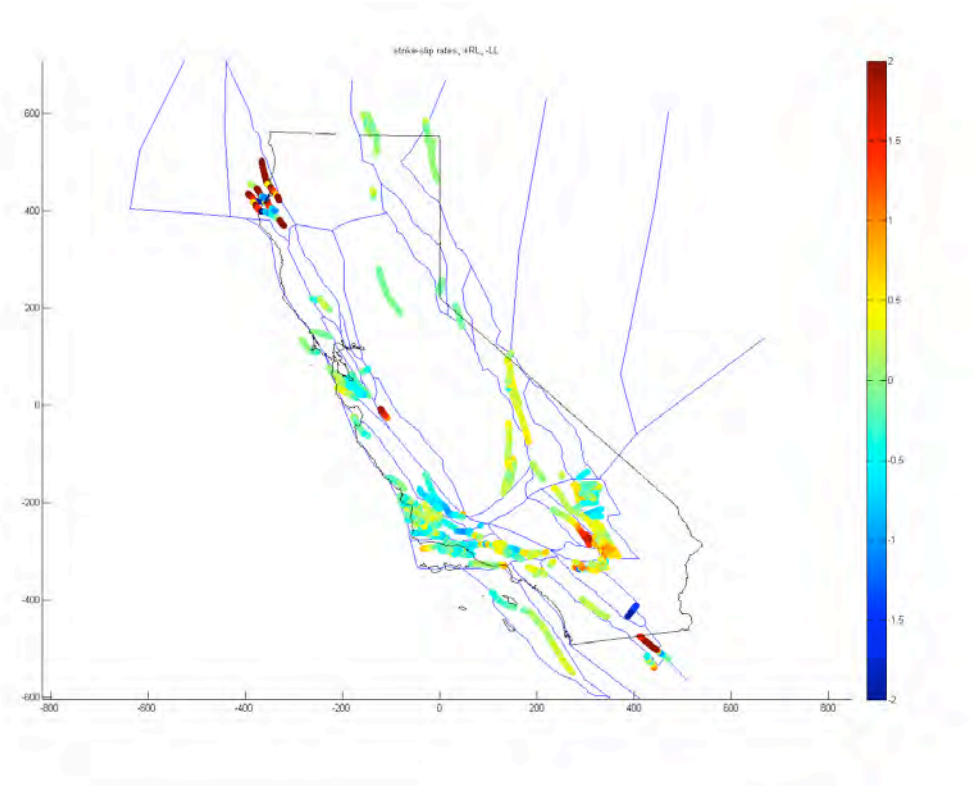


Figure 4. Example of strike-slip rates (mm/yr) computed on off-block-boundary faults by integrating block strain rates across the faults. From Johnson model.

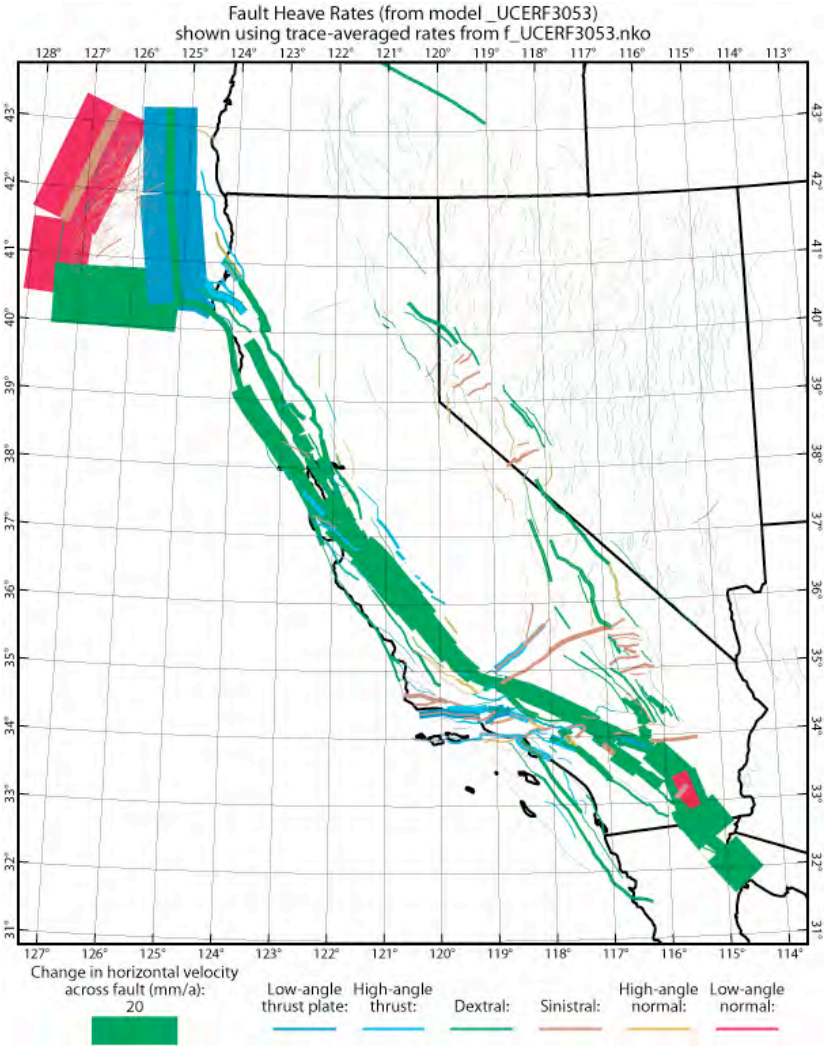


Figure 5. Preliminary Neokinema fault-based model by Peter Bird. Slip rate is indicated by width of bar.

Task D2 – Geologic Slip Rate Compilation

Task Leader: Tim Dawson (CGS)

Objective: Compilation of pure geologic slip rates and associated supporting data to use as additional constraints in the geodetic block and deformation models as well as examine the slip rate data for potential biases. For background information regarding this task, please refer to the main text of this report.

Task Status: Approximately 140 Quaternary slip rates (plus another 17 entries from a compilation by Peter Bird) have been reviewed and provided to the deformation modelers in a table called UCERF3_Geologic_Slip_Rates_version_3_2011_08_03.xls. (Click to download or find it here) The table is a compilation of Quaternary geologic slip rates for faults in the UCERF3 fault model. The table, which is derived from a larger database, is being compiled to provide additional documentation regarding geologic slip rates used in UCERF3, for use in comparing geologic slip rates to geodetic rates, and as possible additional constraints for geodetically-based deformation models. The database is also being compiled to examine potential biases in geologic slip rates used by past WGCEPs.

The table provided to the modelers is not intended to be a comprehensive table of all reported geologic slip rates. It is intended to provide representative, Quaternary slip rates at locations along faults on or near the UCERF3 block model boundaries (Figure D2-1). In general, slip rates were compiled from the literature using the following criteria: 1.) Quaternary rates of < 1 m.y., with Late Quaternary rates preferred 2.) Rates with reported dating constraints and/or feature offset information 3.) From sites that could be reliably located so that longitude and latitude information can be reported.

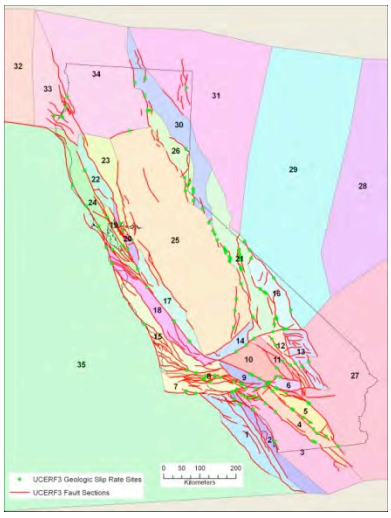


Figure D2- 1. UCERF3 block boundaries and geologic slip rate sites.

This table does not include slip rates that rely on assumptions of characteristic slip, are heavily model dependant (such as using assumptions of horizontal to vertical slip to derive horizontal slip rates from amounts of vertical offset), or slip rates that are in need of revision due to revised dating at a site (e.g. Salyards and others 1992, Pallett Creek slip rate). Rates that are somewhat suspect, because they may be derived from features offset by a limited number of earthquakes (1 event, or slip rates with large open intervals) that may not represent a longer-term average are also excluded, or noted in the comments section as such. Multiple slip rates (either for the same time frame, or over multiple intervals of time) at a single or nearby sites are not reported in this table. Instead, either a single representative slip rate or averaged slip rate is reported. The full database (which is not yet fully populated) will consist of these other slip rate categories, plus additional rates that are currently missing the necessary supporting data, or need additional review. The full database will be used to examine the dataset for potential biases in slip rates and evaluate the quality of the slip rates based on quality of the offset data and dating constraints as described in the original Task description.

The table also includes selected entries excepted from an extensive database of slip rates (including long-term rates) Peter Bird (UCLA) has compiled. Entries attributed to him were selected based on the criteria described above.

An associated analysis is also in progress to assign the "Grey faults" (Quaternary faults in the model with unknown slip rates) with slip rates. This will involve assigning the general USGS slip rate categories given to all faults in the Quaternary Fault and Fold Database to the Grey faults. An initial analysis (presented at the Northern California Fault Model Workshop) showed relatively good agreement between the generalized slip rate categories and faults with assigned geologic slip rates in the fault model. If this relation holds with additional analysis, then it seems reasonable to assign the generalized slip rates categories to faults without slip rates. The database of pure geologic slip rates will be used to support this effort in associating geologic slip rates to the USGS-assigned slip rate categories.

A description of the columns in the table is provided below.

Description of Table Columns:

FAULT_NAME	Name of UCERF3 fault or fault section

BLOCK BOUNDARY RELATION	For convenience, each reported slip rate is associated with the nearest block boundary in the UCERF3 Block Model, or is noted as not associated with a block boundary. Note that the block boundaries are highly simplified from the fault model so relating a slip rate site to a block boundary is not necessarily straight forward.
SITE NAME	Name of the slip rate site, if given, or simply related to the source reference.
Longitude	Longitude of the slip rate site, given in decimal degrees.
Latitude	Latitude of the slip rate site, given in decimal degrees.
MEASURED COMPONENT OF SLIP	Measured component of slip is expressed either as vertical separation (V), or horizontal separation (Hz). Some sites will have this reported separately and will need to be combined if one wants to calculate net slip rate, or extension rates.
FAULT SENSE OF MOTION	Sense of motion on the fault, as given in the UCERF3 fault model. N normal, RL, right lateral; LL, left lateral; R; reverse or thrust.
PREF SLIP RATE	Either the reported preferred value, or the middle of the reported range.
REPORTED RATE	The slip rate as reported by the investigators from the source material. Note that only horizontal or vertical slip rates are reported in this table. Net slip rates need to be calculated either using the fault parameters assigned in the UCERF3 fault parameter database, or using the preferred parameters of the original investigators.
MAX SLIP RATE	Maximum slip rate, if reported, or calculated from source material.
MIN SLIP RATE	Minimum slip rate, if reported, or calculated from source material.
GENERAL COMMENTS	Descriptive information intended to provide background information, and context of the reported slip rate. This column may provide information about other reported rates for a given fault section, and why these rates are not included in the table.

SHORT CITATION	Short citation of source data. Full citations have not yet been compiled, please contact Tim Dawson if you need a full reference.
SITE TYPE	1: trench data, 2: geologic offset, from surface data, or cross sections.
COMPILER	Denotes compiler of data (either Tim Dawson, or Peter Bird)

Based on additional meetings with the group charged with developing a geodesy-based deformation model, a table that specifies block boundary faulting styles and assigns general slip rate categories was developed and provided to the deformation modelers. The table is called UCERF3_Block_Boundary_assignments_2011_09_version_3.xls and can be found at http://www.wgcep.org/data

Faulting style is assigned based on the USGS Quaternary Fault and Fold Database (QFFD), or where no faulting style is specified, assigned based on attributes such as fault orientation, regional patterns of deformation, and similarities to other faults in the area. Slip rate categories are based on the QFFD assignments. However, in some cases, the categories are more finely divided where additional information is available. The purpose of this table is to provide additional constraints on faulting styles and slip rates for the UCERF3 block models. The table may also be used as a basis to check geologic rates against the results of the block modeling exercise.

Task D3 – Line Integral Tools

Task Leader: Tom Parsons (USGS)

This task is intended to offer testing opportunities for the fault and deformation models. We anticipate using the same line integral and strain tensor tools as applied for UCERF-2, and an additional test using a more independent comparison of vertical deformation between observed values and those calculated from the deformation model. Brief summaries are given below of already-applied methods as well as the uplift-subsidence testing.

Line Integral Analysis

The long-term rate model can be tested against the known plate boundary rate. We use the method of Humphreys and Weldon (1994) to accumulate uncertainty along line-integral paths across the plate boundary, and use several input values, including uncertainties in the rake and orientation of the faults, deformation between stable North America and California, and block rotations, from Humphreys and Weldon (1994) where the model does not contain the required data. Fault slip rates are taken from deformation model. Strong differences between the model and plate boundary sum are not expected because past Working Group models, upon which this one will be built, have been "tuned" to match the known plate rate, by choosing "preferred" values from a broad range of uncertain slip rates that approximately add up to the plate rate.

Line integrals are very sensitive to the path chosen. One could test possible differences between closely spaced paths, by a Monte Carlo sampling approach, like that used by Humphreys and Weldon (1994) to determine cumulative uncertainty in each path. This was not done by WG-07 because it was clear from qualitative examination of the data that only Transverse Ranges paths would change by more than a few millimeters per year. In addition, line integral paths that cross rotating blocks must correctly account for rotations that are not explicitly included in our deformation model. We have used the rotations determined by Humphreys and Weldon (1994), but it is unlikely, particularly in southern California, that all of the rotations are known and well characterized. The addition of GPS-driven modeling in UCERF-3 may resolve this problem. Appendix P of the WG-07 report has a full description of the line integral approach taken.

Strain Tensor Analysis

To test deformation and seismic source models, strain tensors across the Pacific - North American plate boundary can be constructed and compared to predictions from the far field plate motion. For WG-07, we used the Kostrov (1974) method as presented in Aki and Richards (1984). Molnar (1983; 1979; et al., 2007; Chen and Molnar, 1977) and many others have discussed the relative merits of using symmetrical strain tensors (as we did) versus asymmetrical tensors or a combination of rotational and irrotational components of the deformation field. We finessed this issue to some extent by comparing principal strain axes from our symmetrical strain tensors to those resulting from a single ideally-oriented (plate boundary parallel) fault, with the plate rate of slip, embedded in the same volume as the distributed deformation we consider. The fact that the distributed deformation almost

exactly equaled the strain inferred from the Pacific - North America plate motion in both rate and style suggested that symmetrical tensors adequately captured the deformation. For WG-07, we analyzed ten 3D volumes spanning our model, oriented perpendicular to the plate boundary

For the entire region, the WG-07 deformation model accounted for ~95% of the plate motion. This is almost certainly within the calculation uncertainty, which includes the slip rates on the faults, the rate of background seismicity and aftershocks, the depths of the faults and the thickness of the block being deformed. For the entire region, the WG-07 seismic source model accounted for ~70% (64.6% plus an estimated 5% aftershocks that are not included in the model) of the plate motion. This is very consistent with the global average seismic component of strike slip plate boundaries (Bird and Kagan, 2004). Appendix P of the WG-07 report has a full description of the strain tensor approach taken.

Vertical Strain Analysis

The reasoning behind examining vertical strain implied by the UCERF deformation models (Fig. D6f2) is that these measures are largely independent of the data used in model construction. This is in contrast with testing against the horizontal plate boundary rates because the deformation model tends to be assembled with the plate boundary budget in mind; thus the extent that they match is not necessarily an independent test. Significant mismatches between observed uplift and subsidence, or recent topography could indicate problems with the fault and/or deformation models, and potentially help with distinguishing between, or weighting of competing models.

Progress to date on new tools includes completed programming that converts UCERF fault and deformation model database formats into 3D elastic dislocations (Fig D6f1) that can be slipped according to their estimated slip rates and rakes. Thus the implications on long-term deformation can be examined and compared with observables such as overall plate boundary displacement rates. We have begun to assemble a considerable uplift and subsidence rate database (see reference section). A very complete database for southern California already exists (Niemi et al., 2008), and there are a number of published estimates for northern and Central California as well.

Another potential use of the dislocation model tools is in mapping stress concentrations that result from slipping the faults. If very large stress concentrations result from slipping the model faults, this would be an indication that the deformation model may be incomplete in terms of absorbing plate boundary stress.

Figure D6f1. 3D dislocation model of California faults as of UCERF2.

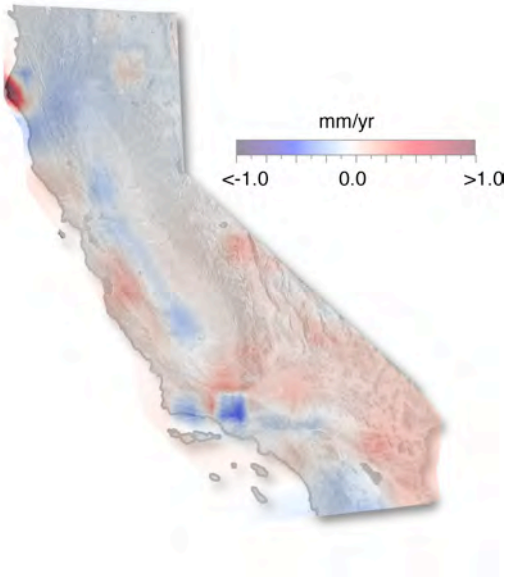


Figure D6f2. Calculated uplift and subsidence resulting from slipping faults in the UCERF2 deformation model.

References cited and assembly of vertical change database:

Aki, K., and Richards, P.G. (1980), Quantitative seismology; Theory and Methods, v. 1, Freeman and Co., New York, 557 p.

Barbour, M. G., T. Keeler-Wolf, and A. A. Schoenherr (2007) Terretrial vegitation of California, University of California Press.

Bawden, G. W., W. Thatcher, R. S. Stein, K. W. Hudnut, and G. Peltzer (2001), Tectonic contraction across Los Angeles after removal of groundwater pumping effects, Nature 412, 812-815 (23 August 2001), doi:10.1038/35090558

Bird, P., and Kagan, Y.Y. (2004), Plate-tectonic analysis of shallow seismicity: Apparent boundary width, beta, corner magnitude, coupled lithosphere thickness, and coupling in seven tectonic settings, Bull. Seis. Soc. Amer., v. 94, no. 6, 2380-2399, doi:10.1785/0120030107.

Bowersox, J. R. (2004), Late Neogene Paleobathymetry, Relative Sea Level, and Basin-Margin Subsidence, Northwest San Joaquin Basin, California, Search and Discovery Article #30029, Adapted from extended abstract for presentation at AAPG Annual Convention, Dallas Texas, April 18-21, 2004.

Burgmann, R., G. Hilley, A. Ferretti, and A. Novali (2006), Resolving vertical tectonics in the San Francisco Bay Area from permanent scatterer InSAR and GPS analysis Geology, v. 34; no. 3; p. 221–224; doi: 10.1130/G22064.1.

Carmichael, I.S.E., Lange, R.A., Hall, C.M., Renne, P.R. (2006), Faulted and tilted Pliocene olivine-tholeiite lavas near Alturas, NE California, and their bearing on the uplift of the Warner Range Bulletin of the Geological Society of America 118 (9-10), pp. 1196-1211.

Chaytor, J.D., Goldfinger, C., Meiner, M.A., Huftile, G.J., Romsos, C.G., and Legg, M.R. (2008), Measuring vertical tectonic motion at the intersection of the Santa Cruz-Catalina Ridge and Northern Channel Islands platform, California Continental Borderland, using submerged paleoshorelines, Bulletin of the Geological Society of America 120 (7-8), pp. 1053-1071.

Chen, W.-P., and Molnar, P. (1977), Seismic moments of major earthquakes and the average rate of slip in Central Asia, Jour. Geophys. Res., v. 82, no. 20, 2945-2969.

Cichanski, M. (2000) Low-angle, range-flank faults in the Panamint, Inyo, and Slate ranges, California: Implications for recent tectonics of the Death Valley region Bulletin of the Geological Society of America 112 (6), pp. 871-883.

Covault, J., Normark, W., Romans, B., & Graham, S. (2007). Highstand fans in the California borderland: The overlooked deep-water depositional systems Geology, 35 (9) DOI: 10.1130/G23800A.1

Dzurisin, D., Poland, M.P., Bürgmann, R. (2002) Steady subsidence of Medicine Lake volcano, northern California, revealed by repeated leveling surveys Journal of Geophysical Research B: Solid Earth 107 (12), pp. ECV 8-1 - 8-16.

Gillespie, A. R., Quaternary subsidence of Owens Valley, California, Geomorphology/Paleoclimatology, 356-382.

Grant, L. B., L. J. Ballenger and E. E. Runnerstrom (2002), Coastal Uplift of the San Joaquin Hills, Southern Los Angeles Basin, California, by a Large Earthquake since A.D. 1635, Bulletin of the Seismological Society of America; March 2002; v. 92; no. 2; p. 590-599; DOI: 10.1785/0120010119

Grove, K., and Niemi, T.M. (2005) Late Quaternary deformation and slip rates in the northern San Andreas fault zone at Olema Valley, Marin County, California Tectonophysics 401 (3-4), pp. 231-250.

Hogarth. L. J., J. Babcock, N. W. Driscoll, N. Le Dantec, J. K. Haas, D. L. Inman, and P. M. Masters (2007), Long-term tectonic control on Holocene shelf sedimentation offshore La Jolla, California, Geology, 35; no. 3; p. 275–278; doi: 10.1130/G23234A.

Humphreys, E. D., and Weldon, R. J. (1994), Deformation across the Western United States: A local determination of Pacific-North America transform deformation, J. Geophys. Res., v. 99, 19,975-20,010.

Keller, E.A., M. Duffy, J.P. Kennett, and T. Hill (2007), Tectonic geomorphology and hydrocarbon induced topography of the Mid-Channel Anticline, Santa Barbara Basin, California, Geomorphology, 89, 274–286.

Knott, J. R. and D. S. Eley (2006), Early to middle holocene coastal dune and estuarine deposition, Santa Maria Valley, California, Physical Geography, 2006, 27, 2, pp. 127–136.

Knott, J.R., Sarna-Wojcicki, A.M., Machette, M.N., Klinger, R.E. (2005), Upper Neogene stratigraphy and tectonics of Death Valley - A review Earth-Science Reviews 73 (1-4), pp. 245-270.

Kostrov, B. V. (1974) Seismic moment and energy of earthquakes, and seismic flow of rock, Izv. Acad. Sci. USSR Phys. Solid Earth I., 23-44.

Leon, L.A., Christofferson, S.A., Dolan, J.F., Shaw, J.H., and Pratt, T.L. (2007), Earthquake-by-earthquake fold growth above the Puente Hills blind thrust fault, Los Angeles, California: Implications for fold kinematics and seismic hazard Journal of Geophysical Research B: Solid Earth 112 (3), art. no. B03S03.

Lerch, D.W., Miller, E., McWilliams, M., Colgan, J. (2008), Tectonic and magmatic evolution of the northwestern Basin and Range and its transition to unextended volcanic plateaus: Black Rock Range, Nevada Bulletin of the Geological Society of America 120 (3-4), pp. 300-311.

McCrory, P.A. (2000), Upper plate contraction north of the migrating Mendocino triple junction northern California: Implications for partitioning of strain Tectonics 19 (6), pp. 1144-1160.

McLaughlin, R.J., and Clark, J.C. (2004), Stratigraphy and structure across the San Andreas fault zone in the Loma Prieta region and deformation during the earthquake US Geological Survey Professional Paper (1550 E), pp. E5-E47.

Meigs, A., Yule, D., Blythe, A., and Burbank, D. (2003) Implications of distributed crustal deformation for exhumation in a portion of a transpressional plate boundary, Western Transverse Ranges, Southern California Quaternary International 101-102 (1), pp. 169-177.

Merrits, D. and W. B. Bull (1989), Interpreting Quaternary uplift rates at the Mendocino triple junction, northern California, from uplifted marine terraces, Geology, 17, 1020-1024.

Merritts, D.J. (1996), The Mendocino triple junction: Active faults, episodic coastal emergence, and rapid uplift Journal of Geophysical Research B: Solid Earth 101 (3), pp. 6051-6070.

Molnar, P. (1983), Average regional strain due to slip on numerous faults of different orientations, Jour. Geophys. Res., v. 88, no. B8, 6430-6432.

Molnar, P., Anderson, R. S., and Anderson, S. P. (2007), Tectonics, fracturing of rock, and erosion, Jour. Geophys. Res., v. 112, F03014, doi:10.1029/2005JF000433.

Molnar, P. (1979), Earthquake recurrence intervals and plate tectonics, Bull. Seis. Soc. Amer., v. 69, no. 1, 115-133.

Montgomery, D. R. (1993), Compressional uplift in the Central California Coastal Ranges, Geology, 21, 543-546.

Muhs, D.R., Wehmiller, J.F., Simmons, K.R., and York, L.L. (2003) Quaternary sea-level history of the United States, Developments in Quaternary Science 1 (C), pp. 147-183.

Niemi, N.A., Oskin, M., and Rockwell, T.K. (2008), Southern California earthquake center geologic vertical database motion, Geochemistry, Geophysics, Geosystems 9 (7), art. no. Q07010

Orange, D. L. (1999), Tectonics, sedimentation, and erosion in northern California: submarine geomorphology and sediment preservation potential as a result of three competing processes, Marine Geology, Volume 154, Issues 1-4, 2 February 1999, Pages 369-382.

Orme, A. R., (1998), Late Quaternary tectonism along the Pacific coast of the Californias: a contrast in style, Geological Society, London, Special Publications; 1998; v. 146; p. 179-197; DOI: 10.1144/GSL.SP.1999.146.01.10.

Orme, A.R., and Orme, A.J. (2008), Late Pleistocene shorelines of Owens Lake, California, and their hydroclimatic and tectonic implications, Special Paper of the Geological Society of America (439), pp. 207-225.

Peterson, C.D., Jol, H.M., Vanderburgh, S., Phipps, J.B., Percy, D., Gelfenbaum, G. (2010), Dating of late Holocene beach shoreline positions by regional correlation of coseismic retreat events in the Columbia River littoral cell, USA, Marine Geology 273 (1-4), pp. 44-61.

Pinter, N., C. C. Sorlien, and A. T. Scott (2003), Fault-related fold growth and isostatic subsidence, California Channel Islands, American Journal of Science, Vol. 303, April 2003, P.300-318; doi:10.2475/ajs.303.4.300

Pirazzoli, P. A. (1991) World Atlas of Holocene sea level changes, Elsevier, 300 pp.

Platt, J.P., Kaus, B., and Becker, T.W. (2008), The mechanics of continental transforms: An alternative approach with applications to the San Andreas system and the tectonics of California Earth and Planetary Science Letters 274 (3-4), pp. 380-391.

Punke, M. L. (2005), Paleoenvironmental Reconstruction of an Active Margin Coast from the Pleistocene to the Present: Examples from Southwestern Oregon, Ph.D. Thesis, Oregon State.

Ryan, H.F., Parsons, T., and Sliter, R.W.(2008), Vertical tectonic deformation associated with the San Andreas fault zone offshore of San Francisco, California, Tectonophysics 457 (3-4), pp. 209-223.

Saleeby, J., Saleeby, Z., Nadin, E., and Maheo, G. (2009), Step-over in the structure controlling the regional west tilt of the Sierra Nevada microplate: Eastern escarpment system to Kern Canyon system, International Geology Review 51 (7-8), pp. 634-669.

Stein, R.S., W. Thatcher, and R.O. Castle (1979), Initiation and development of the southern California uplift along its northern margin, Tectonophysics, 52, Issues 1-4, 10 February 1979, Pages 301-302.

Stokes, M., and García, A.F. (2009), Late Quaternary landscape development along the Rancho Marino coastal range front (southcentral Pacific Coast Ranges, California, USA), Journal of Quaternary Science 24 (7), pp. 728-746.

Sylvester, A.G. (1995), Nearfield vertical displacement in the creeping segment of the San Andreas fault, central California, 1975 to 1994, Tectonophysics 247 (1-4), pp. 25-47.

Talling, P.J., Sowter, M.J. (1999), Drainage density on progressively tilted surfaces with different gradients, Wheeler Ridge, California Earth Surface Processes and Landforms 24 (9), pp. 809-824.

Trecker, M. A., L. D. Gurrola, and E. A Keller (1998), Oxygen isotope correlation of marine terraces and uplift of the Mesa Hills, Santa Barbara, California, USA, in, Stewart, I. S., and Vita-Finzi, C. (eds.), Coastal Tectonixcs, Geological Society of London, Special Publications, 146, 57-69.

Trexler Jr., J.H., Park, H.M., Cashman, P.H., Mass, K.B. (2009), Late Neogene basin history at honey lake, northeastern California: Implications for regional tectonics at 3 to 4 Ma Special Paper of the Geological Society of America (447), pp. 83-100.

Wakabayashi, J., Sawyer, T.L. (2001), Stream incision, tectonics, uplift, and evolution of topography of the Sierra Nevada, California Journal of Geology 109 (5), pp. 539-562.

Whipple, K. X., and C. R. Trayler (1996), Tectonic control of fan size: the importance of spatially variable subsidence rates, Basin Research, 8, 351–366.

Wisely, B.A., and Schmidt, D. (2010) Deciphering vertical deformation and poroelastic parameters in a tectonically active fault-bound aquifer using InSAR and well level data, San Bernardino basin, California Geophysical Journal International, 181 (3), pp. 1185-1200.

Yeats, R.S. (2004), Tectonics of the San Gabriel Basin and surroundings, southern California, Bulletin of the Geological Society of America 116 (9-10), pp. 1158-1182.

Task R1 - Evaluate Along-Strike Distribution of Average Slip

Task Leader: Glenn Biasi (UNR)

Introduction

The focus of this task is to supply D_{sr} (displacement D on section s of rupture r) estimates for use in the grand inversion. The inversion develops a list of all possible ground ruptures based on mapped fault locations, lengths, and distances to other faults that could support fault-to-fault jumps. Faults are discretized into panels ~7 km wide that extend in depth nominally to the base of the seismogenic crust. Faults may join in mid-panel, as might occur when a normal fault ends at a continuing strike-slip fault. Ruptures and details of fault intersections are available through the SCEC VDO (http://www.wgcep.org/tools-scec_vdo).

Background and Overview of Strategy

The shape of the displacement profile.

The displacement profile is a fundamental aspect of assigning *Dsr*. The simplest displacement profile is a boxcar shape for which the displacement is the average everywhere. Among more realistic shapes , ruptures have been found on average (Biasi and Weldon, 2006) to follow an empirically deduced *sqrt(sin(x/2L))* (a.k.a., *sinesqrt,* or informally, the "rainbow") shape after normalizing ruptures by length and average displacement. Individual ruptures are more variable, but as an average shape it is fits remarkably well. We find, using the larger dataset of Wesnousky (2008) that it also characterizes subsets of the available data, including the shortest (<30 km) and longest (>200 km) ruptures.

Additional ruptures to update the rupture profile were identified at a workshop in March 2-3, 2011 in Menlo Park. A review was also recommended of the parametric values for events in the Wesnousky (2008) compilation. New and reviewed events will also contribute to task R2, especially whether or not average displacement (AD) saturates with increasing rupture length (L) as length exceeds the thickness of the seismogenic crust. Ground-rupturing earthquakes most likely to contribute to the AD-L discussion included several large and great historical Chinese and Mongolian earthquakes for which new data have been developed. The functional shape of the AD vs. L curve and how it flattens with increasing length will be developed under Task R2.

We examined the origin of the *sinesqrt* shape itself. Rupture in an ideal solid should yield an elliptical rupture profile. Studies indicate that friction near the crack ends modifies the elliptical shape and flattens the slip gradient. The result is a flatter taper of rupture ends than the ellipse that is empirically similar to the *sinesqrt*.

Applying the displacement profile to multi-segment and multi-fault ruptures

Displacements on individual faults of the Landers rupture were unusually large for their individual mapped fault lengths and standard regressions of length versus displacement. On the other hand, the average displacement for the rupture as a whole is roughly consistent with expectations. Using ruptures in the Wesnousky (2008) database, we developed evidence showing that the disproportionate displacements seen in the Landers rupture are a common feature of multi-fault ruptures (Figure 1). When faults or segments combine into larger ruptures, their displacements increase out of proportion to segment length but similar to displacements expected for the net rupture. It seems that the individual components generally "know" in advance or respond in the event to the larger scales of the combined event. A similar result is observed for the small set of available non-strike-slip events (Figure 2). The mechanism(s) for this scaling change may be dynamic, and in any case involve stress accumulation on scales similar or larger than the combined earthquake rupture. With respect to assigning *Dsr* in ruptures, the full rupture length is thus the more relevant dimension than would be the individual segment lengths linking to make the rupture.

Figure 3 was developed to explore the scale of overlaps and secondary structures necessary to provide continuity of displacement in a generally linear fault system. The model approximates two sub-parallel strands of a strike-slip system. Abutting segments of a linked rupture with *sinesqrt* displacement shapes would have zero displacement where they join. This raises problems trying to match geodetic slip rates along a complete fault system. However, with 5% overlap of the component ruptures, about 60% of the rupture average displacement is recovered in the middle section. This amounts to only ~5 km of overlap in a 100 km long rupture. With an overlap of less than 10%, the sum of the two displacements would equal the mean for the entire rupture. Thus small overlaps in sub-parallel fault ends can accommodate apparently discontinuous slip and leave very little slip deficit to be filled with smaller earthquakes.

The way that AD is estimated from observed rupture profiles may contribute to the discrepancy between seismic estimates of magnitude and estimates from length and displacement. Rupture displacements at the surface can exhibit short-wavelength spatial variability that can translate into a biased and smaller AD estimate. Ruptures are likely to be less variable and have fewer lows at depth as they are at the surface. To explore this question, we developed "enveloped" average displacement estimates for available ruptures. We found that for a few generally smaller events enveloped average displacements are larger by 30% or more, but the effect is smaller for the larger events. For example, the difference averages about 10% for average displacements of 2-3 m, and less than ~5% for events larger than 3 meters. Thus enveloped displacement profiles do not appear to be the primary cause for a discrepancy between geologic and seismic moment estimates.

Implementation Plan

Application of the displacement profile to multi-segment and multi-fault ruptures

-- Five broad cases exist for fault-to-fault ruptures:

SS-SS

SS-Normal

SS-Reverse

Normal-Normal

Reverse-Reverse

-- D(L) for the full rupture should scale with full rupture length. Uncertainty in D(L) may be taken empirically from the range of $D_{observed}$ around the D(L) vs. L regression. The COV of D(L) may be taken from Hecker et al. (2011) if the empirical range around the regression for D(L) is too large.

-- Communication of stress between faults of a rupture means that some continuity of stress direction can be assumed. That is, the slip motions in a multi-fault rupture will be solving a coherent problem of stress relief.

-- The available rupture data indicate that multi-fault rupture displacements consist of multiple rainbows. However, when they link, a gross continuity of displacement in the net rupture profile is commonly observed. Continuity of displacement is maintained through overlap of the linking faults and through a system of secondary linking structures generally too small and too poorly known to model separately.

-- For the *sinesqrt* shape to be used, fault ends require consistent slip-rate gradients to avoid an artificial slip deficit or excessive rates of smaller earthquakes.

-- The section size used to discretize ruptures limits the degree to which tapered ruptures can be implemented.

-- Fault extensions and changes in geometry with time are not accommodated. In theory this means that the stress state of the affected faults may become inconsistent.

-- Some fault-to-fault cases are under-determined and cannot be resolved without other information or an ad hoc choice. A simple example is a normal fault that T's into a strike slip fault that continues past the T. How much SS motion is absorbed in normal fault opening is not determined by the geometry.

Applying the displacement profile to multi-segment and multi-fault ruptures

Approach A: "One-size-fits-all"

-- Take D(L) from a regression the full rupture length. D(L) would come from Task R2. Uncertainty in D(L) may be taken empirically from the range of $D_{observed}$ around the regression.

-- Apply a full *sinesqrt* piecewise as though the pieces connected collinearly end-to-end.

-- Apply the displacement in the direction of slip on each section. So SS is generally horizontal, but normal and reverse are applied in the dip direction.

-- Accept the geometric problems in this solution - esp. that the horizontal components of slip will not match.

-- Estimate magnitude from the summed areas, displacements, and down-dip widths as recommended from Task R2.

Approach B: "Fault-style specific"

-- Take D(L) from a regression the full rupture length, as in Approach A.

-- For sub-parallel strike-slip segments/sections:

Apply rainbows to each involved fault separately

Scale each upward until the average displacement over the total is satisfied

Geometry at ends of segments, if realistic, will include enough overlap for an accommodation zone. Slip rates taper to zero where the faults taper to zero offset.

-- For normal or reverse faults joining strike slip (e.g., the example rupture)

Apply the full displacement of the SS where the two join to the second fault and taper it to the end. A half *sinesqrt* taper on the normal or reverse fault is proposed.

If it joins yet another fault, boxcar to third fault then taper.

Apply slip in the slip direction of the individual sections.

Scale each up so that the summed displacement over the total rupture returns average D from the first step.

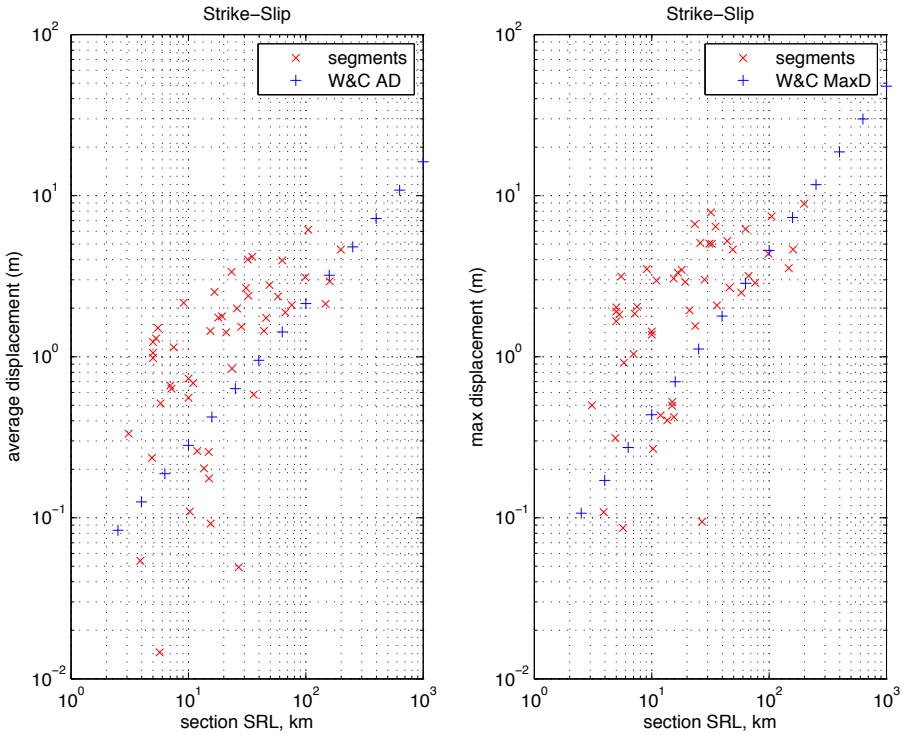


Figure 1. Average displacement vs. section rupture length (left) and maximum displacement vs. section rupture length for strike-slip ruptures in the Wesnousky (2008) database that have step-overs of 1 km or more. Blue "+" are the AD vs. SRL and MaxD vs. SRL "all" regression of Wells and Coppersmith (1994).

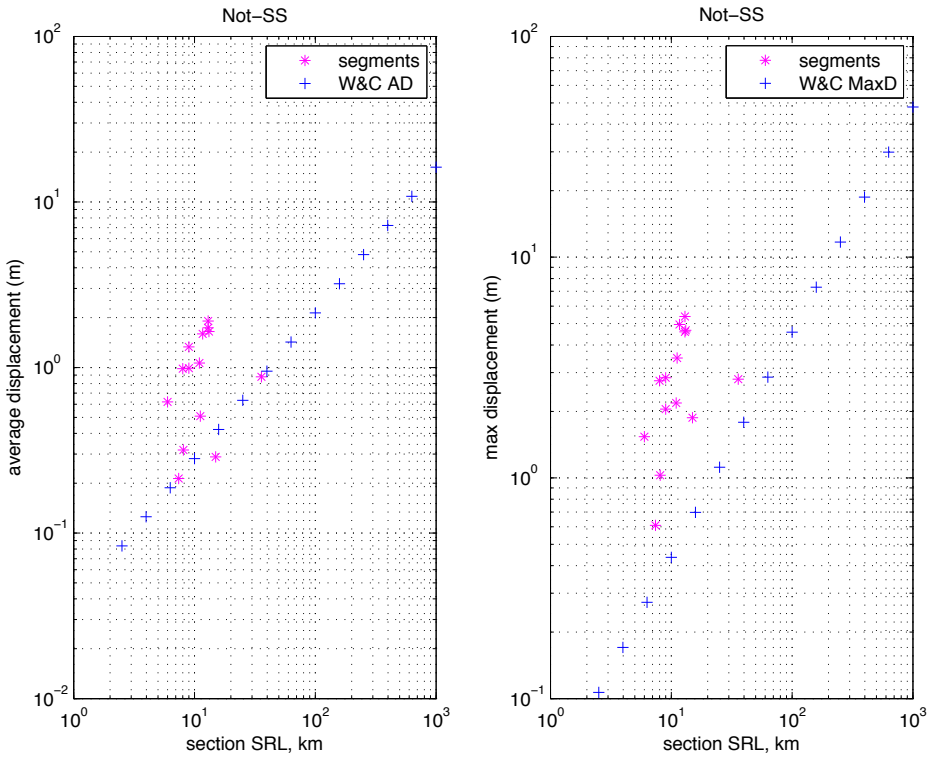


Figure 2. Average displacement vs. section rupture length (left) and maximum displacement vs. section rupture length for reverse and normal slip ruptures in the Wesnousky (2008) database that have step-overs of 1 km or more. Details are as in Figure 1.

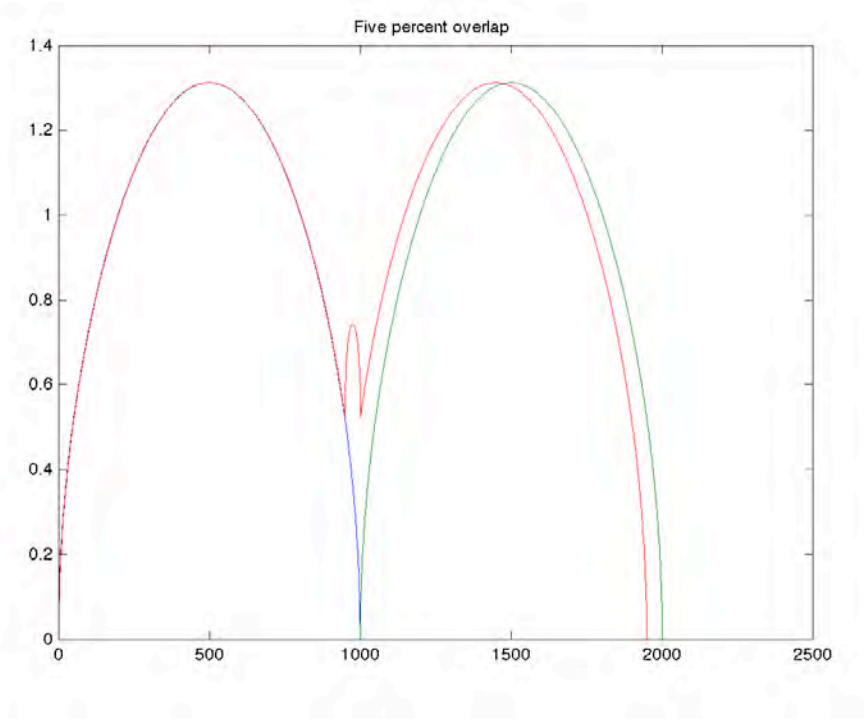


Figure 3. Overlapping "rainbows" illustrate the scale of structures to consider when assessing the problems posed at step-overs and ends of overlapping faults. In this example the displacement profiles overlap by only 5% of the total rupture length, but 60% of the average displacement for the linking ruptures is recovered in the overlapping region. Summed displacements for less than 10% overlap would total the average for the rupture.

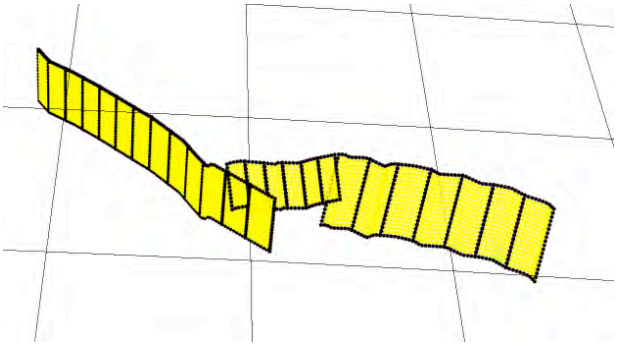


Figure 4: Complicated multi-fault rupture. Strike-slip and two reverse vaults are linked.

Task R2 – Evaluate Magnitude-Scaling Relationships and Depth of Rupture

Task Leader: Bruce Shaw (LDEO)

Please see the following document:

http://www.wgcep.org/sites/wgcep.org/files/Task_R2_v4.pdf

Task R3 - Paleoseismic Recurrence Interval Estimates

Task Leader: Tom Parsons (USGS)

This task is aimed at developing mean interevent times and their uncertainties at points where paleoseismic data are observed. Methods that are likely to be used to compute the long term large earthquake rates on California faults require a mean recurrence interval estimate, and prefer a probability density function to compare with. These include the earthquake rate inversion method of Field and Page (2010), and physics-based earthquake simulators under consideration (e.g., Ward, 2000; Rundle et al., 2006; Dieterich and Richards-Dinger, 2010).

There are a number of approaches to this problem of varying complexity. Most commonly, variants of maximum-likelihood techniques are applied to observed series to estimate recurrence parameters (e.g., Nishenko and Buland, 1987; Davis et al., 1989; Wu et al., 1991; Ogata, 1999). To account for dating uncertainty, Ellsworth et al. (1999) developed a process in which carbon-dating-PDFs of paleoseismic intervals are bootstrapped, and then the results are used to develop Brownian Passage Time parameters for recurrence interval and coefficient of variation using a maximum likelihood technique. Similarly, Biasi et al. (2002) draw repeatedly from allowable paleointerval ranges to find most likely distribution parameters, but also use a Bayesian approach to assess the true intervals. Console et al. (2008) use a Monte Carlo method to account for open intervals and event time uncertainty. Parsons (2008a) uses a Monte Carlo approach to expand the range of possible solutions considered past those contained in the observed intervals, and forces all the acceptable solutions to fit all of the observations. This is most useful for limited series.

The UCERF2 exercise calculated earthquake rates according to fault slip rates. Theses values were then checked for consistency with paleoseismic event rates. Two methods were used, in southern California, the stringing-pearls method of Biasi and Weldon (2009) were applied along the San Andreas fault. For the rest of California, the Monte Carlo approach (Parsons, 2008a) was used to test observed intervals against different recurrence models.

Database

The paleoseismic database contains lists of intervals within which events of unspecified size caused surface ruptures. An initial step under this task will be to reassess the database of paleoseismic information, and potentially add new observations. Uniform, consistent criteria for adding or omitting data based on quality and/or publication status will be developed. A peer-review, published constraint for data inclusion is preferable. A further database issue that

needs addressing is that paleoseismic records are incomplete because many earthquakes do not reach the surface and some of those that do reach the surface are not be recorded. To accommodate this problem many studies (including UCERF2) use the probability that a rupture reaches the surface, based on observation of historical ruptures, as the probability that an event will be observed in a trench. Because ruptures with small displacements are more difficult to preserve and be interpreted in the geologic record and because there are hiatuses in trench records, the probability that an event will be recognized in a trench must be lower than the probability that the event reaches the surface.

We propose to look at a representative suite of trench studies in different types of geologic materials to better understand what factors affect the resolution of paleoevents and thus develop general rules to estimate the completeness of the record for different sized events. Three approaches are likely to provide useful results: 1) Apply uniform semi-quantitative measures of event quality (e.g., Scharer et al. 2007) to all trench sites to compare event quality between different types of sites, 2) Make summaries of the amount of offset recognized in trench studies to get a better idea of what the lower threshold of recognition of offset is in different geologic settings. For example, if the smallest offsets mapped in alluvial deposits is 10 cm, that is likely the lower resolution and surface ruptures that are likely to be less than 10 cm are not likely to be found. 3) Apply methods developed to access the completeness of stratigraphic records to trenches to assess the completeness of our trench records and to identify significant hiatuses in the records. For example, at the Frazier Mountain site on the San Andreas fault there are on average one distinguishable clastic unit per decade, but some portions of the section preserve only a few layers per century and other portions many units per decade. By focusing on the resolution through time we can assess what parts of the past are well sampled and what portions are not, and thus infer when our records are likely to be complete or incomplete, and when we are likely or unlikely to recognize events that are seen in neighboring sites.

MRI calculations

After database issues are fully resolved and updated, we will recalculate recurrence intervals at each paleosite. Optimally, we would have enough earthquake intervals to unequivocally define the shape of recurrence distributions on faults. However, doing that requires at least ~25-50 intervals to gain the necessary resolution (e.g., Matthews et al., 2002), and there are presently no well-resolved continuous series of large earthquakes that long. A large sample is needed because earthquake recurrence-time distributions are skewed asymmetrically about their means (e.g., Nishenko and Buland, 1987; Hagiwara, 1974; Kagan and Knopoff, 1987). For example, if one wants to characterize mean recurrence from a limited sample, that sample has highest probability of being drawn from the mode (most frequent value) of the distribution, which is effective for normal parent distributions where the mode equals the mean (Fig. R31). However, if one averages the small sample drawn from a skewed parent distribution, then the resulting value tends to fall somewhere between the mode and the actual mean. Tests with commonly used earthquake recurrence distributions show that this effect causes sample means to be ~60%-70% of the parent mean value (Parsons, 2008a) (Fig. R31).

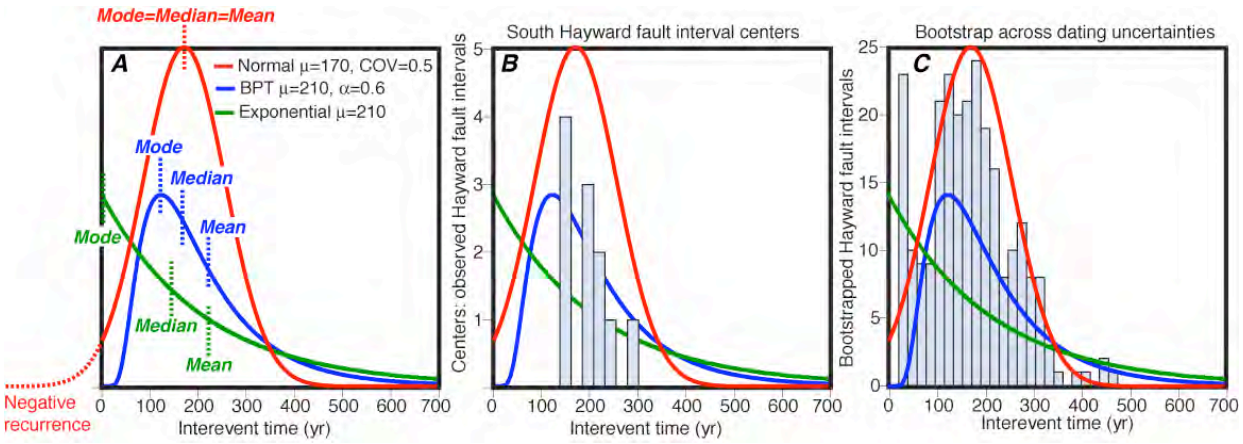


Figure 3: (a) Example earthquake recurrence distributions are compared with a normal distribution. The blue curve is Brownian Passage Time (BPT) (Kagan and Knopoff, 1987; Matthews et al., 2002), which is a commonly used time-dependent distribution. The green curve is an exponential distribution and is used for time independent earthquake probability calculations. Earthquake recurrence distributions are skewed such that the modes are not the same as the means as in the normal distribution; thus a small sampling will most likely have some bias towards the mode, and the sample mean will often underestimate the parent distribution mean. The normal distribution is unacceptable for earthquake recurrence because it can allow negative recurrence times. In (b), a histogram of south Hayward fault paleoseismic event times (interval centers) appears most consistent with all distributions. In (c), a histogram of a bootstrap over full dating uncertainties is shown that results in a mode equal to the mean as in a normal distribution with the exception of allowing short intervals like the exponential.

We anticipate enhanced use of paleoseismic observations for UCERF3. The UCERF2 rate model was tested for consistency with paleoseismically-determined rates, but it was not constrained by them. A desire expressed by UCERF3 earthquake rate modelers is to have recurrence probability density functions (PDF's) for every site. This need requires an assumption of recurrence distribution shape up front. However, this assumption must be made to produce earthquake probability values anyway. Methods we are likely to employ for this are those of Biasi (2002) and Parsons (2008a), which for time dependence give a range of possible mean recurrence intervals and coefficients of variation. Each of these parameter combinations defines a unique recurrence PDF. The full array of combinations could be used to constrain rate models, or more simply, the most likely combination could be applied.

Direct probability calculations can be made at points based on paleoseismic information. In some cases where there are quality paleoseismic data, observed intervals can be developed into

rate models that are independent of slip-rate based earthquake rate calculations. Because these values are calculated assuming an underlying recurrence model, they can have narrower uncertainty ranges than those caused by fault geometry, slip-rate, and segment boundary definitions (Figure R32). These values could be reasonably given some weight in the logic tree at the probability-calculation step, particularly if unsegmented fault models will be used.

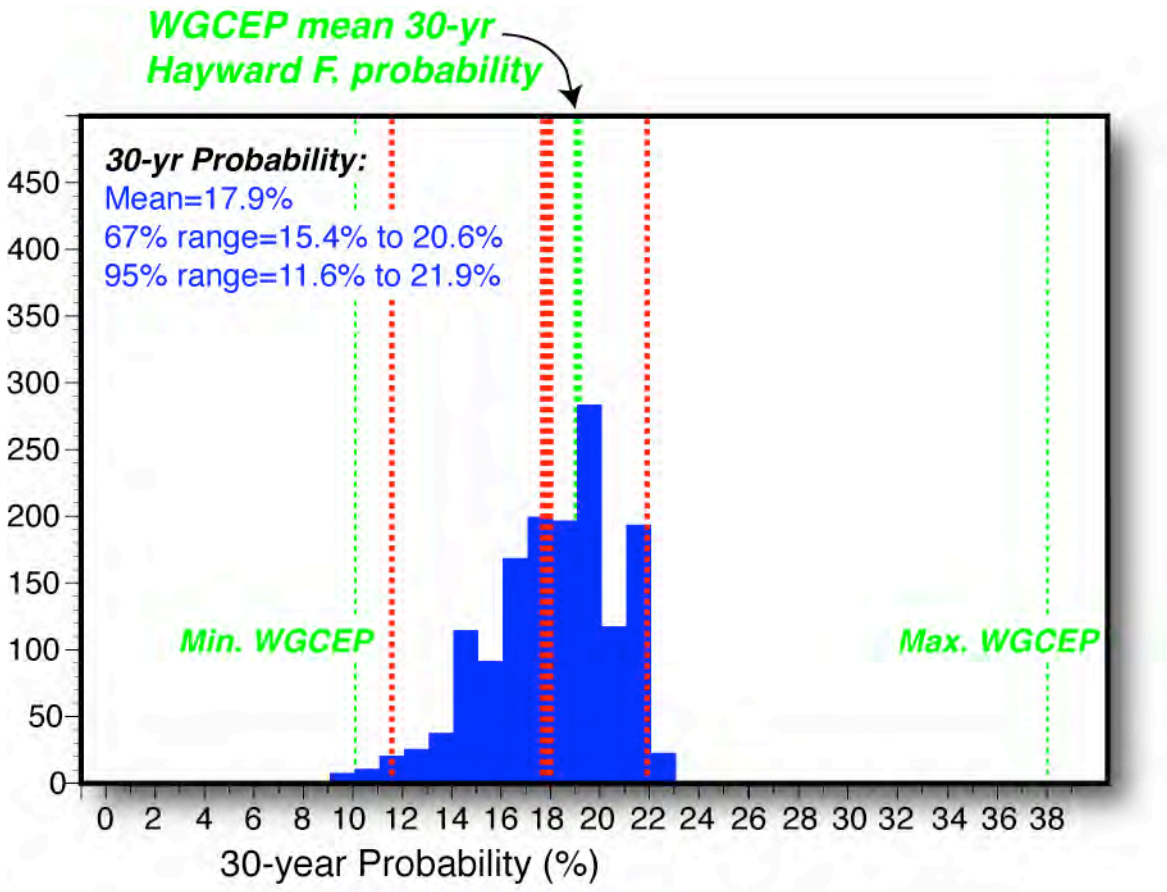


Figure R32: Comparison of Hayward fault probability calculated by UCERF2, and by direct use of intervals from paleoseimic observations. The range of possible probabilities is narrower in the paleoseismic example.

On longer faults with multiple paleosites, earthquake rupture histories can be found by linking multiple observations along the same fault strand; the "stringing pearls" method of Biasi et al. (2009). This moves us out of a point process and creates a second dimension of rupture extent, adding a magnitude constraint on fault segments. This can limit the solution space for inversions because they would then be required to produce rates of different magnitude events that match the earthquake history (and its uncertainties) rather than just a magnitude threshold.

Lastly, the paleoseismic database can be used to test recurrence MRI's against one another for consistency. Some models will fit better than others, and the difference can be quantified. For example, Parsons (2008b) found a much better fit using BPT models than exponential for the Hayward fault series (Figure R33). Weighted solutions for aperiodicity and recurrence model PDF choices can be assessed at each paleosite, and extended to rupture processes using earthquake rate model solutions if desired.

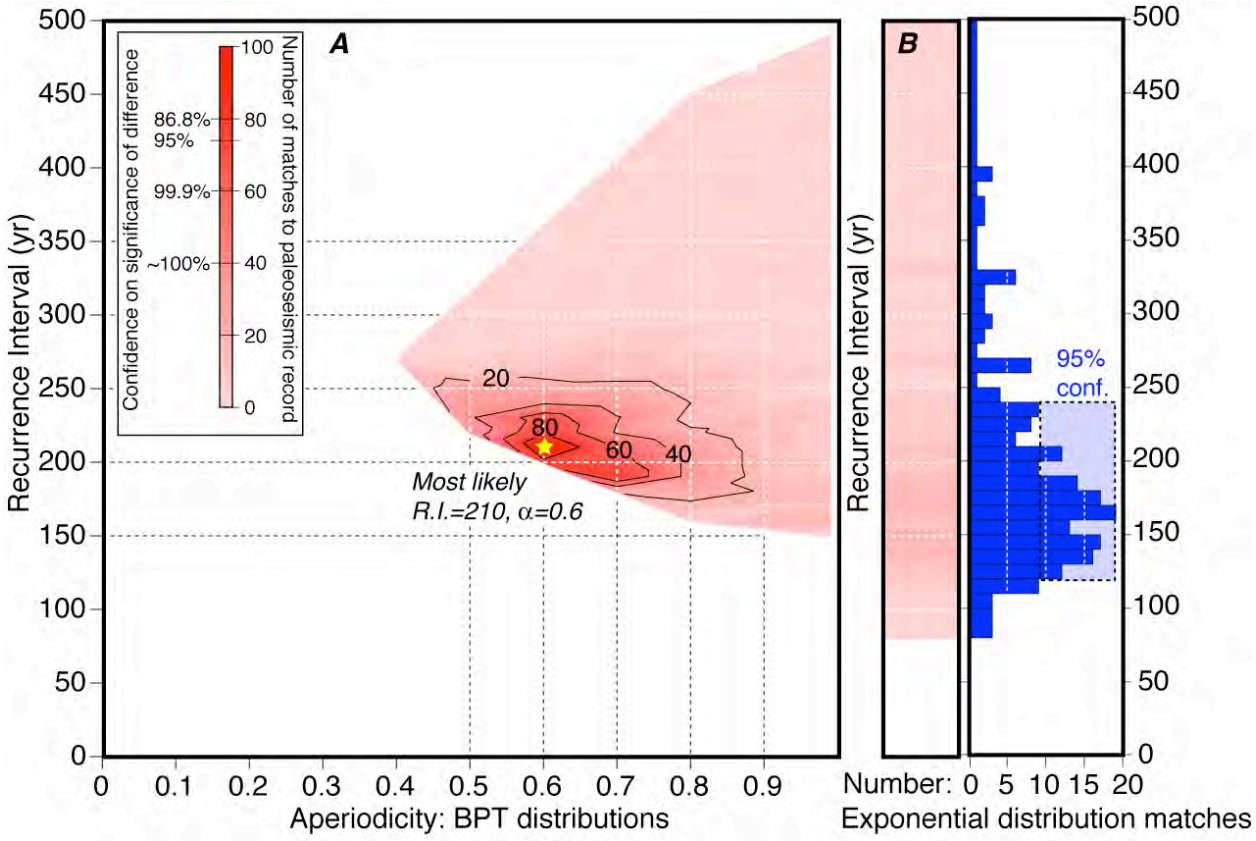


Figure R33: Contours of matches to south Hayward fault paleoseismic event series of different (a) time dependent (Brownian Passage Time) and (b) time independent (exponential) recurrence distributions. The best-fit distributions are time dependent, with recurrence intervals of μ~210 yr, and coefficient of variation α~0.6. Confidence (Z-test) on the significance of relative proportions is keyed to the contour intervals. The best-fit exponential distributions have significantly fewer matches, leading to the conclusion that earthquake recurrence on the south Hayward fault is time dependent, possibly from a stress renewal process. A histogram of exponential distribution matches is shown in (b), with 95% confidence of significance shaded, which gives the same information as the adjacent contour mapping, but in more detail.

References:

Biasi, G. P., R. J. Weldon II (2009), San Andreas Fault Rupture Scenarios from Multiple Paleoseismic Records: Stringing Pearls, Bulletin of the Seismological Society of America; April 2009; v. 99; no. 2A; p. 471-498; DOI: 10.1785/0120080287

Biasi, G. P., R. J. Weldon II, T. E. Fumal, and G. G. Seitz (2002) Paleoseismic event dating and the conditional probability of large earthquakes on the southern San Andreas fault, California, *Bull. Seismol. Soc. Am. 92,* 2761-2781, 2002.

Console R., M. Murru, G. Falcone, and F. Catalli (2008). Stress interaction effect on the occurrence probability of characteristic earthquakes in Central Apennines, J. Geophys. Res., 113, B08313, doi:10.1029/2007JB005418.

Davis, P. M., D. D. Jackson, and Y. Y. Kagan (1989), The longer it has been since the last earthquake, the longer the expected time till the next?, Bull. Seismol. Soc. Am., 79, 1439–1456.

Dieterich, J.H. and K.B. Richards-Dinger (2010). Earthquake recurrence in simulated fault systems, Pure And Applied Geophysics, in press.

Ellsworth, W. L., M. V. Matthews, R. M. Nadeau, S. P. Nishenko, P. A. Reasenberg, and R. W. Simpson (1999), A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities, U.S. Geol. Surv. Open File Rep. OF99-520, 22 pp.

Field, E. H., and M. T. Page (2010), A relatively objective, reproducible, and extensible methodology for estimating earthquake-rupture rates on a fault or fault system, submitted to B.S.S.A.

Hagiwara, Y., Probability of earthquake occurrence as obtained from a Weibull distribution analysis of crustal strain, *Tectonophysics*, *23*, 323-318, 1974.

Kagan, Y. Y., and L. Knopoff, Random stress and earthquake statistics: time dependence, *Geophys. J. R. Astron. Soc. 88*, 723-731, 1987.

Matthews, M. V., W. L. Ellsworth, and P. A. Reasenberg, A Brownian Model for Recurrent Earthquakes, *Bull. Seismol. Soc. Am., 92,* 2233 - 2250, 2002.

Nishenko, S. P., and R. Buland (1987), A generic recurrence interval distribution for earthquake forecasting, Bull. Seismol. Soc. Am., 77, 1382–1399.

Ogata, Y. (1999), Estimating the hazard of rupture using uncertain occurrence times of paleoearthquakes, J. Geophys. Res., 104, 17,995– 18,014.

Parsons, T. (2008a), Monte Carlo method for determining earthquake recurrence parameters from short paleoseismic catalogs: Example calculations for California, Journal of Geophysical Research, v. 113, doi:10.1029/2007JB004998.

Parsons, T. (2008b), Earthquake recurrence on the south Hayward fault is most consistent with a time dependent, renewal process, Geophysical Research Letters, v. 36, doi:10.1029/2008GL035887.

Rundle, P. B., J. B. Rundle, K. F. Tiampo, A. Donnellan, and D. L. Turcotte (2006). Virtual California: fault model, frictional parameters, applications, Pure Appl. Geophys. 163, no. 9, 1819–1846.

Scharer, K. M., R. J. Weldon, II, T. E. Fumal, and G. P. Biasi (2007), Paleoearthquakes on the Southern San Andreas Fault, Wrightwood, California, 3000 to 1500 B.C.: A New Method for Evaluating Paleoseismic Evidence and Earthquake Horizons, Bulletin of the Seismological Society of America; August 2007; v. 97; no. 4; p. 1054-1093; DOI: 10.1785/0120060137.

Ward, S. N. (2000). San Francisco Bay area earthquake simulations: A step toward a standard physical earthquake model, Bull. Seismol. Soc. Am. 90, 370–386.

Wu, S.-C., C. Cornell, and S. R. Winterstein (1991), Estimation of correlations among characteristic earthquakes, Proc. 4th Int. Conf. Seismic Zonation, Stanford, California, 11 – 18.

Task R4: Probability of Seeing Events in a Paleo Trench

Task Leader: Ray Weldon (UO)

Statement of the Problem: Paleoseismic records are incomplete because many earthquakes do not reach the surface and some of those that do reach the surface are not be recorded. To accommodate this problem many studies (including UCERF2) use the probability that a rupture reaches the surface, based on observation of historical ruptures, as the probability that an event will be observed in a trench. Because ruptures with small displacements are more difficult to preserve and be interpreted in the geologic record and because there are hiatuses in trench records, the probability that an event will be recognized in a trench must be lower than the probability that the event reaches the surface.

Proposed Solution: We propose to look at a representative suite of trench studies in different types of geologic materials to better understand what factors affect the resolution of paleoevents and thus develop general rules to estimate the completeness of the record for different sized events. Three approaches are likely to provide useful results: 1) Apply uniform semi-quantitative measures of event quality (eg Scharer et al. BSSA, 2009) to all trench sites to compare event quality between different types of sites, 2) Make summaries of the amount of offset recognized in trench studies to get a better idea of what the lower threshold of recognition of offset is in different geologic settings. For example, if the smallest offsets mapped in alluvial deposits is 10 cm, that is likely the lower resolution and surface ruptures that are likely to be less than 10 cms are not likely to be found. 3) Apply methods developed to access the completeness of stratigraphic records to trenches to assess the completeness of our trench records and to identify significant hiatuses in the records. For example, at the Frazier Mtn site on the SAF there are on average 1 distinguishable clastic unit per decade, but some portions of the section preserve only a few layers per century and other portions many units per decade. By focusing on the resolution through time we can assess what parts of the past are well sampled and what portions are not, and thus infer when our records are likely to be complete or incomplete, and when we are likely or unlikely to recognize events that are seen in neighboring sites.

Task R5: Solve the "Grand Inversion"

Task Leader: Morgan Page (USGS)

Introduction

The purpose of the "grand inversion" is to solve for the long-term rate of all possible ("on-fault") ruptures in California. We describe several applications of this methodology: first, we apply it to northern California using UCERF2 ingredients and assumptions. We then demonstrate feasibility for the statewide problem by presenting a minimally constrained model for the entire state that employs a parallelized version of the simulated annealing algorithm. The description here assumes familiarity with the section "Earthquake Rate Models" in the main body of this report, as mathematical details are not duplicated here.

Setting up the Inversion: Data and Constraints

Except where otherwise noted, the components currently used in the inversion are as given in UCERF2 (these are to be swapped out later as newer components become available). For example, in defining D_{sr} (slip on the s^{th} subsection in the r^{th} rupture) we use the average of the magnitudes from the Hanks and Bakun and Ellsworth B relationships, as well as the tapered slip model for variations along strike. The inversion methodology solves for the rates of ruptures that are consistent with the following data and constraints:

1) **Slip Rates**. The average slip in each rupture a section participates in, multiplied by the rate of each rupture, must sum to the long-term slip rate for that section, as given by the deformation model. This constraint is applied to each fault section. The slip rates applied here come from the UCERF2 deformation model 2.1.

2) **Paleoseismic Event Rates**. The total rate of all ruptures a section participates in, multiplied by the probability each rupture is paleoseismically visible, must sum to the mean paleoseismic event rate for that section. This constraint is applied to all fault sections where there is paleoseismic data from UCERF2 (7 in Northern California, 24 total in the entire state).

3) ***A priori* Rupture Rates**. For this feasibility study, we apply an *a priori* rate to each rupture, which makes the inversion over determined (thus, a unique solution exists). As rupture rates are what we are solving for, this constraint is applied weakly in order to keep the solution as close as possible to these a priori rates (while matching the other data more precisely). Different rupture-rate constraints are used in the 2 Northern California models that we will present:

> Model 1: *UCERF2 constraint*. Of the solutions that come closest to satisfying the other constraints, we choose the one that is closest (L^2 norm) to equivalent UCERF2 rates (the mapping of UCERF2 ruptures to the equivalent inversion ruptures is far from trivial; details are available upon request, but we are confident that conclusions here do not depend on any subjectivity in this mapping).
>
> Model 2: *Smoothness constraint*. Of the solutions that come closest to satisfying the other constraints, we choose the one that is closest (L^2 norm) to a set of "smooth" rupture rates (see below for further details).

4) **Nonnegativity Constraint**. Rupture rates cannot be negative. This is a strict constraint that is not included in the system of equations but is strictly enforced in the simulated annealing algorithm, which does not search any solution space that contains negative rates.

5) **Magnitude Distribution**. For this feasibility study we assume to know the target magnitude frequency distribution exactly. The total magnitude-frequency distribution for the region is assumed to be Gutenberg-Richter with a *b*-value of 1.0 (including aftershocks). Since we are solving for the rate of on-fault, full-seismogenic-thickness ruptures, we must subtract from this the "off-fault" seismicity (representing everything else). For this purpose we assume the off-fault distribution implied by UCERF2 is correct. Subtracting this off-fault magnitude-frequency distribution from the assumed regional total gives us the target on-fault distribution for the inversion (see Figure 1). Thus the target magnitude distribution we use represents a distribution that will make the region Gutenberg-Richter, provided that the UCERF2 background magnitude distribution is correct.

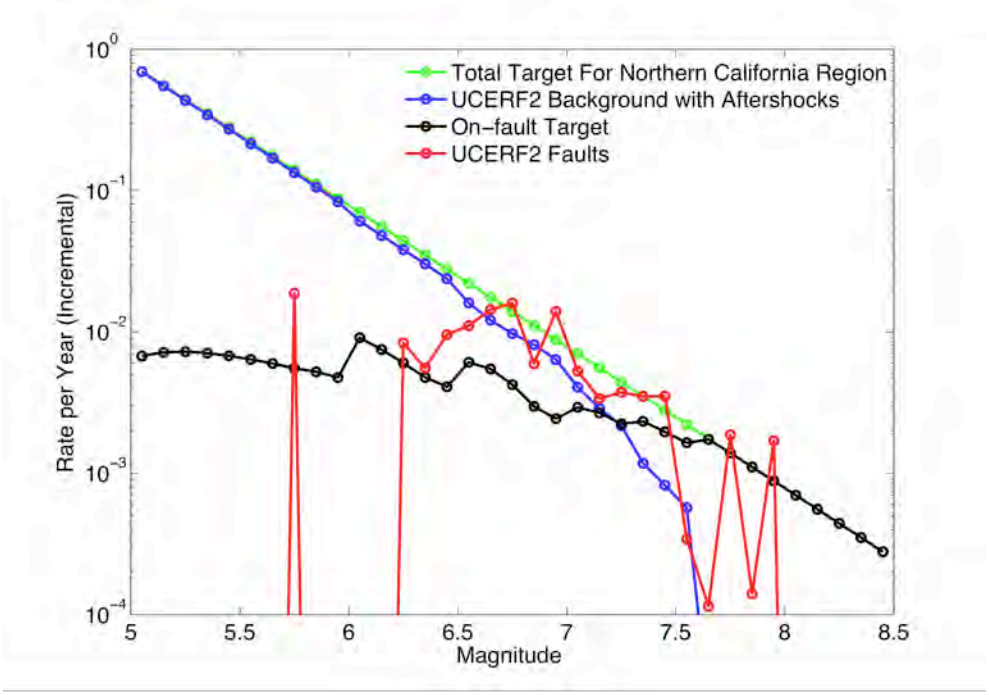


Figure 1. Target Magnitude Distributions compared to UCERF2. The green line shows the total target magnitude distribution for the Northern California region, which is a Gutenberg-Richter distribution with *b=1.0*. The magnitude distribution for background (off-fault) seismicity in UCERF2 is shown in blue. Aftershocks have been added to this distribution. Subtracting the blue from the green gives the target distribution for on-fault seismicity, which both Model 1 and Model 2 in this report are constrained to match. Finally, the actual on-fault magnitude distribution implied by the UCERF2 ruptures is shown in red. Its overprediction relative to the black rates are a problem known as the "bulge". This overprediction results in a deviation from Gutenberg-Richter statistics for total seismicity in the Northern California UCERF2 model.

With the exception of the nonnegativity constraint, these constraints are linear and can be combined to form a matrix equation of the form $Ax=d$. Our task is to set up this matrix equation and solve for x, the rate of all ruptures, given A and d.

The Simulated Annealing Algorithm

We solve for x using a simulated annealing algorithm. This algorithm simulates the slow cooling of a physical material to form a crystal. It has the following steps:

1. Set x equal to initial solution x_0. This could be a random starting solution (different each time the algorithm is run), a starting solution of zeros, or, as in the cases presented here, the UCERF2 rates (Model 1) or a smooth starting model (Model 2).

2. Lower the "temperature", T, from 1 to 0 over a specified number of iterations. We lower the temperature linearly, although different approaches to annealing specify different functions for the temperature. The more iterations (i.e., the more slowly the temperature is lowered), the better the final solution will be. Over each simulated annealing iteration, the following is done:

   a. One element of x (one rupture rate) is chosen at random. This element is then perturbed randomly. It is here that the nonnegativity constraint is applied – if it becomes negative, the element is set to 0. Our algorithm does not use smaller perturbations as the temperature is lowered (this was tested but did not result in faster convergence times).

   b. The misfit for the perturbed x, x_{new}, is calculated, and from this the "energy" of that solution: $E_{new} = (Ax_{new} - d)^2$.

   c. The transition probability, P, is calculated based on the change in energy (between the previous state and the perturbed state) and the current temperature T. If the new model is better, $P=1$. Therefore, a new model is always kept if it is better. If the new model is worse, it is sometimes kept, and this is more likely early in the annealing process when the temperature is high. If $E < E_{new}$, $P = e^{(E-Enew)/T}$.

3. Once the annealing schedule is completed, the best solution x found during the search (the solution with the lowest energy) is returned. (Note that this is a departure from "pure" simulated annealing, which returns the last state found. In some cases the final state will not be the best solution found, since occasionally solutions are discarded for worse solutions).

Simulated annealing works similarly to other nonlinear algorithms such as the genetic algorithm. One advantage of simulated annealing is that there is a mathematical basis for it: besides the analogy to annealing a physical material, the simulated annealing algorithm will find the global minimum given infinite cooling time (Granville *et al*. 1994).

There are several advantages of this algorithm in contrast to other approaches such as the nonnegative least-squares algorithm. First, the simulated annealing algorithm scales well as the problem size increases. (In fact, it would not be computationally feasible for us to use the nonnegative least-squares algorithm to solve a problem of this size.) It is designed to efficiently search a large parameter space without getting stuck in local minima. Next, quite importantly, for an underdetermined problem the simulated annealing algorithm gives multiple solutions (at varying levels of misfit depending on the annealing schedule). Thus both the resolution error (the range of models that satisfy one iteration of the data) and the data error (the impact of parameter uncertainty on the model) can be sampled. Finally, simulated annealing can allow us to include other nonlinear constraints in the inversion apart from nonnegativity; for example, we

can easily incorporate an inequality constraint on the magnitude-frequency distribution, as described in the main report.

Sample Northern California Inversion Solutions

We next show 2 models solving for rupture rates on the major faults in Northern California. The faults are discretized into sections. These sections are for numerical tractability and do not represent geologic segments. At maximum, the length of the fault sections are half the seismogenic width (this varies along the faults; on average the a fault section is about 7-km along strike). Sections are linked together to form ruptures; the smallest ruptures contain two fault sections (thus, all rupture lengths are greater than or equal to the seismogenic thickness). All viable ruptures are generated from the digitized fault sections that satisfy the following rules:

1) All fault sections connect within 5 km or less.
2) Strikes cannot vary by more than 45 degrees between neighboring subsections.
3) Strikes cannot change by more than 90 degrees along the rupture length.
4) Rakes cannot change by more than 90 degrees between subsections.
5) Ruptures cannot include a given subsection more than once.

These are sample rules for feasibility; the parameters in the rules are adjustable in the code (as there are certainly too many ruptures to define by hand). We include all faults in Northern California in the UCERF2 fault model with the exception of the creeping segment of the San Andreas Fault. The total number of fault sections is 696. Using the rules described above results in 17,777 ruptures.

Currently we are not including any penalty for multi-fault ruptures. A concurrent task (R6) is developing rupture weights based on Coulomb stress. A metric has been developed that quantifies the degree to which given sections are kinematically consistent when they rupture together. These could easily be added to the inversion through an a priori rupture rate constraint.

We present results from two models for the Northern California long-term rupture rates. These models differ only in the *a priori* rupture-rate constraints applied. Each model is generated by using the simulated annealing method described above for 1,000,000 iterations. This takes approximately 1 hour on a single laptop computer. The solutions are stable – annealing for a longer time does not change the solution significantly, and the algorithm returns the same solution each time despite the stochastic nature of the algorithm due to the number of constraints used, which form an overdetermined problem.

Model 1: A Solution Constrained by UCERF2 Rates

The first inversion solution we present, referred to here as Model 1, is constrained to give the closest solution to the UCERF2 rupture rates that matches the data. The slip-rate fits for this model are shown in Figures 2 and 3. Note that the UCERF2 model included the Mendocino segment as a C-zone, and we have therefore not included it in the UCERF2 fault ruptures.

Figure 3 also includes the slip rate implied by UCERF2 for comparison. Note that in general UCERF2 does not fit the slip rates all that well. This results from the fact that the rates of

floating ruptures on type-B faults (and for the un-segmented option on type-A faults) do not change along strike, leading to a rainbow pattern for slip rates along the original fault sections (although slip rates are generally matched when averaged over the entire original segments). For type-A faults, recall that UCERF2 gave 50% weight to two different models – one based on expert opinion event rates (the "*a priori*" model), and another where these expert-opinion rates were modified to be moment-balanced. The extent to which these two models differ leads to discrepancies in the UCERF2-implied slip rates in Figure 3.

Model 1, the solution found with the simulated annealing inversion method, underpredicts the slip rates of small, isolated faults. The target magnitude distribution only allows a limited number of small-magnitude ruptures (it has a low *b*-value), which the inversion uses both to match sudden slip-rate drops at the ends of segments and slip-rate constraints on segments without larger ruptures.

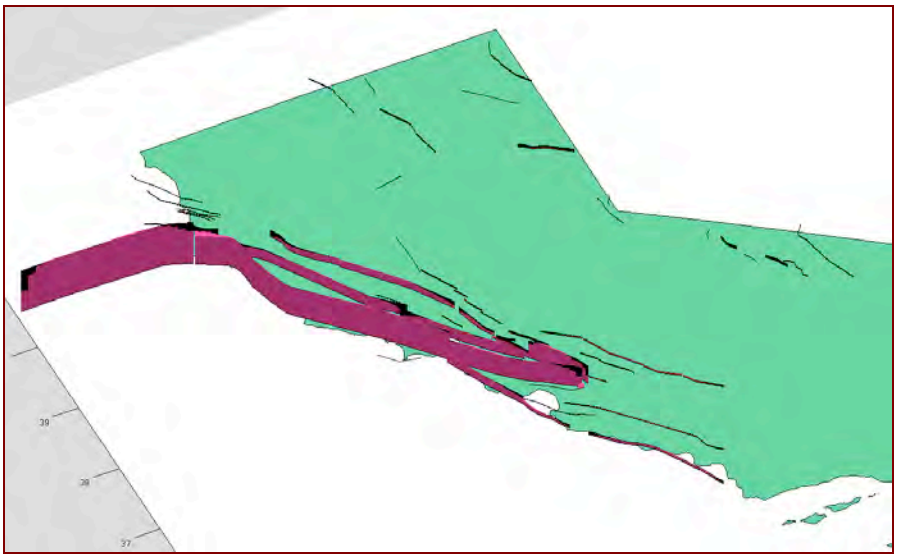


Figure 2. Slip Rate Fit for Model 1. Target slip rates are shown in black vertically projected from their associated fault traces. Synthetic slip rates for Model 1 are shown in magenta; where they overlap with target slip rates the resulting color is deep purple.

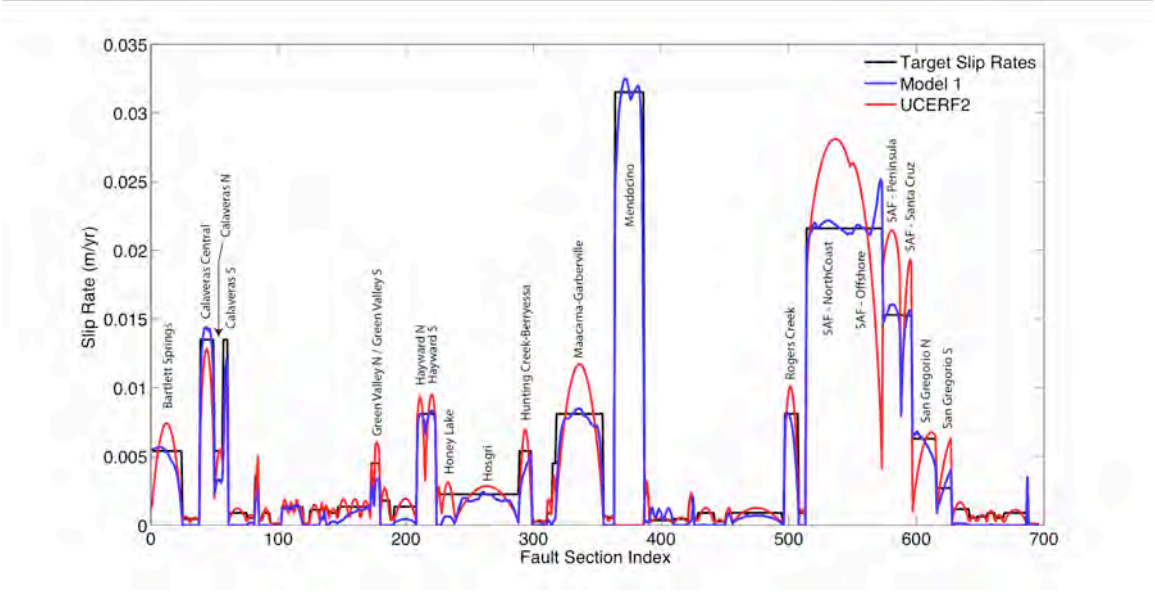


Figure 3. Comparison of slip-rate fit for Model 1 and UCERF2 slip rates. Fault Segments are ordered alphabetically; major segments are marked. UCERF2 slip rates tend to overpredict segment centers and underpredict segment endpoints due to the uniform rate of floating ruptures in UCERF2. Model 1 slip rates tend to underpredict slip rates for short, isolated faults, in part due to the imposed magnitude-frequency distribution.

The event rates produced by Model 1 and the UCERF2 solutions are shown in Figures 4 and 5. Both models match the paleoseismic data within the error bars, although for most segments the Model 1 event rate is closer to the paleoseismic mean.

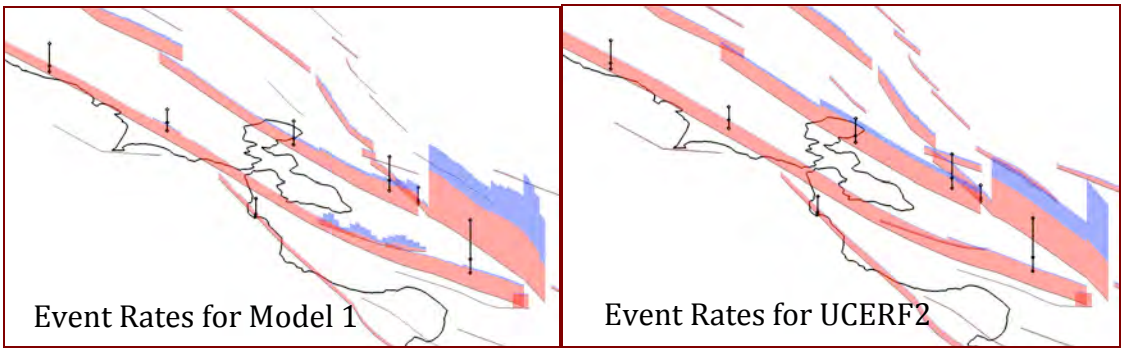


Figure 4. Event Rates for Model 1 (left) and UCERF2 (right). Blue bars are total event rates (the total rate of earthquakes a fault section participates in); pink bars show paleoseismically visible event rates determined using the relation of Youngs *et al.* (2003). Vertical black lines show paleoseismic rates and 95% confidence limits.

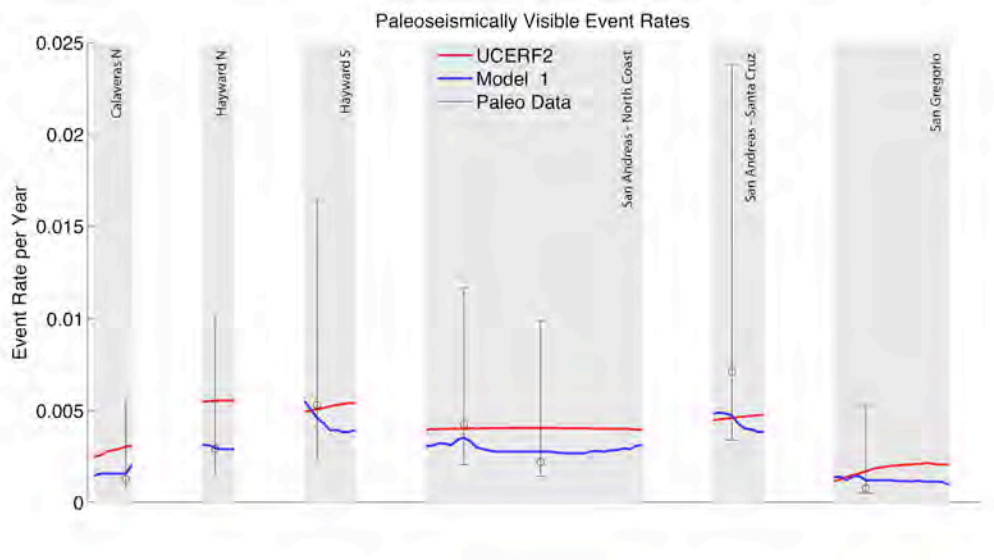


Figure 5. Event rates for Model 1 and UCERF2 on segments which have paleoseismic data. Event rates are paleoseismically-visible event rates found by convolving total event rates with the relation of Youngs *et al*. (2003).

The rates at which the fault sections participate in earthquakes of given magnitude ranges are shown in Figure 6. Both models produce a similar migration of primarily-utilized faults as the magnitudes under consideration change.

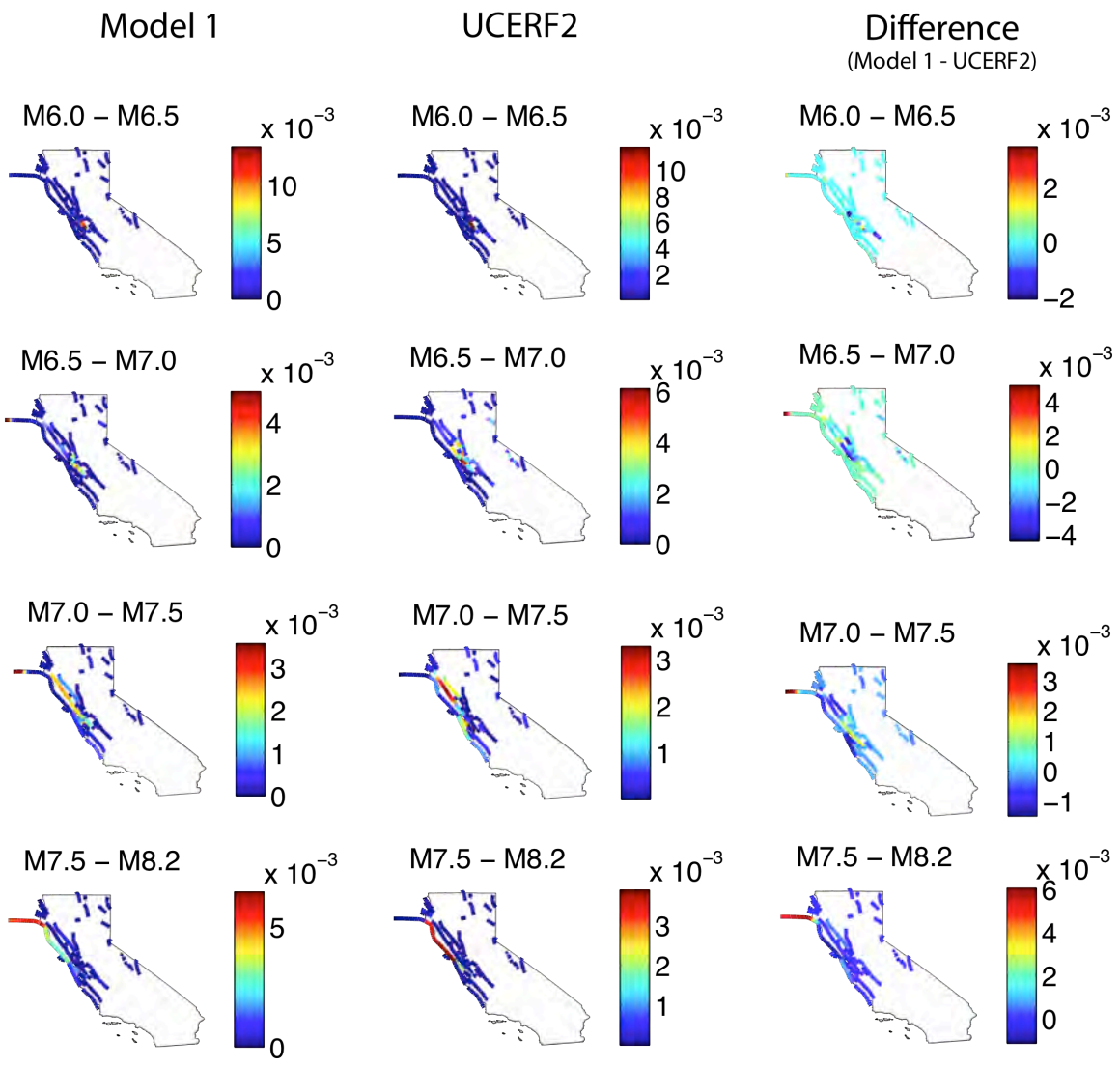


Figure 6. Participation rates (the rate at which a given fault section participations in ruptures of a given magnitude, per year) for different magnitude bins for Model 1 and UCERF2 . The third column plots the difference between the two models. Note that the color scales change. A high-resolution version of this figure is available here: http://www.wgcep.org/sites/wgcep.org/files/TaskR5_Fig6_Model1.pdf

Figure 7 compares the magnitude distribution of Model 1, which is constrained to match the target magnitude distribution, and the UCERF2 on-fault seismicity solution. Model 1 slightly overpredicts the largest ruptures (up to the largest rupture discretized in the model, which effectively defines the maximum magnitude allowed). The UCERF2 model has higher rates between magnitudes 6.5 and 7 relative to the target distribution; recall that the target distribution, when combined with the off-fault UCERF2 seismicity and its aftershocks, would give a Gutenberg-Richter distribution for the region. The UCERF2 model on the whole had an overprediction of earthquakes around magnitude 6.5, and this "bulge" problem was most pronounced in Northern California. Model 1 (and Model 2 in the next section) shows that this can be alleviated using the inversion method without sacrificing fits to the slip-rate and event-rate data.

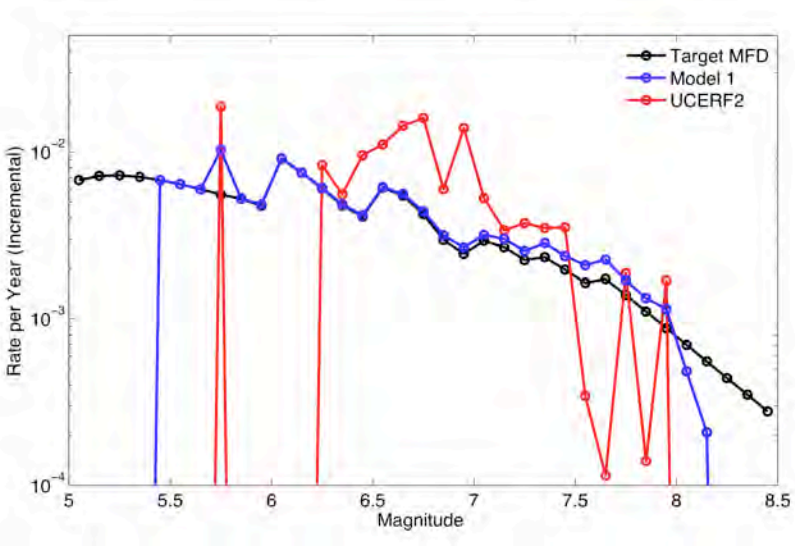


Figure 7. Incremental magnitude-frequency distributions for Model 1 and UCERF2. The target magnitude frequency distribution for Model 1 is shown in black. It is found by taking a Gutenberg-Richter distribution, with *b*=1, and subtracting off background (off-fault) UCERF2 rates (corrected to include aftershocks).

Model 2: "Smoothest" Solution

The second model we present has a different *a priori* rupture-rate constraint. We constrain Model 2 to be the closest solution to a set of "smooth" rupture rates that satisfies the data. This "smooth" solution attempts to create as uniform a magnitude distribution in space as possible. The other constraints, including the target magnitude distribution, are identical for Models 1 and 2.

The "smooth" rupture rates are found as follows: The target magnitude distribution defines a total available event rate for each magnitude bin. These rates are portioned, for each magnitude bin, into each fault section-rupture bin (for ruptures in that magnitude bin) in direct proportion to the slip-rate for each fault section. The rate constraint for each rupture is then found by summing over the rates assigned to each section-rupture bin that each rupture includes. Thus this starting model takes into account the varying slip rates among the sections and the varying overlap of ruptures from section to section. It attempts to distribute the available event rates keeping the regional magnitude distribution as uniform as possible.

The resulting "smooth" rupture rates do not match the data well because, in order to match the slip-rate data, the inversion will actually need to vary the magnitude distribution along strike (for example, more small ruptures are required near fault ends that do not link to other faults, or near sudden slip-rate changes). So there is still work for the inversion to do. However, the resulting solution will contain as little magnitude distribution changes along-strike as possible.

The slip-rate and event-rate fits for Model 2, which are quite similar to the fits for Model 1, are shown below in Figures 8 and 9.

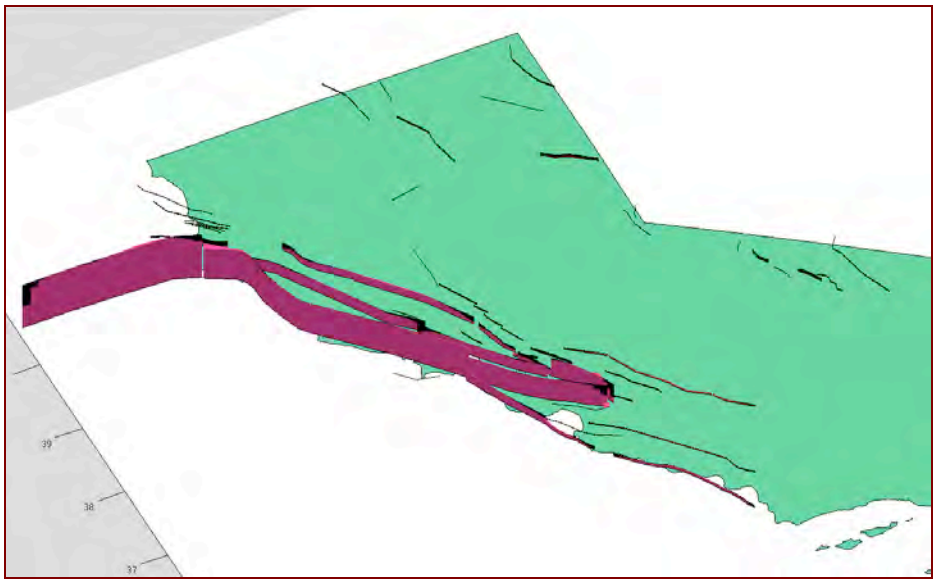


Figure 8. Slip-rate fit for Model 2. Target slip rates are shown in black vertically projected from their associated fault traces. Synthetic slip rates for Model 1 are shown in magenta; where they overlap with target slip rates the resulting color is deep purple.

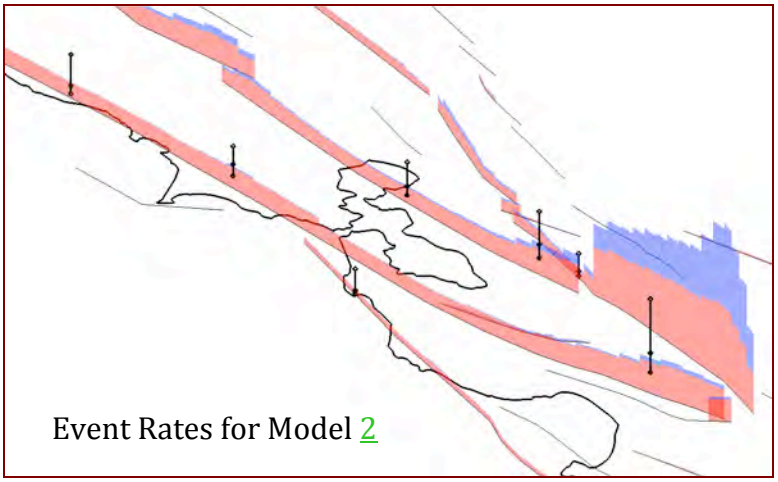


Figure 9. Event Rates for Model 2. Blue bars are total event rates; pink bars show paleoseismically visible event rates determine using the relation of Youngs et al. (2003). Vertical black lines show paleoseismic rates an 95% confidence limits.

Figure 10 shows the Model 2 participation rates for different magnitude bins and UCERF2 participation rates for comparison. There are quite apparent changes in fault participation for different magnitude increments, as there was with Model 1. This demonstrates that subregional changes in the magnitude distribution are required to match the data. The rupture-rate constraint, which should smooth some of the spatial magnitude distribution heterogeneity, is applied very weakly in Model 2 (it makes the problem overdetermined but does not add substantially to the misfit). By increasing the weight of this constraint, we can reduce the magnitude distribution changes in space, at the expense of fit to the slip rates.

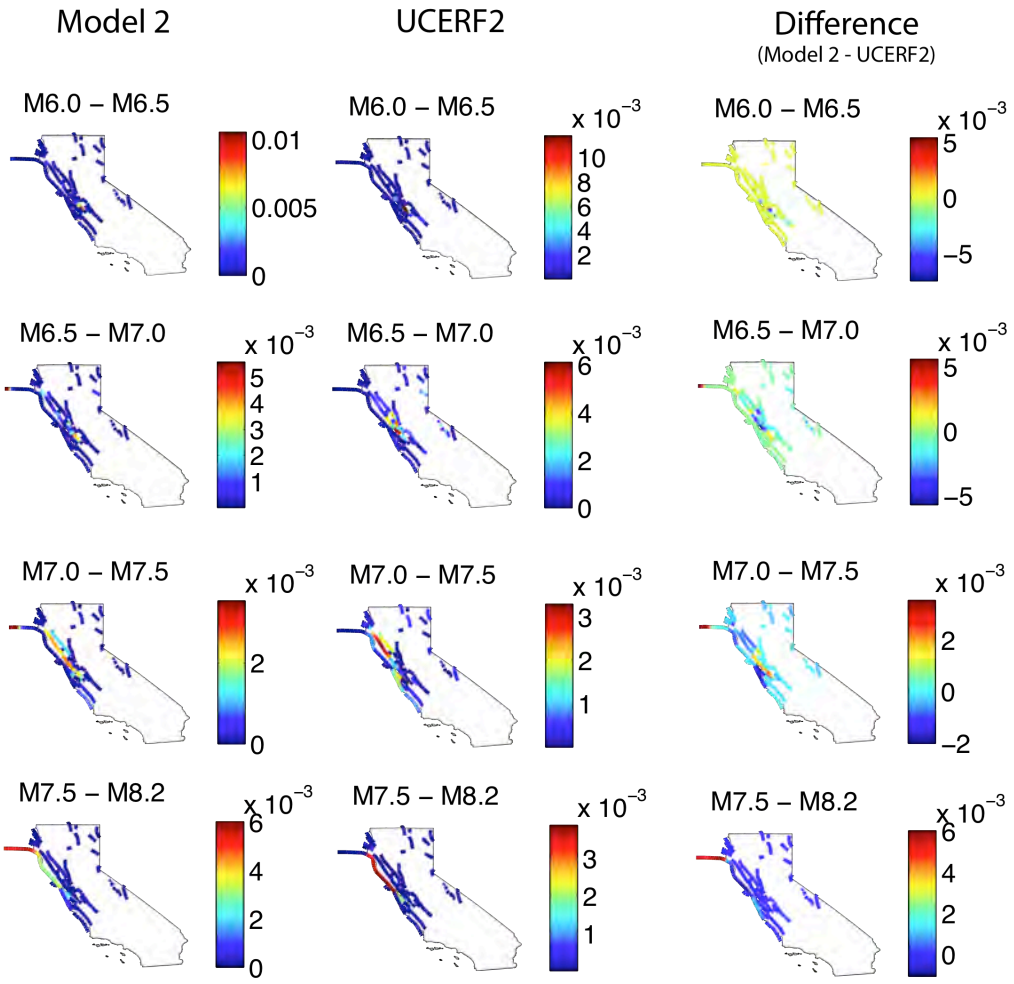


Figure 10. Participation rates for Model 2, UCERF2, and the difference for different magnitude bins. Note that the color scales change. A high-resolution version of this figure is available here:
http://www.wgcep.org/sites/wgcep.org/files/TaskR5_Fig10_Model2.pdf

The magnitude distribution of Model 2, shown in Figure 11, shows an excellent fit to the target magnitude distribution.

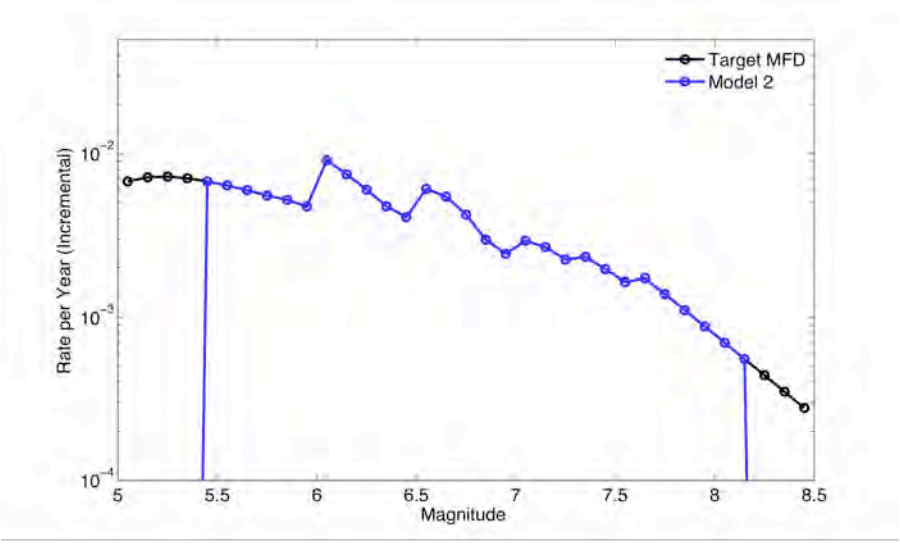


Figure 11. The magnitude distribution of Model 2 (blue) matches the target magnitude distribution (black) well for magnitude bins where there are ruptures parameterized.

Discussion

Model 1 and Model 2 are very similar in their fits to the data, in their magnitude distributions (by design), and in their participation rates for different magnitude bins. The actual rupture rate solutions are quite different, as shown below in Figure 12. Rupture rates themselves, given the over 17,000 ruptures in the model, are not well constrained by the data alone. However, many of the ruptures are nearly identical, and when looking at linear combinations of the ruptures rates (such as fits to the slip rates), or functionals of the rupture rates (perhaps hazard) we are likely to find something much more robust.

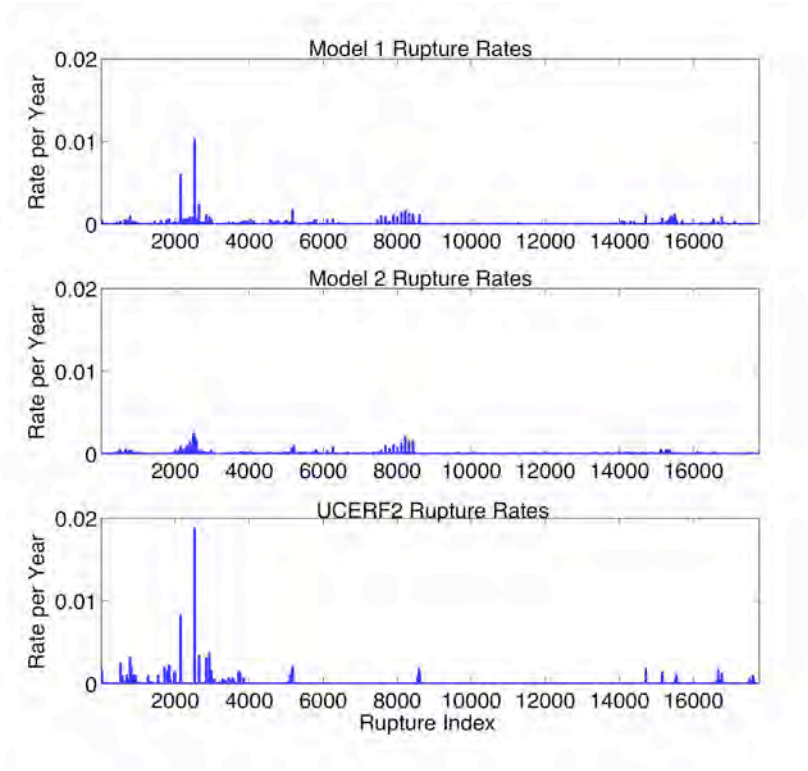


Figure 12. Rupture rates for Model 1, Model 2, and nearest ruptures to UCERF2 ruptures. The rupture-rate constraint difference between these models leads to different rupture-rate solutions, even though the data fits and participation rates are not substantially different.

Both Model 1 and Model 2 do not match slip rates well on all faults. This is due to a fundamental incompatibility between the slip rates, as defined on each fault section, and the target magnitude distribution. To match all the changes in slip rates (which are not tapered in this model between fault segments or at fault ends), the inversion must use smaller ruptures. It also must use smaller ruptures to satisfy slip rates on short, isolated faults. The target magnitude distribution used in Models 1 and 2 has quite a low *b*-value and thus does not provide the inversion with a large budget of small ruptures to use to satisfy the slip rates precisely.

Without the target magnitude distribution constraint, the inversion naturally finds quite a different solution. Below, in Figure 13, we show the slip-rate fits of an the inversion without Constraint 4 (the *a priori* rupture-rate constraint) and Constraint 5 (the magnitude-distribution constraint). The slip rates are matched nearly perfectly. This inversion uses many more small earthquakes to achieve this fit, as shown in Figure 14.

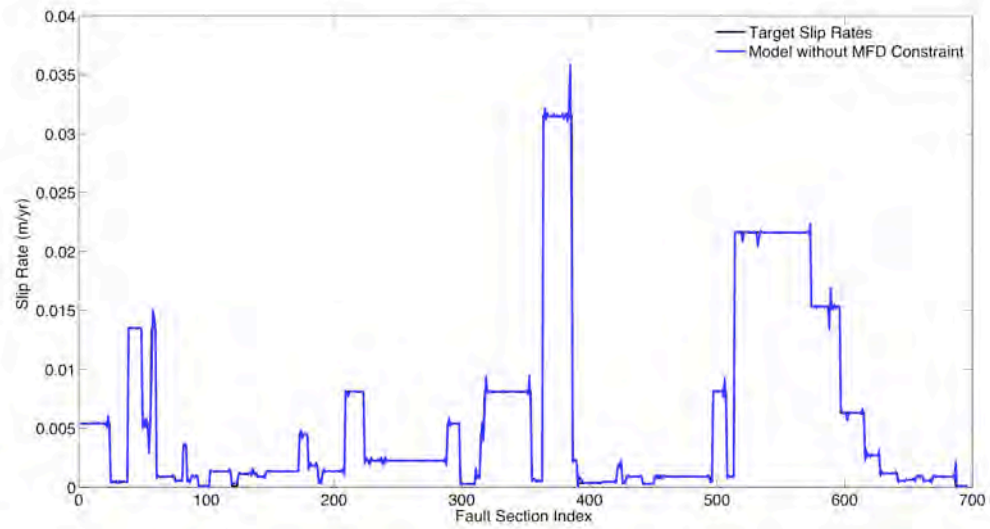


Figure 13. The slip-rate fit for a model run without a magnitude-distribution constraint (or a rupture-rate constraint). The inversion can fit the slip-rate and event-rate constraint nearly perfectly. The slip-rate constraint is somewhat incompatible with the target magnitude distribution used in Models 1 and 2, which do not fit the slip rates this well.

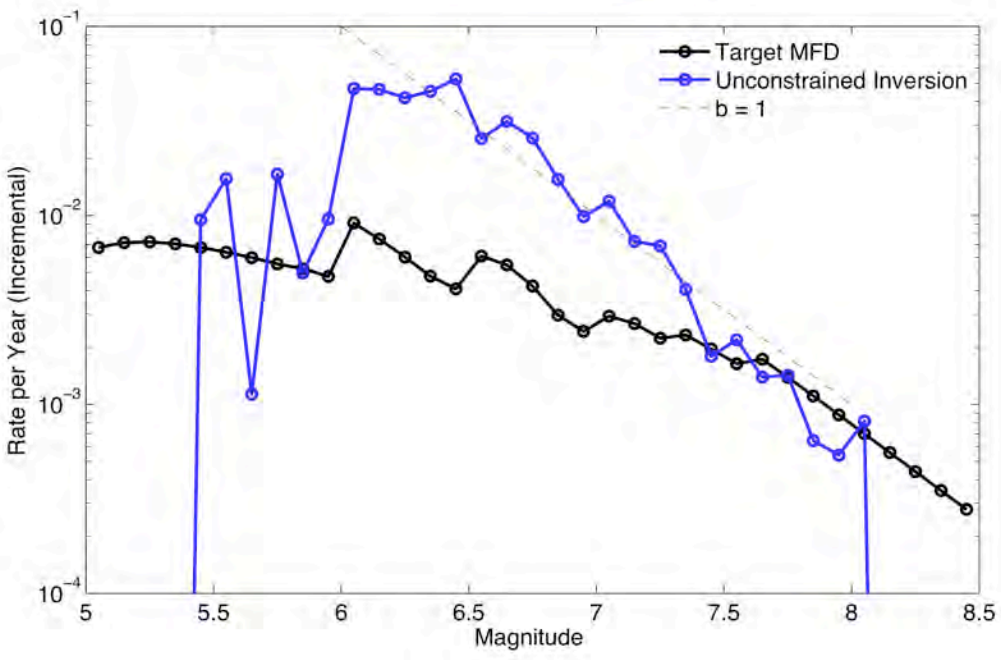


Figure 14. When the inversion is run without a magnitude distribution constraint (such as the target magnitude distribution used for Models 1 and 2 shown in black), the solution has a much higher *b*-value (see the blue line).

Relaxing the magnitude distribution is one possible solution to obtain better slip rate fits. There are other possible solutions as well: tapering the slip rates near fault endpoints, or relaxing the slip-rate constraint itself. In fact slip rates are not well-known at the level of every fault

subsection. One way to handle tradeoffs between slip rates is alternate deformation models; another way would be to incorporate the slip rate constraints over linear combinations of fault sections rather than for each individual fault section. We could also constrain slip rates only at the centers of the original sections and apply a slip-rate smoothing constraint.

Another important point is that our method does not account for ruptures that may only be partly on the parameterized faults. In our model, a rupture is either an on-fault rupture or it is in the "background" seismicity. This contributes to the difficulty of fitting the slip rates on isolated faults that in fact may include ruptures that are partly "off-fault".

California-Wide Inversion

All of the necessary ingredients needed to implement the fully constrained California-wide inversion are not ready; however, to demonstrate feasibility for a problem of this size, we present results from a minimally constrained inversion that uses all of the faults in the UCERF2 deformation model that were given slip rates. This inversion matches the following constraints that we listed above: 1) slip rates (this constraint is weighted by the slip-rate magnitude, so that low slip rates faults are fit well), 2) paleoseismic event rates, 4) nonnegativity, and 5) a magnitude distribution constraint applied to the entire California region. This is the same magnitude constraint used in the Northern California models presented above, which will give a *b=1* Gutenberg-Richter distribution for the entire state if UCERF2 background earthquakes are added to the model. Note that we do not have a rupture-rate constraint, which means this model is underdetermined.

We leave the creeping section of the San Andreas in the inversion and model the creep on this fault section (and others) in the same way as UCERF2, as a reduction in area. This allows for large earthquakes to rupture through the creeping section and link up the Southern and Northern San Andreas fault. One advantage of modeling creep in this way is that we get correct magnitude-area scaling for fault sections with a lot of creep (M6 earthquakes in Parkfield, for example, are longer than M6 earthquakes elsewhere). An alternative way to model aseismicity would be as a reduction in slip rate, which would have the effect of significantly reducing the likelihood of thoroughgoing ruptures on fault sections with a lot of creep.

Parallelization of the Simulated Annealing Algorithm

In order to tackle the additional computational demands of the statewide inversion, a parallel version of the simulated annealing algorithm was implemented by Kevin Milner (USC). This algorithm runs the serial version of simulated annealing (as described above) over a number of processors for a given number of subiterations. Then the best solution among these is kept and redistributed over the processors; this process repeats until convergence criteria (a target misfit or a given number of total iterations) is satisfied.

This algorithm was tested on the USC HPCC cluster; each node in this cluster has 8 processors. The optimal number of threads per node was found to be 4. The algorithm scales well up to 20-50 nodes, but adding nodes beyond this does not improve performance. Using the parallelized algorithm on a cluster results in average speedups of 6-20 relative to the serial algorithm. Graphs of simulated annealing "energy" vs. time (lower energy corresponds to a lower misfit) for a range of models and cluster configurations are shown in Figure 15.

After tuning the parallelized simulated annealing algorithm for optimal performance, we applied it to the minimally constrained statewide inversion described above. This solution we present took approximately 8 hours to run on 20 nodes. Slip-rate fits are shown in Figure 16; they demonstrate that with our improved weighting of this constraints slow- and fast-moving faults are fit equally well. Participation rates for the model are shown in Figure 17.

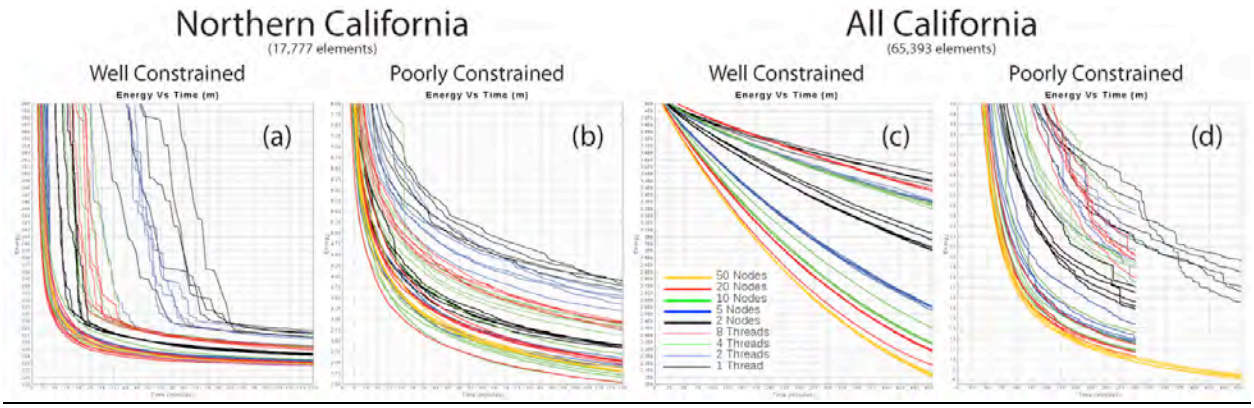


Figure 15. Speedup results for the parallelized simulated annealing algorithm depend on problem size and the amount of constraints on the solutions. Each line shows the average of 5 runs using the number of nodes shown in the legend. More negatively-sloped lines indicate faster convergence speed. Image courtesy of Kevin Milner (USC).

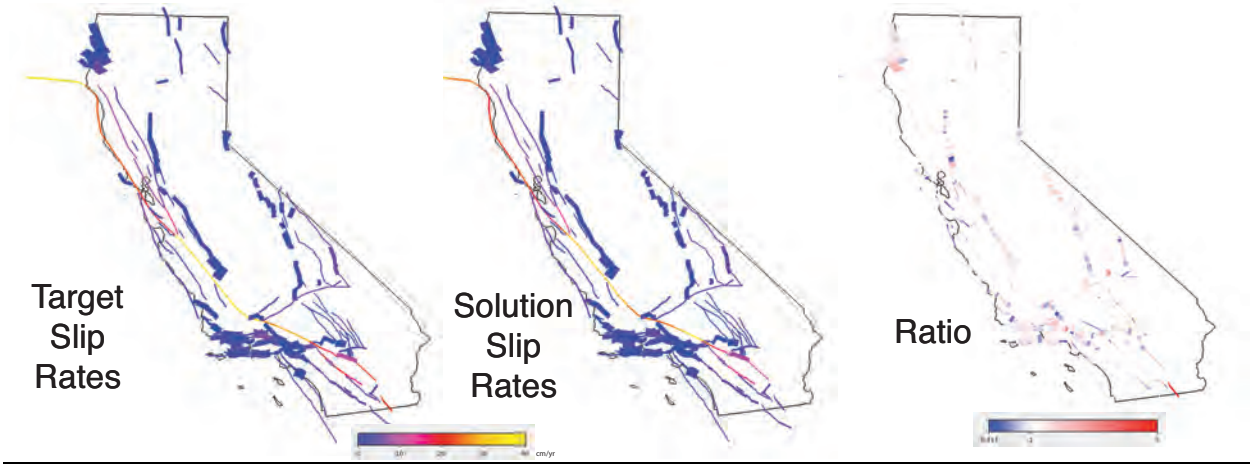


Figure 16. Slip-rate fits for the statewide inversion. The slip-rate constraints were weighted by the slip-rate for each section, so that slow-moving faults are fit as well as fast-moving faults (in terms of the ratio of the target and solution, not the difference). This eliminates the underprediction of low-slip-rate faults that was seen in the Northern California tests.

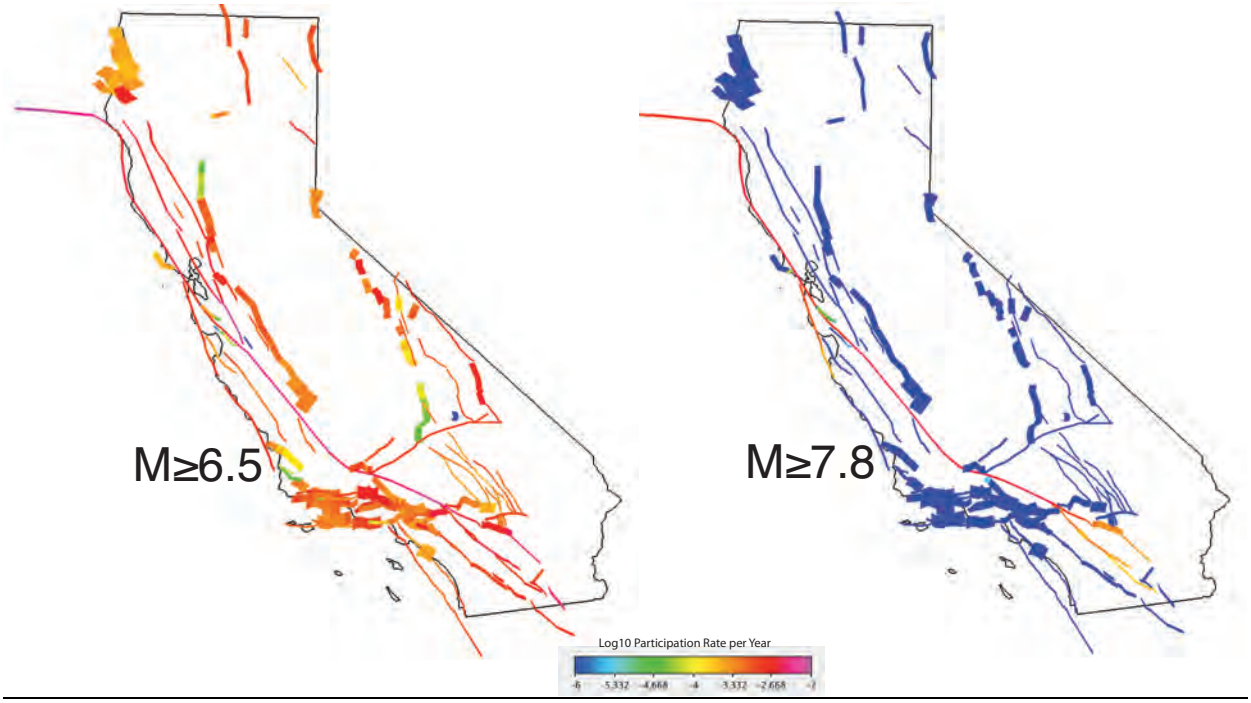


Figure 17. Participation Rates for M≥6.5 earthquakes (left), and M≥7.8 earthquakes (right). The largest earthquakes are placed by the inversion on the fastest-moving faults both because of their slip rates and their connectivity.

Next Steps and Final Remarks

The next steps are to fully implement all of the constraints needed for the statewide inversion. As discussed in the main report, we plan to present 2 alternative, well-constrained models which will be similar to the Northern California Model 1 and Model 2 presented above. One model will stay as close to UCERF2 rates while fitting the slip-rate and event-rate data better. The second model will attempt to match Gutenberg-Richter magnitude distribution on a smaller length scale (and will thus be as "smooth" as possible with regard to spatial variation of the magnitude distribution). We also plan to investigate a 3rd range of models if computational demands allow us; these models will be less constrained (neither "smooth" nor necessarily close to UCERF2 rates) but will explore the range of rates that the data allows.

There are many inputs in the inversion that can be changed to obtain a different solution. The relative weights of different constraints impact the solution, and are largely ad-hoc if the data errors are not well characterized. The inversion relies on inputs from many other project tasks to determine the possible on-fault ruptures, data inputs, and functional forms of slip distributions and probability functions. Many of these inputs will be uncertain and subject to debate. However, it is important to note that all these uncertainties existed in the old methodology for determining rupture rates as well. The inversion methodology described here eliminates the need for voting directly on rupture rates. It provides a means to easily update the model as new data becomes available. Importantly, the inversion methodology provides a mechanism to satisfy all the constraints concurrently. This was lacking in UCERF2 – expert opinion did not simultaneously satisfy slip rates and event rates, and the final model also presented magnitude

distribution problems. The inversion allows all these constraints to be satisfied to the extent that they are compatible.

References

Granville, V.; M. Krivanek, J.-P. Rasson (June 1994). "Simulated annealing: A proof of convergence". *IEEE Transactions on Pattern Analysis and Machine Intelligence* **16** (6): 652–656. doi:10.1109/34.295910.

Youngs *et al*. (2003). A Methodology for Probabilistic Fault Displacement Hazard Analysis (PFDHA), *Earthquake Spectra* **19**, 191-219.

Task R6 - Fault-to-Fault Jumping Probabilities

Task Leader: Glenn Biasi (UNR)

***Summary*:**

Define fault-to-fault jumping probabilities to use as constraints in the Grand Inversion.

***Overview*:**

Fault to fault rupture probabilities refer to the probabilities with which a seismic rupture starts on one fault and jumps or branches to another. Ruptures may in turn jump or branch by further fault-to-fault extension. The Grand Inversion (Field et al, 2011) seeks to consider and invert for some probability of all possible ruptures. An initial list of ruptures is drawn up using some rules for fault-to-fault rupture. A simple rule for fault-to-fault rupture would be to allow fault-to-fault jumps based on simple proximity. Faults closer than some D_{min} may link. In the northern California inversion the rupture set included over 20,000 members. In Southern California the fault system is more extensive and complex and the rupture set is even larger. Task R6 seeks to assist in developing the rupture set and to distinguish relative probabilities of occurrence among its members.

Primary strategies for developing a probability model:

1. Empirical rule set based on observations from mapped ground ruptures and relevant analogs.
2. Probabilities scaled from static linking stress based on Coulomb stress changes.
3. Probabilities or advice based on dynamic modeling.

We also considered the potential contributions of higher resolution fault maps and inferences from distributions of microseismicity at fault junctions. At present these contributions are considered complimentary to the list above and not in themselves a basis for developing probabilities of fault-to-fault rupture.

A mini-workshop was held September 10, 2011 where strategies for developing fault-to-fault ruptures were discussed and a path forward was outlined.

Outline of the plan forward for Task R6, Fault-to-Fault Rupture Probabilities

We envision development of fault-to-fault rupture probabilities on three parallel tracks. These tracks were suggested at the September 10, 2011 UCERF discussion in Palm Springs. Probabilities would be applied to the ensemble of ruptures developed independently for the Grand Inversion using fault-to-fault separation distance and sanity check rules.

***1. Lowest risk, most readily applied*:**

Apply the ensemble of documented and applicable surface ruptures to develop probabilities of jumps and fault-to-fault rupture. This approach uses the fact that real ruptures have found physical solutions to the fault-to-fault problem without having to resolve the physical basis for those solutions.

(a) The simplest probability implementation is based on counting (Wesnousky and Biasi, 2011) that indicates that about half of the ruptures in the Wesnousky (2008) collection include a stepover of 1 km or more. The fraction decreases by about 50% per additional stepover. Assuming for the moment that stepovers may be identified with fault-to-fault jumps, rupture probabilities are then assigned as unity if they have no stepovers, and less by 50% for each successful jump. Rupture probabilities under this model would thus be given weights of 0.5^n for n internal stepovers. The per-step probability used in UCERF would not necessarily be exactly 50% per jump; the factor will be evaluated before probabilities are finalized. The counting approach of Wesnousky and Biasi (2011) has the limitation that the relation of rupture ends to stepovers or fault-to-fault jumps are not considered. Biasi and Wesnousky (in prep) develop revised accounting that can be applied in the same approach as per-step probability decrements. Preliminary results indicate that half of historical ruptures have one end at a geologically defined feature such as a fault end or stepover, and roughly 25% are bounded on both ends.

(b) Empirical data also indicate that the probability of jumping at a stepover is a function of distance. Shaw and Dieterich (2007) propose an exponential functional form $P_{jump} = e^{-r/r0}$ that decays with distance, with a decay length scale of approximately r_0 = 3 km. The scale length was estimated from a model of self-organizing strike-slip faults and also separately from a fit to data from Wesnousky (2006). The distance-derived probabilities of jumping will be multiplied as in the constant probability decrement case described in 1(a) and the differences compared.

2. Requiring some development:

(a) Calculate Coulomb-based stress transfer considering the separation distance and section geometry at fault-to-fault jumps, then scale the result to a probability of jumping. This approach explicitly preserves the geometry and sense of motion of the intersection or nearest approach of faults. Coulomb favorability may be scaled by the favorability of a simpler intersection such as adjoining strike-slip sections. The approach of considering only sections nearest or most influential at the fault-to-fault junction may simplify the Coulomb estimate by removing averaging over all section interactions. A method of converting scaled Coulomb values to probabilities will be required. The probability scale will require a small positive value even for some negative Coulomb changes for ruptures not considered impossible. Dynamic models indicate that ruptures sometimes cross intersections that static stress interactions would not favor.

(b) Coulomb interactions have a more severe distance dependence that is empirically observed. This can make them very sensitive to sometimes poorly known distances of

closest approach. If that proves to be the case, the pairwise interactions of sections may be approximated from the geometry of their rake vectors and the less severe distance relationship of 1(b). A relationship between geometric interaction and probability would be proposed, as in 2(a).

3. Adjustments from dynamics:

(a) Dynamic models indicate that ruptures are more likely to cross stepovers where they meet steep slip gradients at the fault step. For these intersections probabilities can be adjusted according to the geologic knowledge available. Slip gradient criteria for favorability have been given by Elliott et al. (2009).

(b) Stepovers are more likely to be jumped when the transition is compressional than extensional. An adjustment may be proposed for these cases.

(c) Evaluate available dynamic models to see if adjustments are warranted to the distance relationship developed in 1(b).

References

Elliott, A.J., J. F. Dolan, and D. D. Oglesby (2009). Evidence from coseismic slip gradients for dynamic control on rupture propagation and arrest through stepovers, Journal of Geophysical Research 114, B02312, doi: 10.1029/2008JG005969.

Field, E. H. and M. T. Page, Estimating earthquake-rupture rates on a fault or fault system, Bulletin of the Seismological Society of America, Vol. 101, No. 1, pp. 79–92, February 2011, doi: 10.1785/0120100004.

Oglesby, D. (2008), Rupture termination and jump on parallel offset faults, Bull. Seismol. Soc. Am., 98(1), 440– 447, doi:10.1785/0120070163.

Shaw, B. E. and J. H. Dieterich (2007). Probabilities for jumping fault segment stepovers, Geophysical Research Letters, 34, L01307, doi:10.1029/2006GL027980

Wesnousky, S. G. (2006), Predicting the endpoints of earthquake ruptures, Nature, 444(7117), 358– 360, doi:10.1038/nature05275.

Wesnousky, S. G. (2008). Displacement and geometrical characteristics of earthquake surface ruptures: Issues and implications for seismic hazard analysis and the process of earthquake rupture, Bull. Seismol. Soc. Am. 98, no. 4, 1609–1632.

Wesnousky, S. G. and G. P. Biasi (2011). The Length to Which an Earthquake will go to Rupture: an Observational Note, Bulletin of the Seismological Society of America, 101, 1948-1950, doi: 10.1785/0120110013.

Task R7 - Reassess Historical Earthquake Catalog

Task Leader: Tom Parsons (USGS)

This task is intended to reduce uncertainty involving fault assignments in the historic, intensity-based earthquake catalog. There are three issues to consider addressing here: (1) uncertainty on magnitude and location in the historic (1850-1906) earthquake catalog may not be fully represented, as evidenced by comparison with values calculated on parts of the catalog using the method of *Bakun and Wentwort*h [1997]. (2) Potential improvements to the Bakun-Wentworth method may be implementable. (3) Development of PDF's on historic earthquake magnitude and location uncertainties that can be used to fully evaluate long-term earthquake rate changes that have informed empirical probability change models.

A proposed initial task is assembly of available intensity observations into a uniform database (the Global Earthquake Model (GEM) project has offered to host these data free of charge), and then systematic application of the Bakun-Wentworth codes (perhaps with updated attenuation relations if warranted). The proposalBakun-Wentworth method used a grid search to produce contours of confidence on intensity centers that identify most likely epicenters depending on magnitude. This information is used to manually assign events to faults. There can be a degree of ambiguity related to the point process location of large earthquakes because large rupture areas are reduced to a point.

The second part of the proposed task is to modify the Bakun-Wentworth code to invert for solutions using the 3D California fault model. Rupture patches can be systematically added within the 3D fault model, with synthetic intensity values superposed using an attenuation relation. As more patches are added, the magnitude of the rupture grows, and larger intensities are produced from a larger region. Some limited set of fault areas will be most consistent with the observed spatial pattern of intensities. More realistic intensity patterns would be produced instead of the currently used circular intensity patterns. The resulting output would then be best-fit magnitude and fault-area assignments that lack potential interpretation bias that is necessary in current practice.

References

Bakun , W. H., and C. M. Wentworth (1997). Estimating earthquake location and magnitude from seismic intensity data, *Bull. Seismol. Soc. Am. 87*, 1502-1521.

Task R8 - Reevaluate Earthquake Catalog

Task Leader: Andy Michael (USGS)

The historic and instrumental earthquake catalog is one UCERF3's primary constraints on earthquake rates and therefore probabilities and is thus worthy of reevaluation even after the

extensive work done during the UCERF2 process [*Felzer*, 2008a; b; *Felzer and Cao*, 2008]. The general process is to produce earthquake catalogs and then calculate earthquake rates as a function of magnitude by either declustering the catalog to remove foreshocks and aftershocks or by fitting a distribution that includes clustering to the complete data set. Updating the catalog and determining rates with a wide variety of approaches may help us understand the why there was a discrepancy between the observed earthquake rates and those predicted by the UCERF2 model and how much of a magnitude offset in the historic data would be required to resolve this discrepancy.

The first step is to compile an updated catalog to include any new analysis since the previous catalog was compiled. This involves a number of issues:

1. A new method for determining ML is being implemented by both the Northern and Southern California Seismic Networks (Urhammer et al., in prep.). This new method will improve the consistency of magnitudes across California but is currently creating temporal inconsistencies in the catalogs because while the method is being used for new events they have not completed the work necessary to recalculate older MLs for consistency. This has been shown to cause problems with determinations of a variety of key parameters such as b-value and the magnitude of completeness [*Tormann et al*., 2010] with a possible solution being proposed by Agnew [2010] with additional ideas in Tormann and Wiemer [2010]. Fortunately for the UCERF3 effort this makes the most difference in magnitudes for events below magnitude 4 and that is below the size used in the UCERF2 catalog. By keeping the minimum magnitude in the UCERF3 catalog at M=4 we maintain catalog completeness at the edges of the catalog and during aftershock sequences. Thus the primary problem would be when determining the b-value. That was done during UCERF2 using smaller events. However, Hutton et al. (2010) confirmed that b-value in southern California is ~1, as previously determined by UCERF2 for the entire state. To avoid the problems caused by these new magnitudes, we propose to compute the b-value for the most recent time period, which uses the new magnitude scheme, in order to determine whether or not the value has changed.

This new formulation started being used in the NCSN on 23 June 2009 but, except for rare events that were reprocessed for other reasons, the NCSN has not yet started reprocessing the existing catalog before that date.

2. The ongoing Historical Earthquake Re-analysis Project at U.C. Berkeley may have produced new magnitudes for events since 1951. At the time of the UCERF2 report this effort was not ready for inclusion in the catalog and we are checking with them to see how far they have gotten on finishing the new analyses.

3. When updating the catalog, we will examine the merged ANSS database for consistency with the NCSN and SCSN catalogs because during the UCERF2 process some differences were uncovered due to events that had been updates in the SCSN catalog that had not been successfully included into the ANSS database.

4. We will examine the final catalog to make sure that nuclear explosion tests have been properly removed. Again, this was an issue when the catalog was developed for the UCERF2 effort.

A number of topics will be considered that result in epistemic uncertainty in the earthquake rates:

A) Different approaches to evaluating historical records to obtain intensity values. Due to the extraordinary effort required to revisit the historical records, we will evaluate the sensitivity of the earthquake rates to possible systematic uncertainties in the intensities and the resulting historical catalogs. These changes would affect all of the historical earthquake catalogs.

B Different historical earthquake catalogs. As was done for the UCERF2 effort (Felzer, 2008b) the effect of using magnitudes determined by different approaches to analyzing intensity data [*Bakun*, 1999; 2000; 2006; *Bakun and Wentworth*, 1997; *Toppozada et al*., 2002] on the resulting earthquake rates will be considered.

C) Different declustering techniques. In the UCERF2 effort the Gardner and Knopoff declustering method [1974] was used as has been the practice in the USGS National Seismic Hazard Maps. However, it is recognized that different declustering approaches can affect results and so we will experiment with a variety of methods in order to determine the rates.

D) Rather than decluster the catalog, rates could be determined by fitting models, which include both a background rate and clustering to the data. However, this approach is prone to parameter tradeoffs between the background rate and the clustered components and we will explore these methods and their tradeoffs.

E) The time-varying magnitude of completeness is a key factor in determining rates and we will examine different methods for determining the magnitude of completeness. We will consider how poor completeness in less active areas of the state affects rate estimates in the more active areas.

F) Determining the rates associated with specific faults or background zones requires associating events with those sources. We will explore the effect of using different methods of associating earthquakes with faults and/or background zones.

G) Finally, we will explore whether allowing variations in the b-value of the Gutenberg-Richter relationship makes a significant improvement in fits to spatially varying rates and how such variations may impact the rates of M>=6.5 earthquakes.

References:

Agnew, D. C. (2010), Comment on “Changes of Reporting Rates in the Southern Californian Earthquake Catalog, Introduced by a New Definition of ML” by Thessa Tormann, Stefan Wiemer, and Egill Hauksson, *Bull. Seism. Soc. Am.*, in press.

Bakun, W. H. (1999), Seismic activity of the San Francisco Bay Region, *Bull. Seism. Soc. Am.*, *89*, 764-784.

Bakun, W. H. (2000), Seismicity of California's North Coast, *Bull. Seism. Soc. Am.*, *90*, 797-812.

Bakun, W. H. (2006), Estimating Locations and Magnitudes of Earthquakes in Southern California from Modified Mercalli Intensities, *Bull. Seism. Soc. Am.*, *96*(4A), 1278-1295.

Bakun, W. H., and C. M. Wentworth (1997), Estimating earthquake location and magnitude from seismic intensity data, *Bull. Seism. Soc. Am.*, *87*, 1502-1521.

Felzer, K. R. (2008a), Calculating California seismicity rates, Appendix I in The Uniform California Earthquake Rupture Foreast, version 2 (UCERF 2), *Open-File Report Rep. 2007-1437I* 42 pp, U.S. Geological Survey.

Felzer, K. R. (2008b), Empirical estimation of regional time variation in seismicity rates, Appendix M in The Uniform California Earthquake Rupture Forecast, version 2 (UCERF 2), *Open-File Report Rep. 2007-1437M* 19 pp, U.S. Geological Survey.

Felzer, K. R., and T. Cao (2008), WGCEP Historical California earthquake catalog, Appendix H in The Uniform California Earthquake Rupture Forecast, version 2 (UCERF 2), *Open-File Report Rep. 2007-1437H*, 127 pp, U.S. Geological Survey.

Gardner, J. K., and L. Knopoff (1974), Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian?, *Bull. Seism. Soc. Am.*, *64*, 1363-1367.

Toppozada, T. R., D. M. Branum, M. S. Reichle, and C. L. Hallstrom (2002), San Andreas fault zone, California: M ≥ 5.5 earthquake history, *Bull. Seism. Soc. Am.*, *92*, 2555-2601.

Tormann, T., and S. Wiemer (2010), Reply to "Comment on 'Changes of Reporting Rates in the Southern Californian Earthquake Catalog, Introduced by a New Definition of ML' by Thessa Tormann, Stefan Wiemer, and Egill Hauksson“ by Duncan Agnew, *Bull. Seism. Soc. Am.*, in press.

Tormann, T., S. Wiemer, and E. Hauksson (2010), Changes of Reporting Rates in the Southern California Earthquake Catalog, Introduced by a New Definition of ML, *Bull. Seism. Soc. Am.*, *100*(4), 1733-1742.

Task R9: Smoothed Seismicity Model

Task Leader: Karen Felzer (USGS)

Introduction

We propose to produce a smoothed seismicity map based on the algorithm of *Helmstetter et al.* (2007) and *Werner et al.* (2011). This algorithm was chosen because it is currently performing best in the prospective 5 year CSEP earthquake forecasting test. We also performed a retrospective test to see how well M≥6 earthquakes over the 50 year period from 1957-2006 could be forecasted by smoothing the 1850-1957 catalog with the *Helmstetter et al.* (2007) algorithm and the National Hazard Map (NHM)/UCERF2 algorithm, respectively. We found that the *Helmstetter* method produced a log-likelihood gain that was 30% higher than that obtained using the NHM/UCERF2 method, with the caveat that we only compared the automated results obtained from both methods, and, at this point in implementation, only smoothed M≥4 earthquakes from the UCERF2 California catalog. The NHM/UCERF2 method also employs secondary adjustments to elongate and narrow the kernel around certain faults, which is not needed with the *Helmstetter* method, as detailed below. The forecasting map produced by *Werner et al.* (2011) for California is given below, and compared with the smoothed seismicity map that was used for the 2008 National Hazard Maps.

Helmstetter (2007) algorithm may be implemented using a power law or Gaussian kernel. The National Hazard Maps have traditionally used a Gaussian kernel, so we plan to use this kernel with the *Helmstetter* technique to preserve consistency. A key innovation of *Helmstetter* (2007) is that while the smoothing constant *d* is set to a constant, 50 km, in the National Hazard Maps, in the *Helmstetter* method *d* is the distance to the nth closest earthquake. This means that *d* will be smaller when earthquakes are dense and broader where earthquakes are further apart. In particular where seismicity is dense and linearly aligned the kernels are focused and the lineation of the hazard is clearly visible even though the kernel around each individual earthquake is circular.

Based on the results of *Werner et al*. (2011) we will use seismicity down to M 2 for the portion of the catalog that is primarily complete to this magnitude (>1984), and use seismicity down to M 4 from 1932-1984. We are still working, however, on what declustering algorithm to use for the catalog. The National Hazard Maps traditionally use the *Gardner and Knopoff* (1974) algorithm, whereas *Helmstetter* (2007*)* used *Reasenberg* (1985), but did not test whether other algorithms produce better results. Clearly whatever declustering algorithm is used should not remove all aftershocks unless major aftershock zones are later replaced by additional modeling; we note that the 1992 Landers earthquake, 1999 Hector Mine earthquake, and 1994 Northridge earthquake, for example, are all significant earthquakes that occurred within active aftershock zones. The hazard posed by aftershocks was also highlighted by the recent devastating aftershock in Christchurch, New Zealand.

If we can gain the assistance of experts in strong ground motion attenuation relationships we hope to test our final results against the observed distribution of precarious rocks.

We also wish to emphasize that because the smoothed seismicity map is based on a relatively short record of seismicity (the majority of events come from after the seismic network started in Southern California in 1932) it indicates where earthquakes are nucleating now, but may not correspond to long term seismicity patterns, nor indicate where aftershocks of future events are more likely to occur. The map is thus designed to be used for general seismicity forecasts on the order of years and for the generation of the spontaneous background seismicity for ETAS aftershock modeling. When individual aftershock sequences are modeled, however, we hypothesize that it is most accurate to use fault and deformation models to forecast the azimuthal distribution of the aftershocks, following *Powers and Jordan* (2010). This is because new aftershocks should be triggered preferentially where strain is highest, which may not correspond uniformly with where recent earthquakes have occurred.

References

Gardner, J. K. and L. Knopoff (1974), Is the sequence of earthquakes in southern California with aftershocks removed Poissonian?, *Bull Seis. Soc. Am.,* 64, 1363-1367.

Helmstetter, A., Y. Y. Kagan, and D. D. Jackson (2007), High-resolution time-independent grid-based forecast for M=5 earthquakes in California, *Seis. Res. Lett.,* 78, 78-86.

Ogata, Y., Space-time point-process models for earthquake occurrences (1998), *Ann. Stat.*, 50, 805-810.

Powers, P. M. and T. H. Jordan (2010), Distribution of seismicity across strike-slip faults in California, *J. Geophys. Res.*, 115, B05305, doi:1029/2008JB006234.

Reasenberg, P., Second-order moment of {C}entral {C}alifornia seismicity, 1969-1982 (1985), *J. Geophys. Res.,* 90, 5479-5495.

Wang, Q., D. D. Jackson, and J. Zhuang, (2010), Are spontaneous earthquakes stationary in California?, *J. Geophys. Res.,* 115, B08310, doi:1029/2009JB007031.

Werner, M. J., A. Helmstetter, D. D. Jackson and Y. Y. Kagan, (2011), High resolution long- and short –term forecasts for California, *Bull. Seis. Soc. Am.,* 101, 1630-1648.

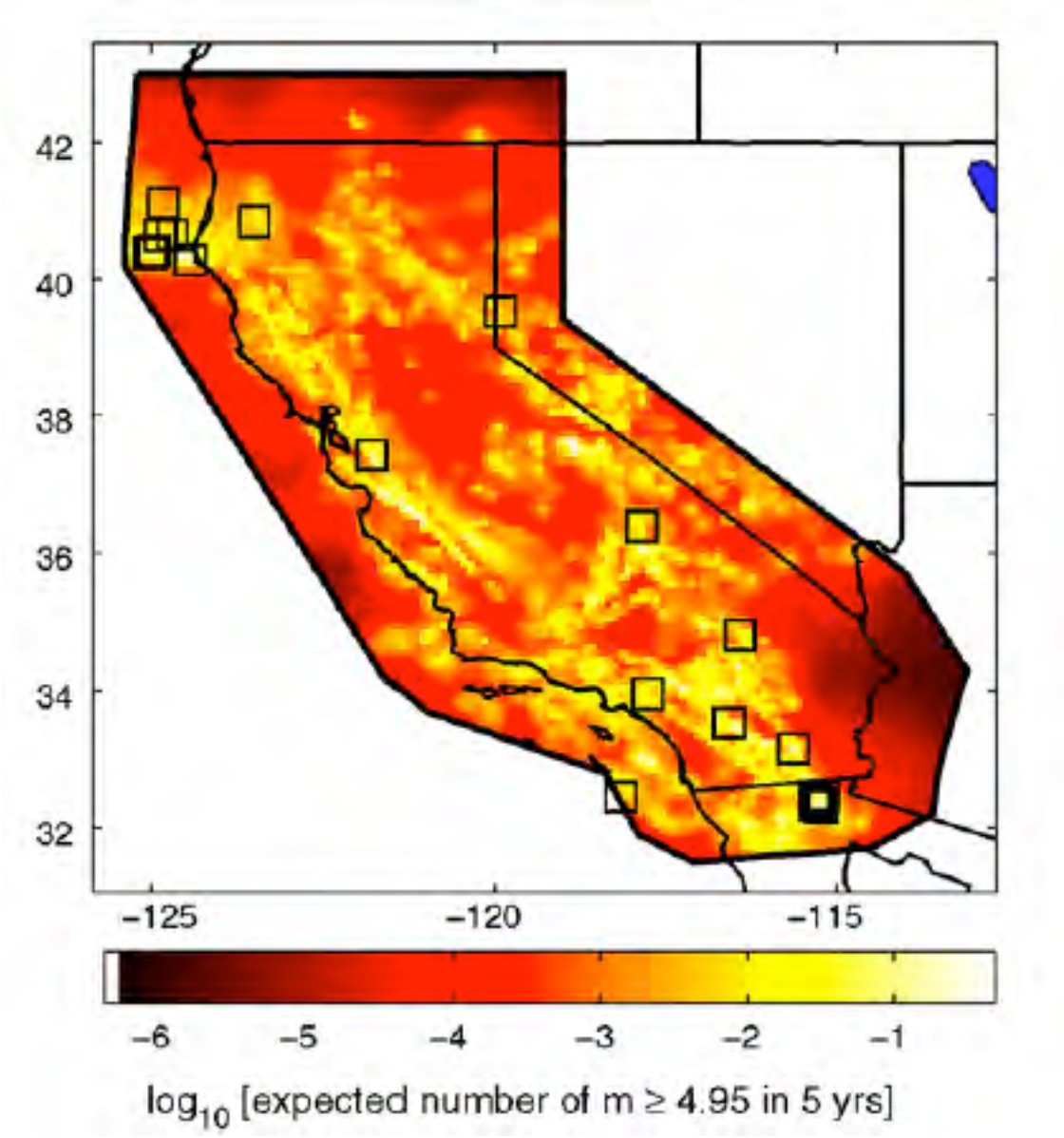


Figure 1: Figure 7 from *Werner et al.* (2011) showing their smoothed seismicity forecast for 2012-2016. Note that the color scale is logarithmic.

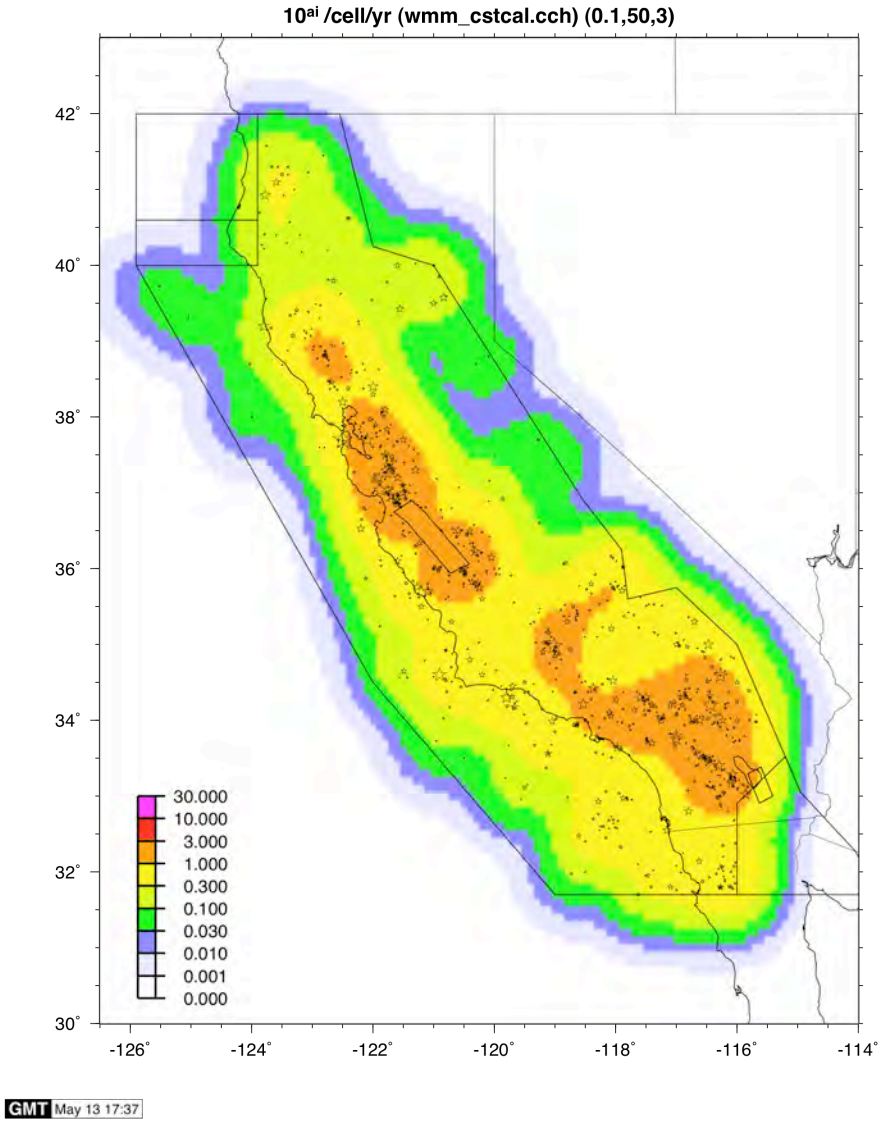


Figure 2: Smoothed seismicity map used for the 2008 National Hazard Maps, based on 1850-2006 catalog data. Figure courtesy of Chuck Mueller.

Task R11: Focal Mechanisms of Off-Fault Seismicity

Task Leader: David D. Jackson (UCLA)

Introduction

Kagan, Jackson, and Rong [2007] employed a smoothed seismicity model to forecast earthquakes in California. The method is further described in Kagan and Jackson [1994].The model is based on evaluating, at each map point, a weighted average of the number of earthquakes per unit time in the vicinity. Weights depend on the magnitude of the earthquakes and their distance from the map point. Their forecast included estimates of the moment tensors of future earthquakes, constructed by weighted averages, with the same weights, of the moment tensors of those nearby earthquakes. We would apply the same technique for all of California to estimate focal mechanisms, and their uncertainties, for all California.

Model Formulation

Our spatial smoothing kernels have the form

$$f(r) = A*(m - m_t)/\sqrt{r^2 + d^2}$$

Where A is a normization constant, r is the distance from a map point to an earthquake, m is the magnitude of that earthquake, m_t is the lower magnitude threshold for the catalog, and d is a constant, related to the uncertainty of location accuracy. For each earthquake, we normalize the moment tensor; then for each map point, we sum the moment tensors times the weight implied by the equation above. By normalizing the moment tensors of each earthquake first, we assure a magnitude weighting given by the equation above, which depends only mildly on magnitude. The variance of the focal mechanism parameters at a map point is determined approximately from the same weighted sum of the variances of the known earthquake focal mechanism. However, the statistics of focal mechanism parameters is not Gaussian, so the error estimates are a bit complicated; details are given in Kagan et al. [2007] and references therein.

Input Data

The only input data needed are locations and focal mechanisms of earthquakes within about 100 km of the region of interest. We'll use a uniform lower magnitude threshold, determined by the smallest magnitude for which all events have measured focal mechanisms. We will not distinguish between on-fault and off-fault earthquakes; all are informative about the focal mechanisms, and there is no danger of double counting because we are only calculating the normalized focal mechanism.

Optional extension of the concept

It is relatively straightforward to include fault orientations and slip directions along with earthquake focal mechanisms as input data. We could convert earthquake occurrence to earthquake rate by dividing by the temporal length of the catalog, subdivide faults into sections, compute tensor moment rates for each section, and compute weighted averages in the same way

we do for earthquakes. Some experimentation would be required, as the effective weight of each fault section would depend in a nonlinear way on its length.

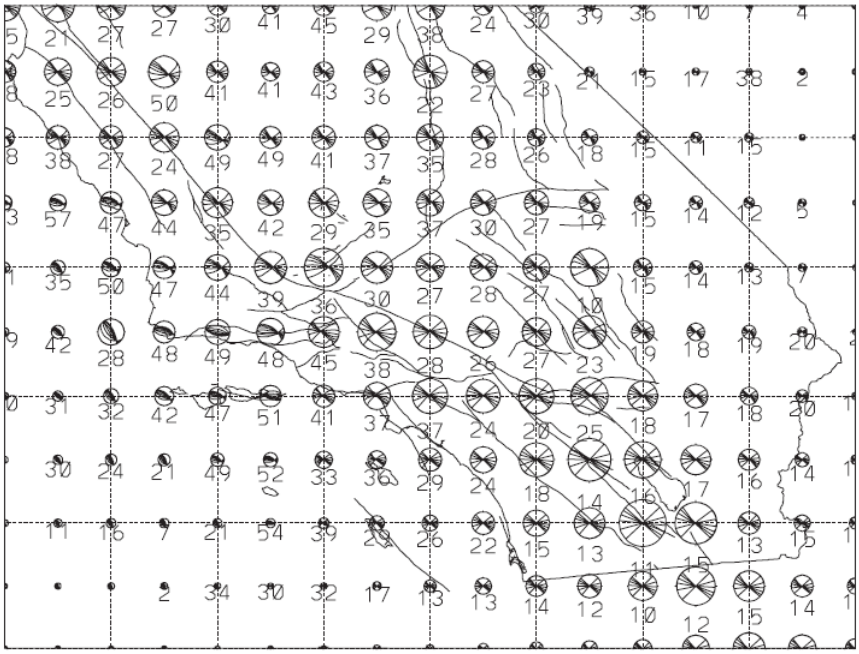


Figure 18. Long-term forecast diagrams of earthquake focal mechanisms in southern California. Lower hemisphere diagrams of focal spheres are shown. Size of the focal mechanism diagram is proportional to forecasted rate of occurrence (see figure 1). Stripes in beach balls are concentrated toward the assumed earthquake fault plane. The numbers below the diagrams of earthquake focal mechanisms correspond to a standard deviation of a weighted 3-D rotation angle. We first calculate the average seismic moment tensor and then compute the rotation of earthquake focal mechanisms with regard to the average double-couple source. Therefore the average rotation angle shows degree of tectonic complexity. Points without beach ball diagrams denote places for which data are inadequate to forecast focal mechanism. From Kagan et al., [2007]. The plot is displayed at URL http://moho.ess.ucla.edu/~kagan/s_cal_fps.ps.

References

Kagan, Y. Y., D. D. Jackson, and Y. F. Rong, 2007. A testable five-year forecast of moderate and large earthquakes in southern California based on smoothed seismicity, Seism. Res. Lett., 78(1), 94-98.

Kagan, Y. Y., and D. D. Jackson (1994). Long-term probabilistic forecasting of earthquakes. *Journal of Geophysical Research* 99, 13,685–13,700.

Task R12 – Distribution of Repeated Slip at a Site on a Fault

Task Leader: Ray Weldon (UO)

Statement of the Problem: The variability of slip at a site on a fault from earthquake to earthquake is a critical but hotly debated parameter. The characteristic earthquake model posits that repeated displacements are very similar, whereas other recurrence models produce less regular repetition of displacement.

Proposed Solution: We intend to collect a global dataset of both repeated historic ruptures and studies of prehistoric ruptures to assess how repeatable slip at a point on a fault is, and if possible understand what controls the variability if it varies from fault to fault. This effort will build on a number of existing summaries and will be a component of a larger effort to collect and interpret information of historical ruptures to assess fault-to-fault jumps, and distribution of slip along strike in ruptures and other parameters we seek to better understand.

Task R13 – Evaluate Physics Based Earthquake Simulators (for rate estimates)

Task Leader: Ned Field (USGS)

Physics-based earthquake simulators represent a viable way of developing an earthquake rate model (e.g., run them for a very long time and look at the rate of each rupture). These are particularly appealing in that they naturally relax segmentation and include multi-fault ruptures. However, the question remains whether these models reliably capture the relevant earthquake physics, and whether their usefulness is diminished by producing a wide range of behaviors among the different simulators (or for alternative parameter settings within a given simulator). At the very least physics-based simulators will be useful exploratory tools, and we plan to use them as such. Fortunately SCEC has a formal working group dedicated to the development, verification, and evaluation of these models, and we are actively working with that group in order to utilize simulators to the maximum extent possible. This group is being led by Terry Tullis, and the leaders of the groups developing different simulators that might be applicable statewide include:

- John Rundle (*Virtual California*; Rundle et al, 2006)
- Steve Ward (*ALLCAL*; Ward, 2000)
- James Dieterich (*RSQSim*; Dieterich and Richards-Dinger, 2010)
- Fred Pollitz (*VISCO-SIM*)

We would first want to convince ourselves that any given simulator is able to reliably reproduce the following (each of which is either imposed, or to some extent well constrained):

- long-term slip rates.
- paleoseismic event rates where available.
- magnitude-frequency distribution (MFD) for entire region.
- magnitude-area and/or slip-length scaling relationships.
- fault-to-fault rupture jumping distances (consistent with observations?).
- Omori decay, at least for small events.

Once a simulator has been "verified" in terms of consistency with the above, we might then want to examine any of the following:

- MFDs at points on faults (Characteristic or Gutenberg Richter?).
- MFD for entire "faults" (assuming faults can be meaningfully isolated and defined).
- Is one 1500-yr sample on a fault (like our SSAF paleo record) indicative of long term behavior?
- Can we run simulators long enough to constrain the long-term rate of "every possible" rupture (at some discretization level)?
- Recurrence-interval statistics at points on a fault, for faults, and for regions.
- Magnitude dependence of recurrence-interval statistics.
- Elastic-rebound predictability (time and/or slip predictable?).
- Sensitivity of large-event statistics to changes in cell size (e.g., in going from ~4 km cells to ~1 km cells).
- Multi-fault rupture behavior (what influences such occurrences).
- Average slip distribution along strike (e.g., is the average over many repeats of the same event broadly tapered (e.g., sqrt(sin) as used in UCERF2) or more flat in the middle? Multi-rainbows for multi-fault ruptures? What's the variability about this average?.
- Does slip continue to penetrate deeper (below the depth of micro seismicity) for longer and longer ruptures?
- The rate of small earthquakes on faults (consistent on the large faults that seem quiet, like parts of the San Andreas?)
- Spatiotemporal clustering, especially for larger events (does ETAS apply at largest magnitudes?; is the fraction of "aftershocks" magnitude independent?).
- Longer-term time dependencies (like implied by the "empirical" model)?
- How do we glean applicable statistical rules from simulators for the purposes of hazard assessment (e.g., assuming a simulator is perfectly correct, how can we use it)?
- Robustness of all of the above with respect to different simulators and alternative parameter settings within a simulator (i.e., what are the epistemic uncertainties).

Further analysis of simulators is awaiting their implementation of the UCERF3 deformation model (at least the preferred logic-tree branch), which is not yet available. Furthermore, additional physical parameters such as fault strength need to be added and tuned for their models, which will take an unknown amount of time. Therefore, much of this work is beyond the scope of UCERF3, but we will continue this activity as time and resources permit. To this end, the analysis tools being developed for the inversions results will be equally applicable to simulator results.

Task R15 – Cascadia Subduction Zone

Task Leader: Art Frankel (USGS)

Plans for Updating the Characterization of the Cascadia Subduction Zone for the National Seismic Hazard Maps and UCERF:

1. We will evaluate the recent results of Goldfinger et al. (2010) from turbidite data that show a recurrence time of about 230 years for M8 and larger earthquakes along the southern portion of the Cascadia subduction zone (CSZ). We are planning a small focused meeting of experts for Fall 2010 to assess the evidence for this higher rate and compare the turbidite results with onshore data, especially from Bradley Lake, Oregon.
2. Based on the results of this meeting, we will develop new magnitude-frequency distributions for Cascadia great earthquakes. These distributions may differ between the northern and southern portions of the CSZ. We will also assess whether multiple distributions should be used to quantify the epistemic uncertainty in recurrence model for any portion of the CSZ.
3. We will evaluate the possibility of temporal clustering of CSZ earthquakes that has been proposed by Goldfinger and Wong.
4. We will evaluate various models for the location of the eastern edge of the rupture zones for great earthquakes on the CSZ. Some scientists have suggested that the up dip limit of tremor events (ETS) may signify the down dip edge of the locked zone. This edge is similar to the geometries that were given substantial weight in UCERF2 and the 2002 and 2008 NSHMs. We will also evaluate recent work using GPS, tide gauge, and microfossil data that provides constraints on the location of the locked zone.
5. We will update the location of the plate interface based on the latest compilation by McCrory.
6. We will reassess our time dependent model for CSZ, which is based on the time since the 1700 earthquake. It remains to be seen how this can be combined with observations of a shorter recurrence time in the southern CSZ.

7. We will hold a regional Pacific Northwest workshop for the update of the NSHM in late 2011 or early 2012. The CSZ issues noted above will be discussed at this workshop, so this workshop will also be important for UCERF 3.

Task P1 and P2 – Address Empirical Model and ETAS explains Empirical Model?

Task Leader: Karen Felzer (USGS)

The UCERF2 report demonstrated that the instrumental earthquake catalog for the state of California (1932-present) has a lower average seismicity rate than the historic catalog (1850-1932). This rate decrease can be documented independently for the north coast, the San Francisco Bay Area, the central coast, and the Los Angeles area. Specifically, UCERF2 found that seismicity rates in the 1906-2006 catalog were ~79% of 1850-2006 rates for the full state, and 46% and 66% of the 1850-2006 rates for the San Francisco Bay Area and Los Angeles areas, respectively. A constrast in seismicity rates in the San Francisco Bay area between 1850-1906 and since the late 1920s has long been noted (*Jaume and Sykes,* 1996). In Southern California it has also been repeatedly recognized that the rate of seismicity along the San Andreas fault is low (*Allen et al.*, 1965).

The observation of a low instrumental seismicity rate in comparison with the historic catalog could be caused by errors in the historic catalog, particularly in the estimation of magnitude error. Because earthquakes follow a power law distribution of magnitudes, normally distributed magnitude error causes more magnitudes to be inflated than deflated, and causes measured seismicity rates to be high. It is doubtful that the entire rate difference can be explained by error, however. It is clear, for example, that the only recorded M>7.5 earthquakes in California: Ft. Tejon, Owens Valley, and San Francisco, occurred in the historic era, suggesting that this was indeed a more active time. Independent work has also found that recorded seismicity rates are low in comparison to what might be expected from GPS measurements. *Ward* (1998), for example, found that the seismic moment release rate recorded in the 1850-1996 catalog was ~73%-86% of what geodetic measurements implied the long term rate should be. *Bird* (2009) likewise found that the seismicity rate of the 1977-2008 California catalog was about 76% of the expected long term average based on the compilation of geodetic, geologic, and stress direction data sets into a deformation model known as NeoKinema.

Variation of the seismicity rate with time is normally expected as a consequence of simple random variability and aftershock triggering. The majority of earthquakes occur as aftershocks (*Gardner and Knopoff,* 1974*)* and aftershock triggering can cause clustering to occur over a range of time scales. This variability can be estimated with the stochastic ETAS model (*Ogata*, 1988). The decreased seismicity rate averaged over the whole state can be recreated with an ETAS model based on the implementation of *Hardebeck et. al.* (2008). A detailed spatial comparison of the instrumental and historic catalog, however, shows that the decreased seismicity rate is not diffused evenly over the state but rather focused on the San Andreas system (*Felzer and Brodsky*, 2005), and that in some places the quiescence is acute, as has long been noted for the Southern San Andreas (*Allen et al.* 1965).

Comparison of normalized spatial seismicity rates calculated from smoothing the M≥4 1932-2010 catalog (see Task R9) with a normalized map of the expected long term seismicity nucleation rates from the Neokinema deformation model (*Bird*, 2009) (Figure 1) confirms that parts of the San Andreas fault, in particular north of San Francisco and in the Carrizo Plain, are severely lacking in seismicity, in some places by up to two orders of magnitude. The same conclusion can be drawn from comparing catalog seismicity within 5 km of the fault with UCERF 2 determined slip rates. Several other faults and regions are also low, including the Hunter Mountain, Death Valley, and Garlock Fault Zones, Ventura County, and Cascadia.

Traditionally the low seismicity rate around the San Francisco Bay area has been attributed to static stress shadowing by the 1906 earthquake, however the stress shadow model prescribes a gradual return to previous seismicity rates which has not been observed (*Reasenberg et al.,* 2003). Nor can the stress shadow model explain why the decrease in seismicity rates initiated in the 1920s rather than in 1906 (*Jaume and Sykes*, 1996). In addition, Figure 1 does not show a strikingly low seismicity rate throughout the Bay Area but primarily only on the San Andreas. The East Bay, in fact, seems to be experiencing a slightly elevated seismicity rate. Comparison of the historic catalog with Neokinema suggests that from 1850-1906 the East Bay experienced a much higher rate than its long term average, while seismicity near the San Francisco portion of the San Andreas was close to its long term rate. The ensuing relative quiescence thus appears to be a combination effect of decay towards long term seismicity rates in the East Bay with a drop below normal rates on the San Andreas proper, as was noted by *Reasenberg et al.* (2003).

It has likewise been proposed that a stress shadow from the 1857 Ft. Tejon earthquake caused seismic quiescence in Southern California, but this stress shadow should have resolved decades ago (*Harris and Simpson,* 1996). Alternatively the characteristic earthquake hypothesis has been supported (*Schwartz and Coppersmith*, 1984). This hypothesis holds that the seismicity rate on the San Andreas fault has been constant but applies only to large earthquakes; small to moderate earthquakes, which make up the vast bulk of seismicity, have a permanently very low rate of occurrence on the fault proper, out of proportion to what would be expected from the Gutenberg-Richter magnitude frequency relationship (*Gutenberg and Richter,* 1944). As a Gutenberg-Richter magnitude frequency distribution has been deduced for M≥7 earthquakes on the Southern San Andreas, from paleoseismic evidence (*Biasi and Weldon,* 2009), the border between ``large'' and ``moderate/small'' earthquakes in this model may reasonably be placed at M 7.

There are several observational challenges to the characteristic hypothesis. First, a permanent lack of M<7 earthquakes on the San Andreas requires a bulge of such earthquakes off of the fault, which has not been observed (*Page and Felzer*, in preparation). Second, observational evidence indicates that like the rate of smaller earthquakes, the rate of M>7 earthquakes on the San Andreas is also low at present, indicating that the two rates may indeed track each other. Since 1932, 6 M≥7 earthquakes have nucleated inside the California UCERF zone and south of the triple junction, but none of these has been on the San Andreas. Assuming that the San Andreas relieves ~40% of the seismic budget over the long term, there is only a 5% chance of observing so many off-fault nucleations if the

current rate of M>7 earthquakes on the fault is not depressed from the long term average. New work by *Felzer* (in preparation) using foreshock statistics demonstrates that if the characteristic threshold is M 7 small earthquakes cannot be depleted in areas where characteristic earthquakes nucleate by more than a factor of 2—3. This means that either characteristic earthquakes do not nucleate in regions where the small earthquake rate is severely depleted, or most of the depletion represents a real, temporary, seismicity rate decrease at all magnitudes.

If strong quiescence's really do occur, then these must be offset by strong clusters at other periods of time. Comparison of the 1932-2010 catalog with the NeoKinema model shows regions currently demonstrating much higher than average seismicity rates do exist. Not surprisingly, these areas are dominated by aftershock zones of significant earthquakes, such as Landers, Hector Mine, San Simeon, Northridge, and Coalinga. It is important to emphasize that these high rates represent seismicity rates averaged over the entire 1932-2010 period. In fact, the total area of the map that is characterized by much lower than expected seismicity rates is fairly balanced by the amount of area with much higher rates, although the total rate in California is low because the current low areas produce high seismicity rates long term. This observation strongly suggests that at some later point in time other areas of the map will become quiescent while the currently quiet parts of the San Andreas and Cascadia will experience 100 times higher than average levels of seismicity. In fact, evidence for the clustering of large earthquakes can be seen in the paleoseismic record of the San Andreas (*Weldon et al.,* 2004) and by the relative proximity of the historic 1812 and 1857 earthquakes on the Southern San Andreas.

An outstanding question is whether aftershock triggering by itself can explain such large swings in seismic activity rate. Our current ETAS model does not produce rate variations that are this large over time periods this long, but it contains many known simplifications, including a broad background kernel, a steady rate of background earthquakes that comprises 40% of the total seismicity rate, a minimum magnitude of 2.5, no modeled faults, and constant values for aftershock parameters. An improved ETAS model is being developed for UCERF 3, which may produce more variability. If we obtain an ETAS model that produces localized seismicity rate swings on the order observed, this would be a valuable tool to model what might happen in the future. If we cannot, we may observe that the current major quiet areas have persisted for years, and thus might be presumed to persist for years to come unless a triggering earthquake in the right location initiates an increase in the seismicity rate. Likewise areas with enhanced seismicity rate often tend to remain more active than average for years or decades, with occasional swings in their total rate of activity as significant aftershock sequences are initiated or decay.

Thus given our current state of knowledge we recommend the following for the short term and long term seismicity rate models:

1) Our work agrees with the conclusions of *Allen et al.* (1965) that the current seismicity catalog, in particular the instrumental portion, is too brief to accurately reflect long term earthquake rates. Some of the rates that we see today are far different than what we would expect over the long term. Thus we recommend that the long term model be produced using long term slip and deformation rates.

Where faults and slip rates are not known regional deformation rates might be used to estimate long term seismicity rates.

2) In contrast, the short term forecast should be based initially solely on rates from the smoothed seismicity catalog. The RELM tests have demonstrated that the current earthquake catalog, with its ongoing areas of activity and quiescence, is inertial enough that it appears to be more predictive than long term slip/deformation rates over periods at least as long as 5 years. Preliminary retrospective tests that we have performed indicate that the seismicity catalog might also be the best predictor for 10 , 30, and 50 year periods, although it would become critical to update the map for new aftershock activity, as described below.

3) As earthquakes continue to occur the short term forecast should be updated to reflect where aftershock triggering is expected. For the ETAS simulation that will be used to implement this forecast the background or ``independent'' earthquakes should be chosen from the short term seismicity rate forecast but the aftershock locations should be sampled from the *a* values on the long term seismicity map. This is consistent with observations that aftershocks often chart new territory whereas short term steady state background events, most of which are probably aftershocks of some unknown or very old event, are less likely to. This method should allow the short term map to evolve towards the long term map with time. An important test of this recommendation is that long term ETAS simulations should be run to verify that the expected long term seismicity map is recreated over a suitably long (5000 years? 10,000 years?) period of time.

References

Allen, C. R., P. St. Amand, C. F. Richter and J. M. Nordquist (1965), Relationship between seismicity and geologic structure in the Southern California Region, *Bull. Seis. Soc. Am.,* 55, 753-797.

Biasi, G. P. and R. J. Weldon (2009), Estimating the magnitude-frequency distribution for the southern San Andreas fault from paleoseismic data, *Seis. Res. Lett.*, 80, 381.

Bird, P. (2009), Long-term fault slip rates, distributed deformation rates, and forecast of seismicity in the western United States from joint fitting of community geologic, geodetic, and stress direction data set, *J. Geophys. Res.,* 114, B11, B11403, doi:10.1029,2009JB006317.

Felzer, K. R. and E. E. Brodsky (2005), Testing the stress shadow hypothesis, *J. Geophys. Res.,* 110, B05S09, doi:1029/2004JB003277.

Felzer (2011), Limits on the frequency of characteristic earthquakes from foreshock statistics, *in preparation*.

Gardner, J. K. and L. Knopoff, Is the sequence of earthquakes in southern California with aftershocks removed Poissonian?, *Bull Seis. Soc. Am.,* 64, 1363-1367.

Gutenberg, B. and C. F. Richter (1944), Frequency of earthquakes in California, *Bull. Seis. Soc. Am.,* 4, 185-188.

Hardebeck, J. L, K. R. Felzer and A. J. Michael (2008), Improved test results reveal that the accelerating moment release hypothesis is statistically insignificant, *J. Geophys. Res.*, 113, B08310, doi:10.1029/2007JB005410.

Harris, R. A. and R. W. Simpson (1996), In the shadow of 1857 – The effects of the great Ft. Tejon earthquake on subsequent earthquakes in southern California, *Geophys. Res. Lett.,* 23, 229-232.

Jaume, S. C. and Lynn R. Sykes (1996), Evoluation of moderate seismicity in the San Francisco Bay region, 1850 to 1993: Seismicity changes related to the occurrence of large and great earthquakes, doi:10.1029/95JB02393.

Michael, A.J., S.L. Ross, D.P. Schwartz, J.W. Hendley II, and Peter H. Stauffer (1999), Understanding earthquake hazards in the San Francisco Bay Region: Major quake likely to strike between 2000 and 2030, *USGS Fact Sheet* 152-99, 4pp.

Ogata, Y. (1988), Statistical models for earthquake occurrence and residual analysis for point processes*, J. Am. Stat. Assoc*, 83, 9-27.

Reasenberg, P. A., T. C. Hanks and W. H. Bakun (2003), An empirical model for earthquake probabilities in the San Francisco Bay Region, California, 2002-2031, *Bull. Seis. Soc. Am.,* 93, 1-13.

Schwartz, D. P. and K. J. Coppersmith (1984), Fault behavior and characteristic earthquakes; examples from the Wasatch and San Andreas fault zones, *J. Geophys. Res.*, 89, 5681-5698.

Ward, S. N. (1998), On the consistency of earthquake moment rates, geological fault data, and space geodetic strain: the United States, *Geophys. J. Int.*, 134, 172-186.

Weldon, R., K. Scharer, T. Fumal, and G. Biasi (2004), Wrightwood and the earthquake cycle: What a long recurrence record tells us about how faults work, *GSA Today*, 14, 4-10.

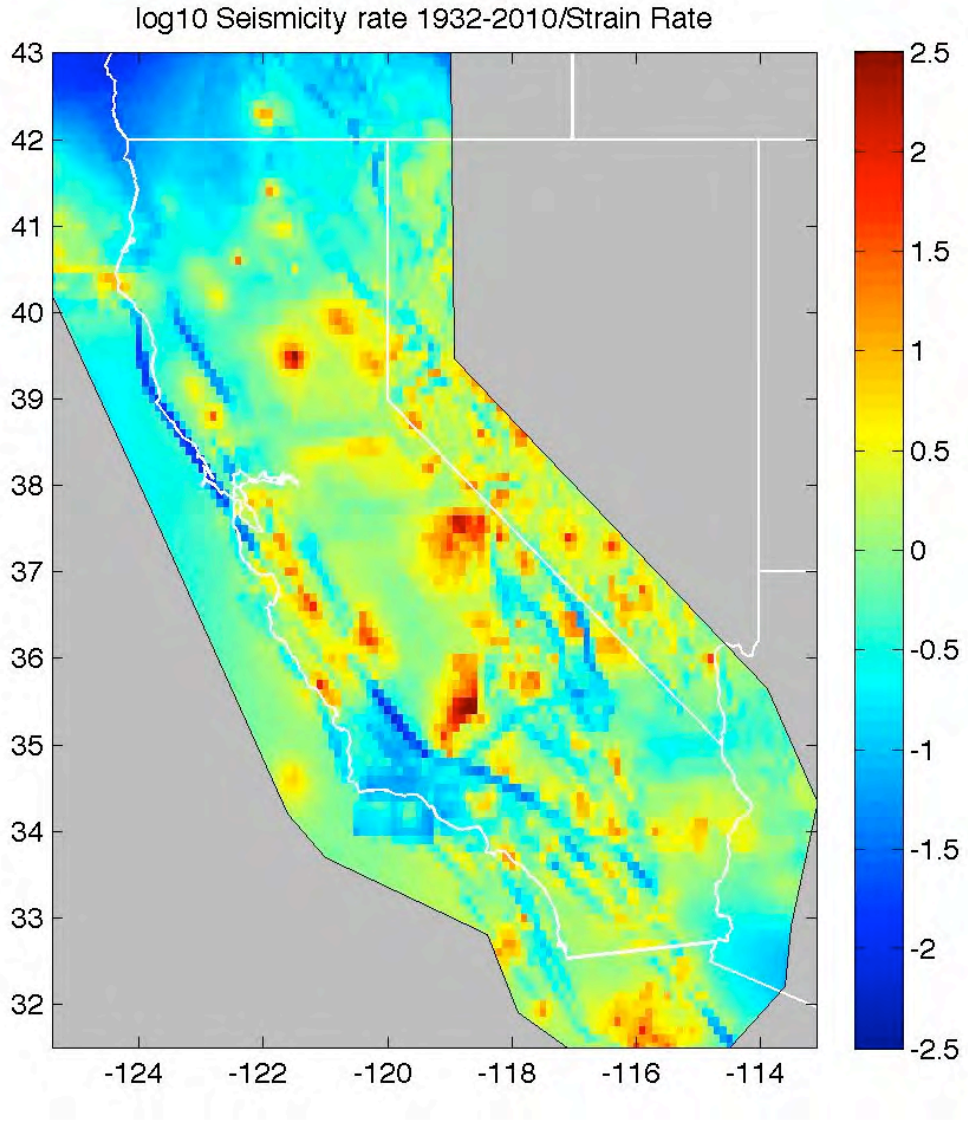


Figure 1: Spatially smoothed seismicity rates from the 1932-2010 catalog (M≥4) divided by NeoKinema long term deformation rates. The total seismicity rate on both maps was normalized before dividing, so that absolute long term rates on either map did not need to be assumed. Results are mapped on 0.1 by 0.1 degree bins and the color scale is logarithmic.

Task P3 (and P7) - Coulomb Stress Explains Empirical Model? (and Evaluate Static Stress Change Models)

Task Leader: Tom Parsons (USGS)

Uncertainty about interactions led UCERF2 to avoid the issue altogether, concluding that the uncertainties in the rate model were larger than the interactions. UCERF2 instead adopted an empirical correction based on seismicity rate changes. Since we know that interactions do occur, this decision is not actually as conservative as it seems, particularly given the large influence of the empirical model, its considerable uncertainties, and arbitrary weighting in UCERF2. These two tasks are aimed at revisiting interactions, which have drawbacks (parameters), and advantages (physics, strong presence in the literature).

This task will proceed in concert with the reevaluation of the empirical model, and the historical earthquake catalog. Once we know what the observed rate changes are, and where large earthquakes occurred, we can readily prepare a statewide stress change map. The extent to which one explains the other will answer task P3. Postseismic mantle relaxation is not as exotic as it was in 1999, and thus can be defensibly (based on a decade of geodesy) incorporated into these calculations. While this adds parameters, these choices do not reverse the sign of the stress change, but tend to amplify the elastic calculations. Ideally we would have a physical basis for application of the empirical model, and importantly, gain insight into to its application in areas of very large uncertainty.

Task P7 is asking whether Coulomb stress changes can be used along side empirical short-term forecasts like ETAS or STEP. Ideally one would get a spatial pattern of stress increase immediately after a main shock. In practice, this has not worked well in prospective tests. Perhaps the best route would be to implement Coulomb calculations in parallel, but offline, to develop a database to compare with empirical methods and observed earthquake occurrence during the operational phase of UCERF3.

Task P5 - Evaluate ETAS implementation

Task Leader: Felzer & Michael (USGS)

Here we present some initial work to test how well the spatial decay of the aftershocks of large mainshocks is recreated by the implementation of the ETAS model. The initial test cases are the 1992 M 7.3 Landers and 1994 M 6.7 Northridge earthquakes. In the ETAS model aftershock locations are distributed according to $(1/(r+\text{dmin}))^{-\text{distDecay}}$ where r is distance in kilometers from the closest point on the mainshock fault plane. We test three different sets of parameters: dmin=0.3, distDecay=1.7; dmin=0.5 distDecay = 2.0; and dmin=0.3, distDecay = 2.5. For each mainshock simulations performed with each of these parameter sets is compared against the data.

There are several serious challenges when working with aftershock data of large mainshocks, particularly in space. The first is that the aftershock catalogs tend to be very incomplete, especially in the first hours to days after the mainshocks, depending on mainshock size and location. The second is that the fault must be modeled as an extended source, but that source may be irregular, with bends and offsets both at the surface and at depth. Slip asperities may also occur at various locations and at various scales. To the extent to which fault and slip irregularities are buried it may be difficult to know their location and extent. This provides a challenge with respect to how to measure the distance between aftershocks and the nearest point on the fault ``plane''. It also ensures that there will be some degree of difference between the real data and ETAS, as the latter currently models the source as smooth slip on a simple plane or set of planes.

Because the real mainshock slip pattern is uncertain and complex, instead of measuring distances between aftershocks and a modeled fault plane we automatically generate ``fault points'' that are in the center of the densest areas of aftershock activity and measure distances between aftershocks and the closest one of these points. The optimal way to find these points, the optimal number, and the sensitivity of the results to the number of points generated is still under development, but the current preliminary method is as follows:

1) Sort all aftershocks occurring in the first week after the mainshock into 5 x 5 x 5 km bins.
2) Sort the bins from most to least densely populated.
3) Calculate a value *Nbin* which is the minimum of either the total area of the fault divided by 10 or the total number of populated bins divided by 4. This calculation is purely arbitrary, chosen because it tends to produce enough points to trace out most of the visually obvious fault trend. Fault area is calculated from the *Wells and Coppersmith* (1994) relationships.
4) Select the *Nbin* most densely populated bins.
5) Place a fault point at the median aftershock location (lat, lon, and depth) in each of these bins.

In the figures and analysis that follows distances are measured to these fault points, both for the real and simulated sequences to keep the analysis comparable.

The following figures provide the map view of the first week of the Landers aftershock sequence, with its associated fault points, and the corresponding map views and solved fault points for ETAS simulations done with the three different parameter sets. For the real mainshock it can be seen that fault points also outline the fault plane of the Big Bear aftershock. This is appropriate since Big Bear was also an active source of aftershocks, with an aftershock density along its fault plane similar to that of the Landers mainshock.

For Landers aftershocks we use a minimum magnitude of 3.0. This value was chosen because it provided enough data for the fault points to trace along most of the rupture, however it is below the completeness threshold of the aftershock catalog, which is known to be initially incomplete below M 4.7. The simulated data, on the other hand, is complete. Thus there is more data on the simulated plots, which will automatically make the cluster of aftershocks along the simulated fault trace appear wider. Figure 2 displays the cumulative probability distribution (CDF) of the distances between the real Landers earthquakes (first week of M≥3 events) and the fault points, compared with the corresponding CDFs of each of the simulations. In the near field, closer than ~2 km, aftershocks are more dense in the real data than in any of the simulations. This suggests than the parameter minDist needs to be decreased further. As we move away from the mainshock fault, however, the real aftershocks become more disperse, with their density decaying at a slower rate than any of the simulations. This indicates that distDecay should also be decreased for this earthquake.

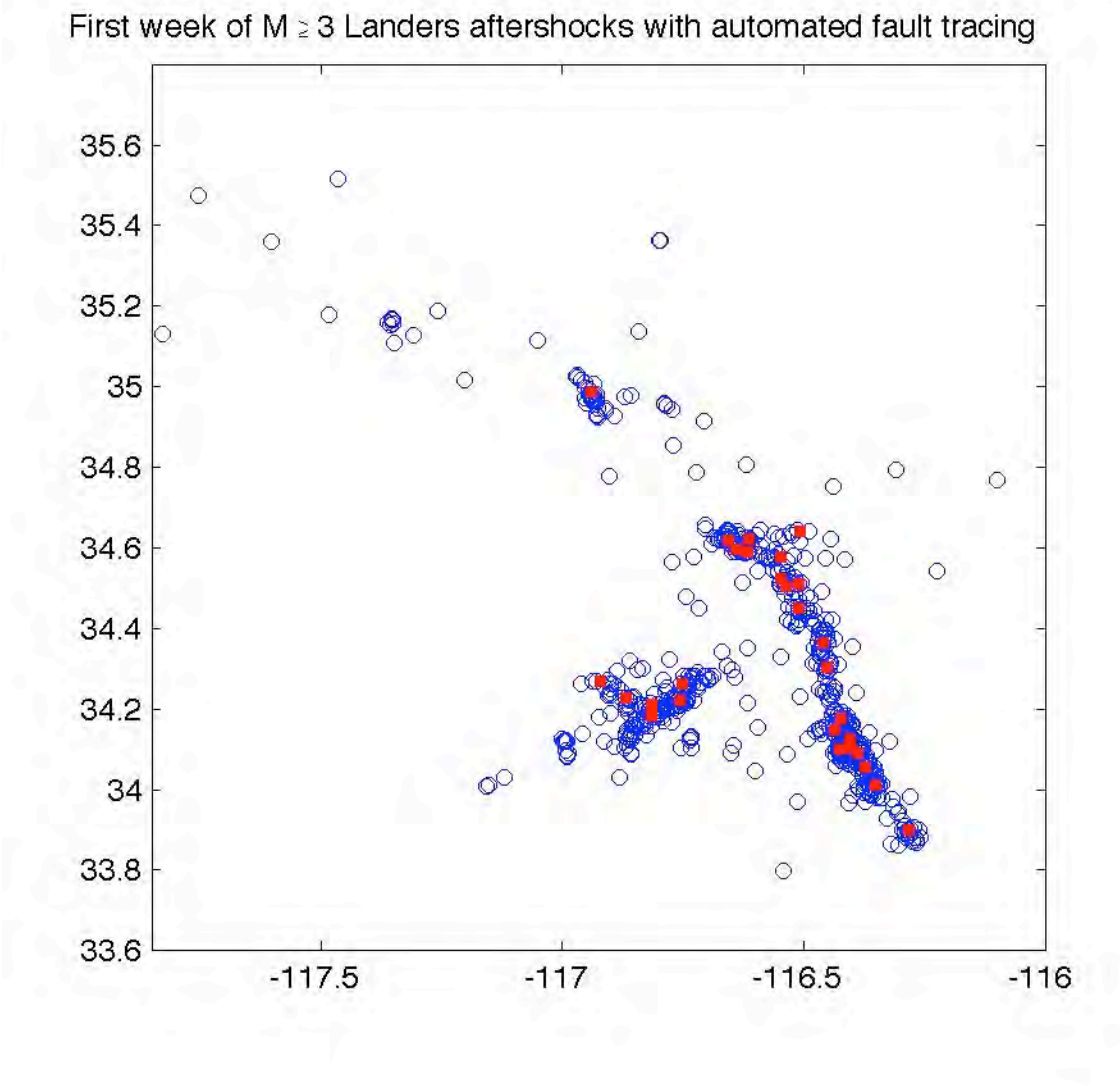


Figure 1a: First week of M≥3 aftershocks of the 1992 M 7.3 Landers earthquake. Aftershocks are given by blue circles and automatically found fault points by red squares.

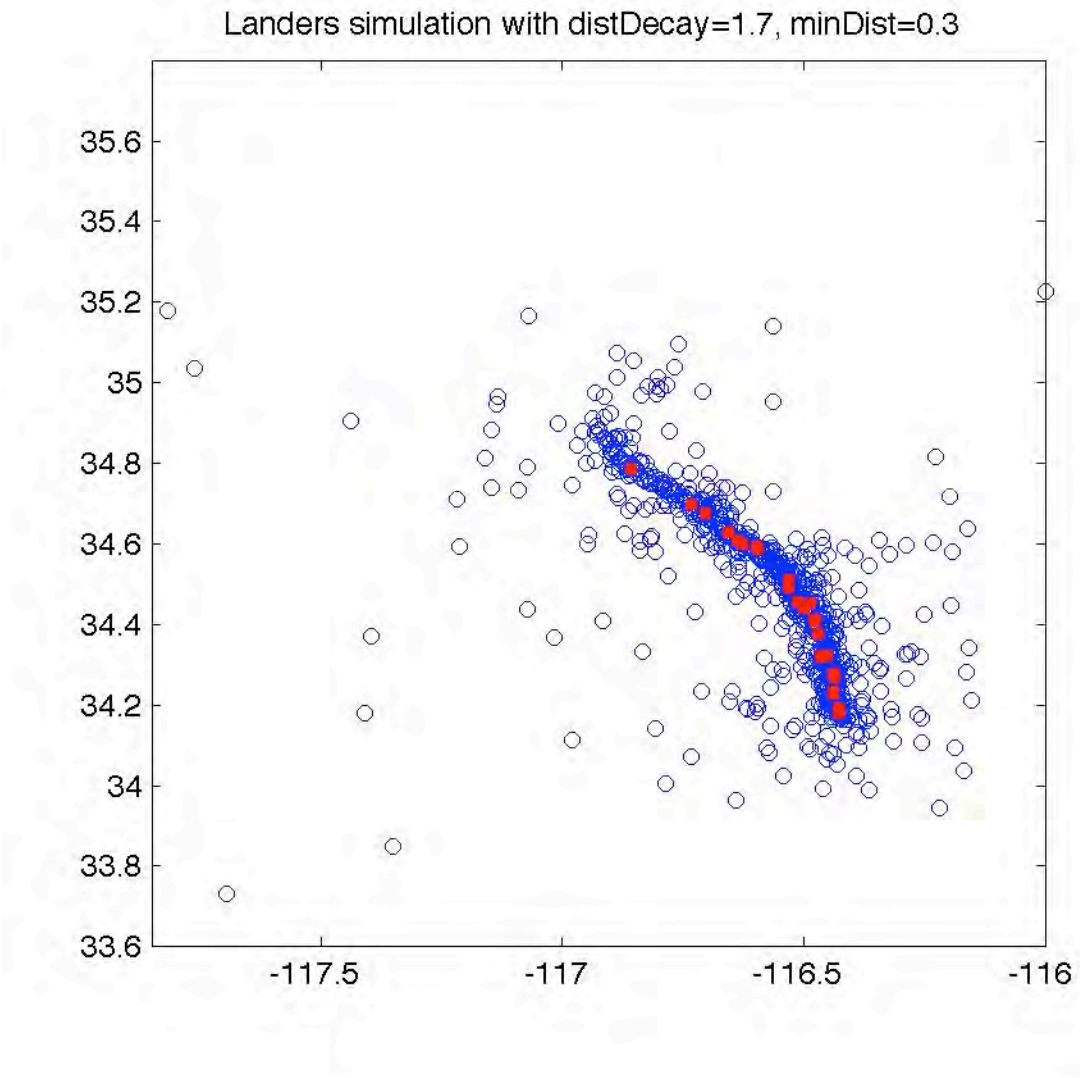


Figure 1b: First week of simulated M≥3 aftershocks with automatically found fault points with distDecay=1.7 and minDist = 0.3.

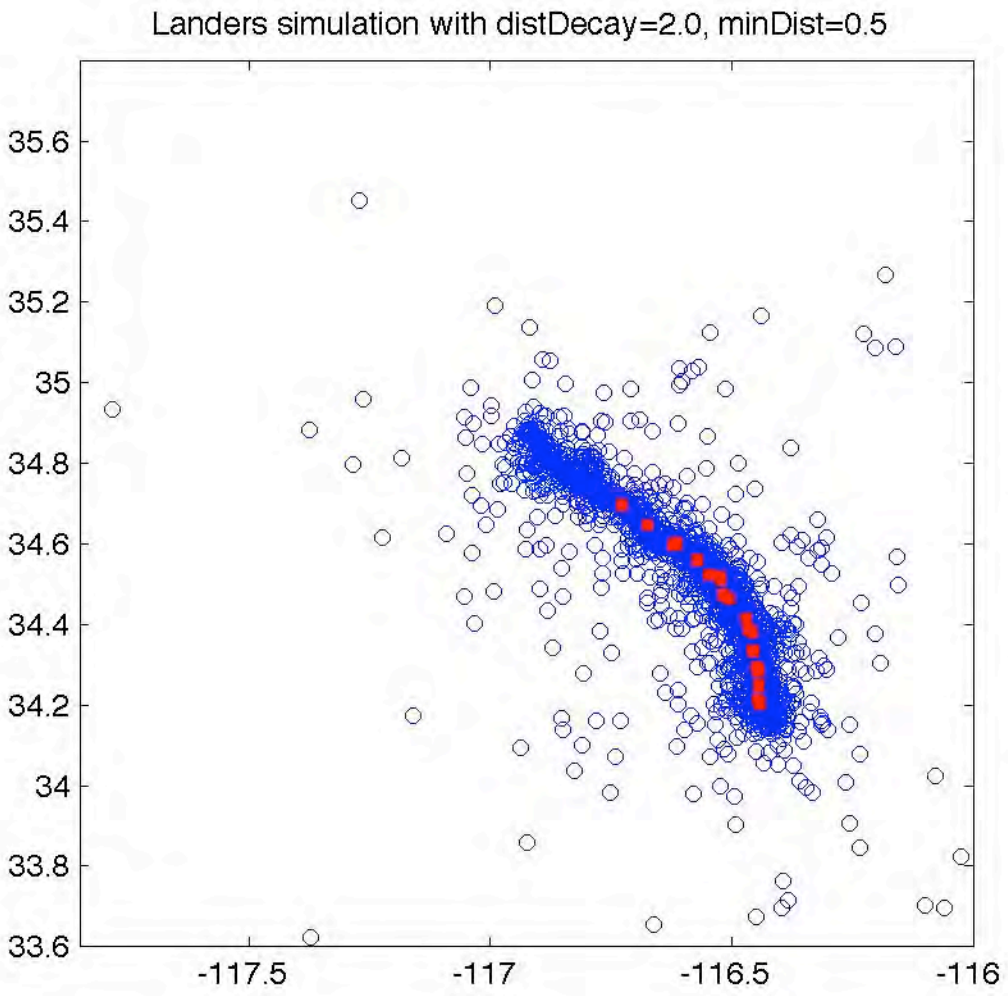


Figure 1c: Same as Figure 1b but with distDecay = 2.0 and minDist=0.5 km.

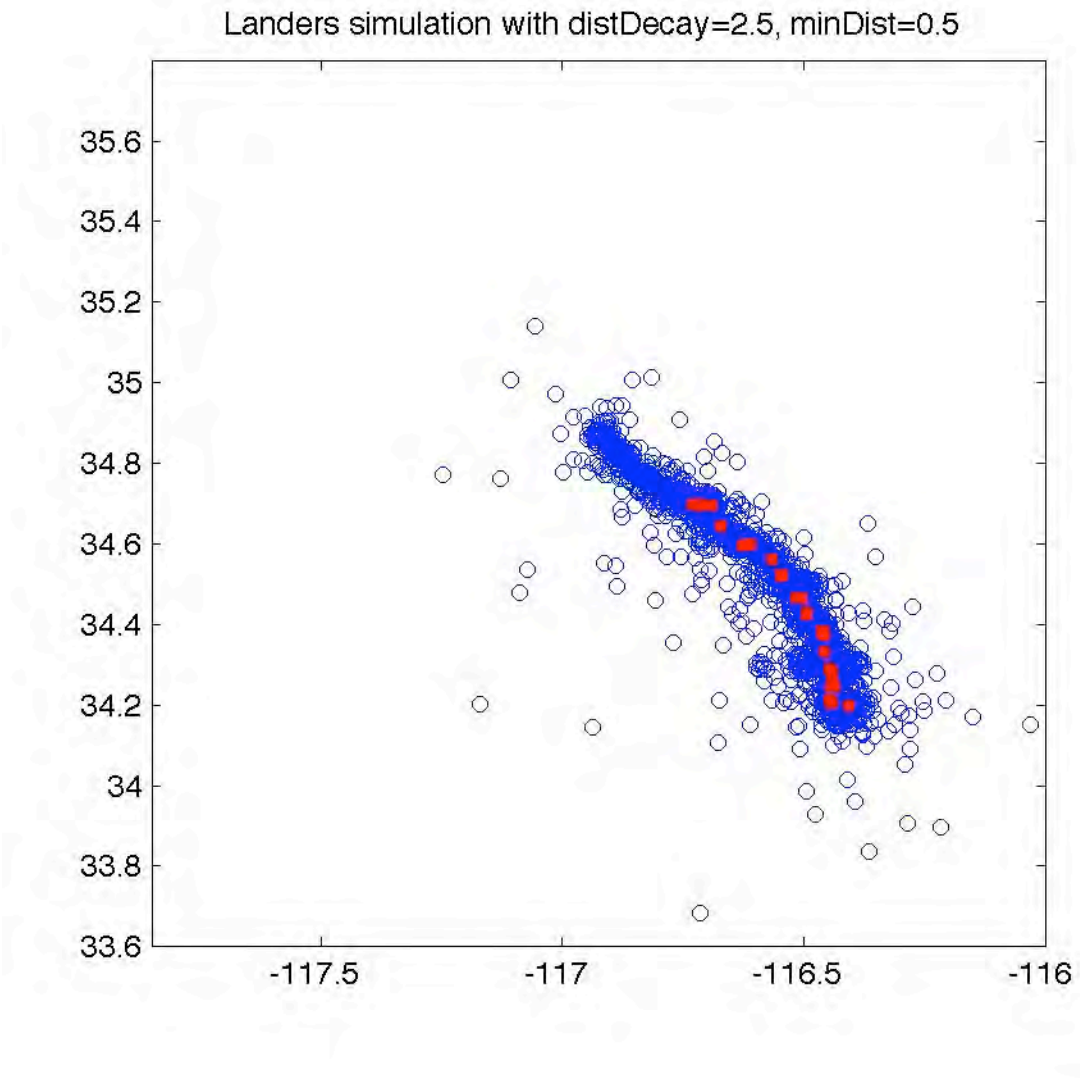


Figure 1d: Same as Figure 1b but with distDecay = 2.5 and minDist=0.5 km.

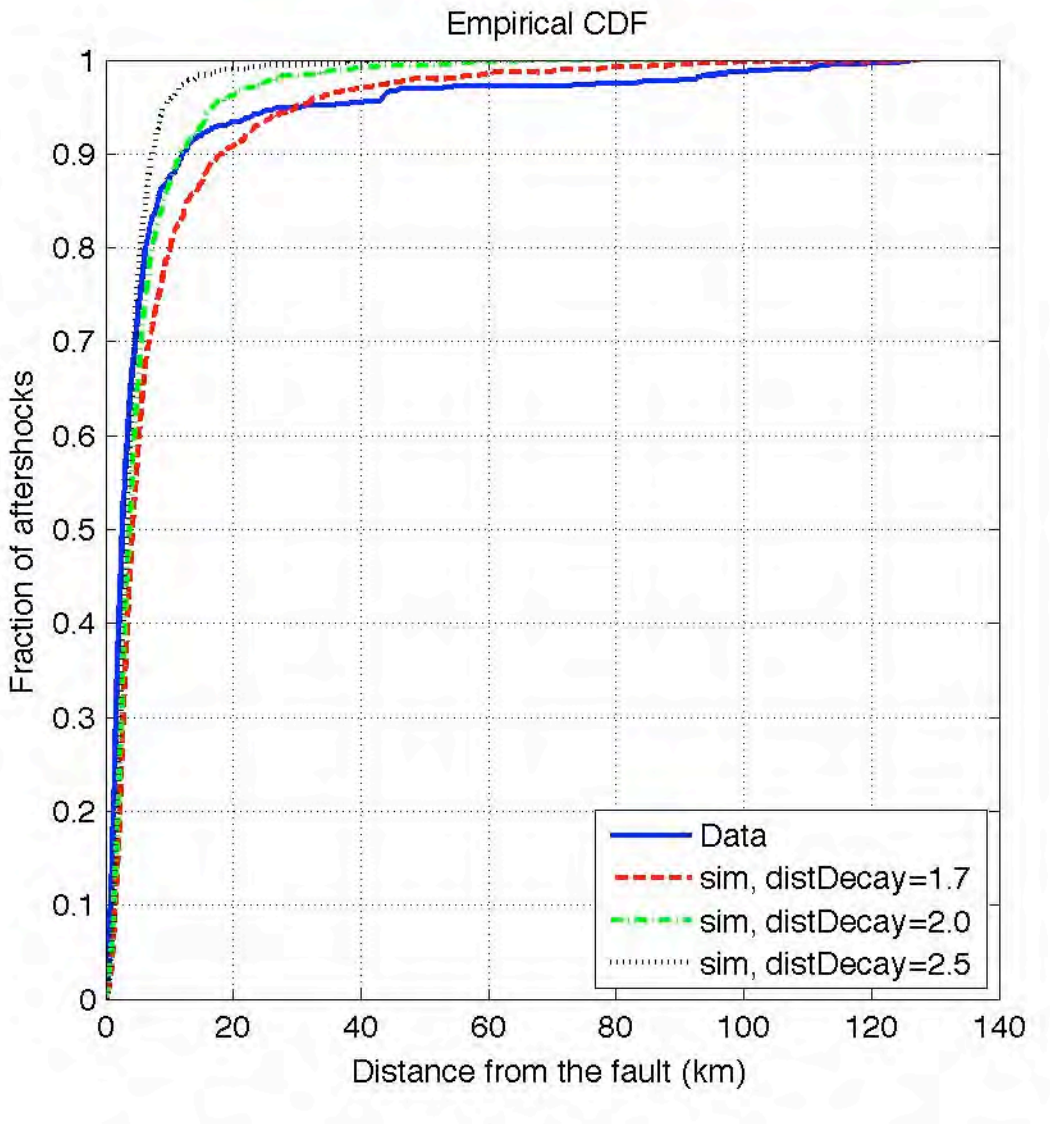


Figure 2a Cumulative probability distribution of distances between aftershock hypocenters and fault points, for the real data and three ETAS simulations

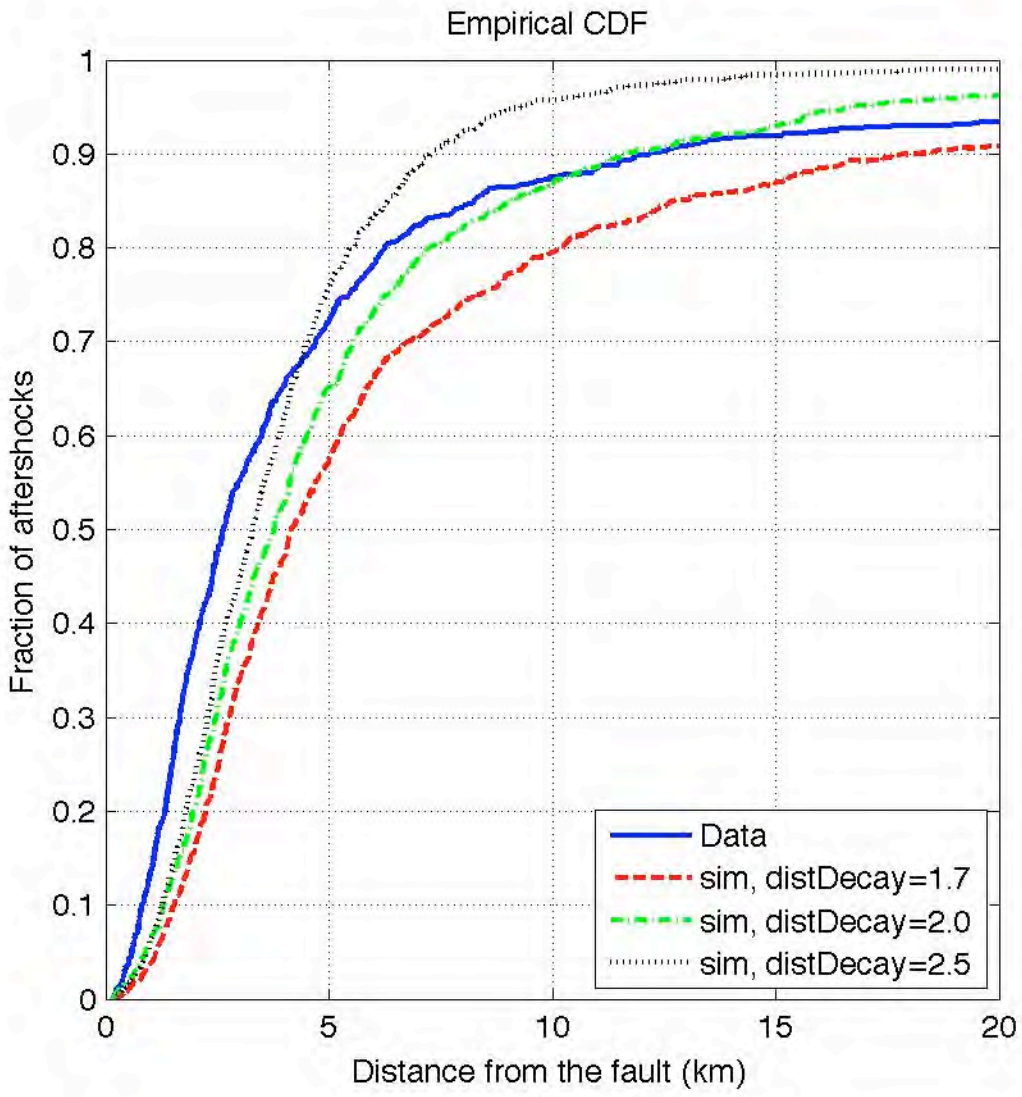


Figure 2b: Near field close up of Figure 2a.

We next inspect the same suite of plots for the Northridge earthquake. In this case at < 2 km distance we see a reasonable fit with the distDecay=2.5, minDist = 0.5 km simulation. Further away the pattern is the opposite of the Landers aftershock result – the aftershock density decays more rapidly, rather than more slowly, than any of the simulations, suggesting the need for a larger value of distDecay.

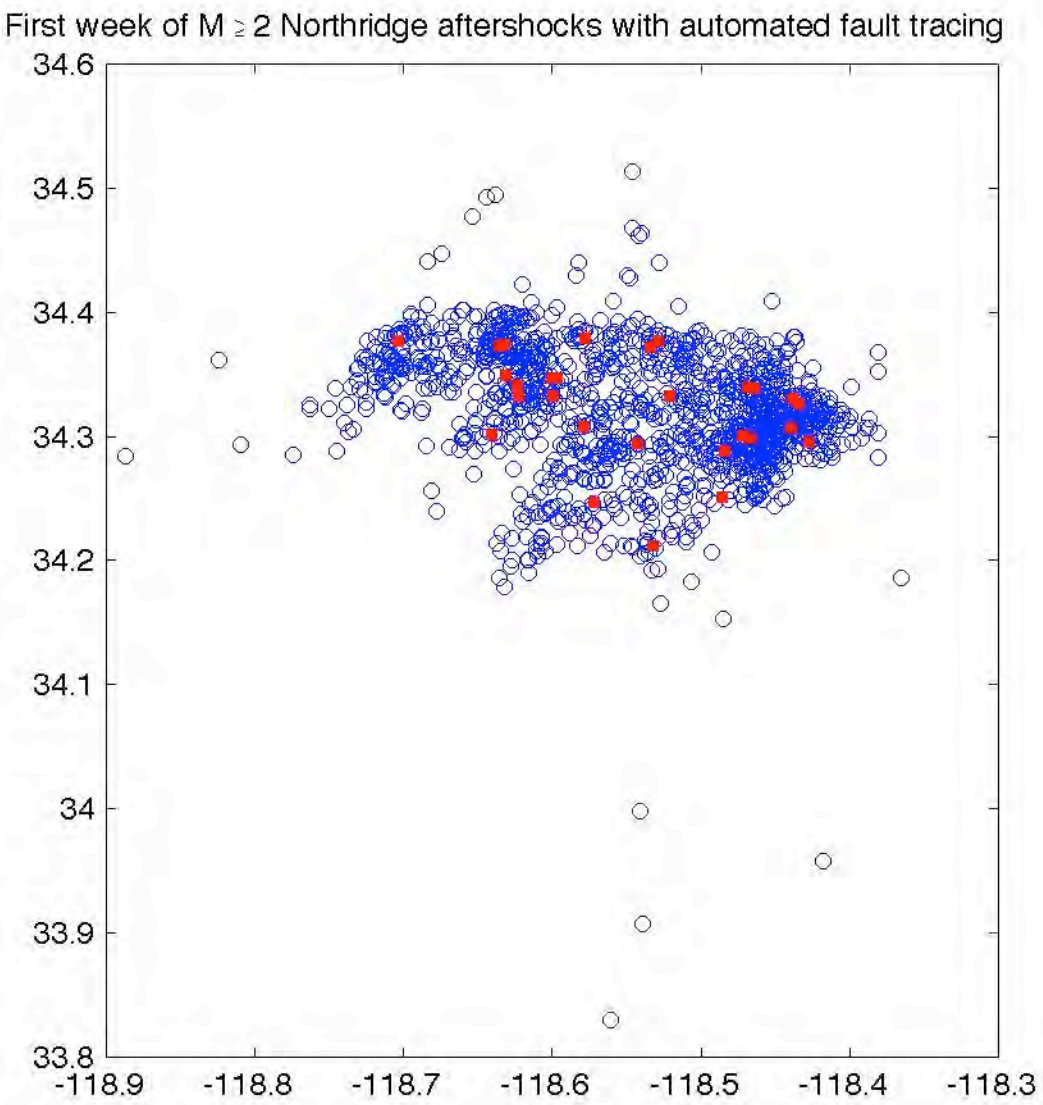


Figure 3a: First week of M≥2 aftershocks of the Northridge earthquake (given by blue circles), with automatically located fault points given by red squares.

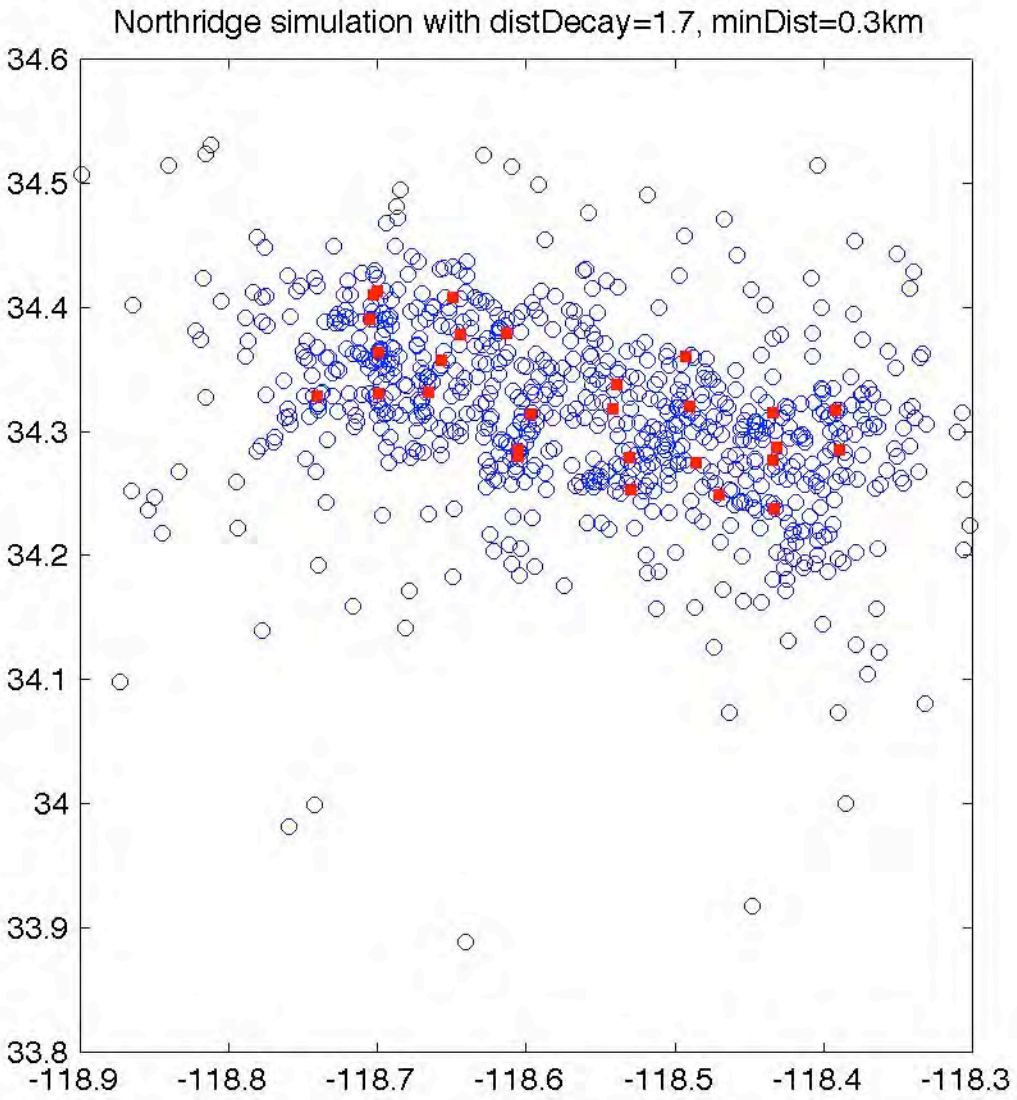


Figure 3b: First week of simulated M≥2.5 Northridge aftershocks, with automatically found fault points. This simulation uses distDecay=1.7 and minDist = 0.3 km.

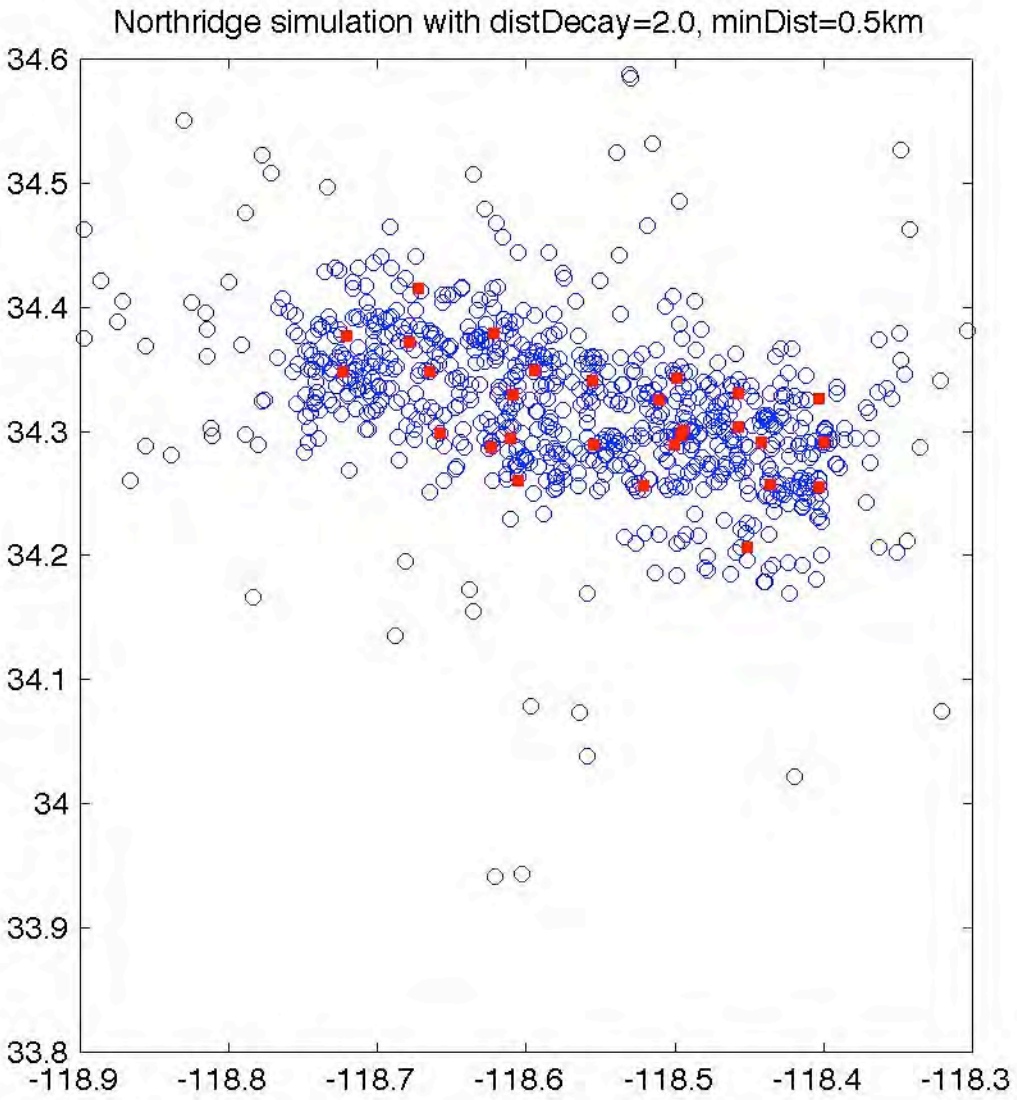


Figure 3c: Same as Figure 3b but with distDecay = 2.0 and minDist = 0.5 km.

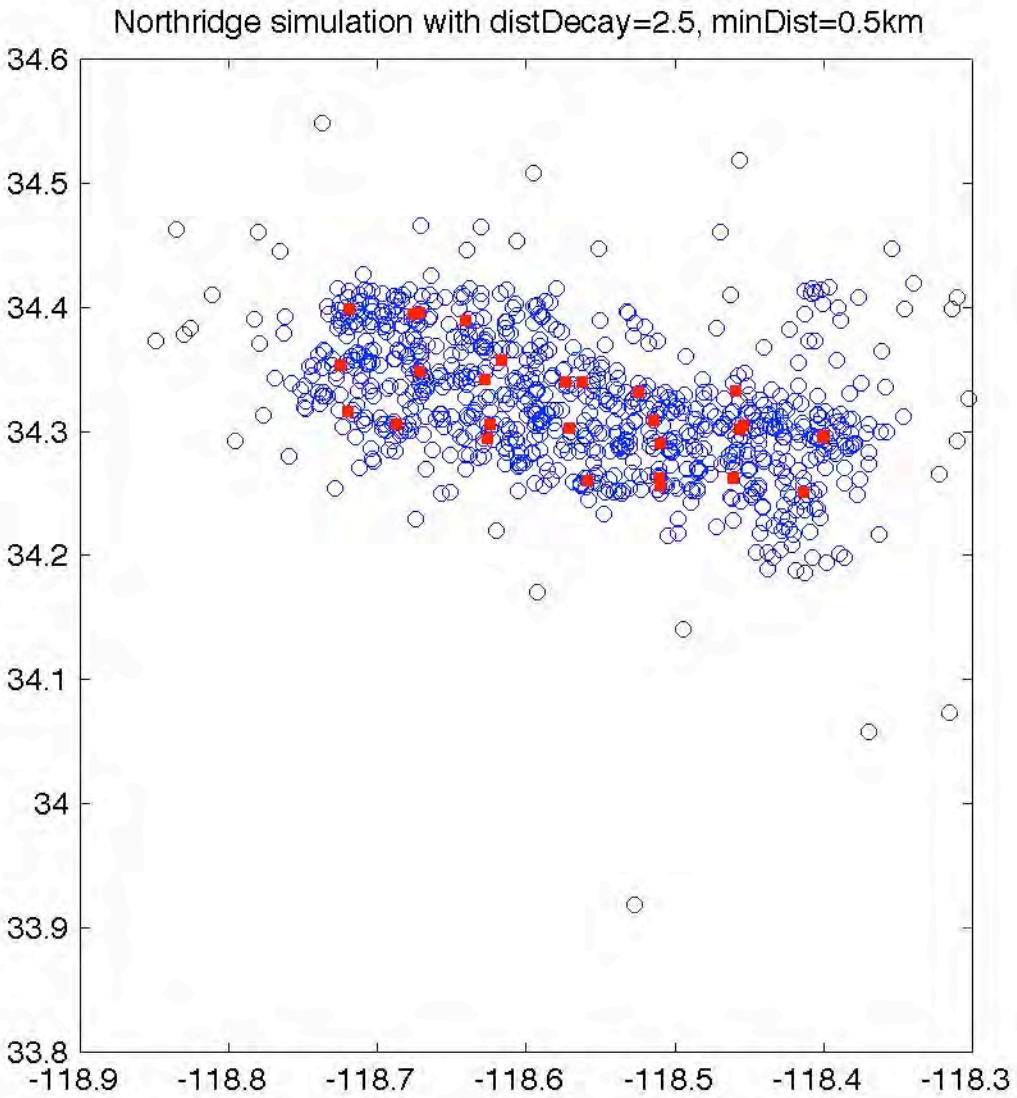


Figure 3d: Same as Figure 3b but with distDecay = 2.5 and minDist = 0.5 km.

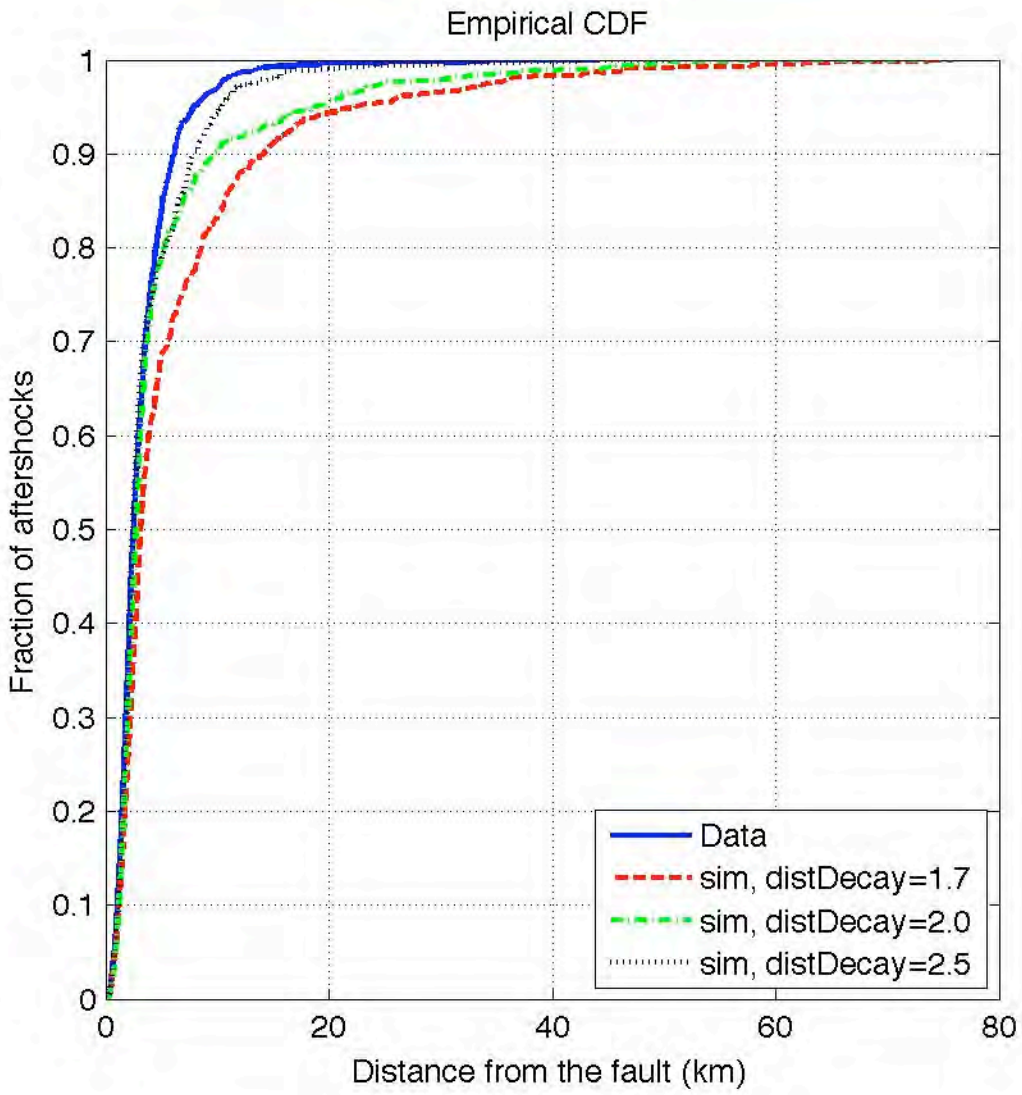


Figure 4a: Cumulative density functions (CDF) of distances between the aftershock hypocenters and mainshock fault points for the real data (blue solid line) and the different simulations.

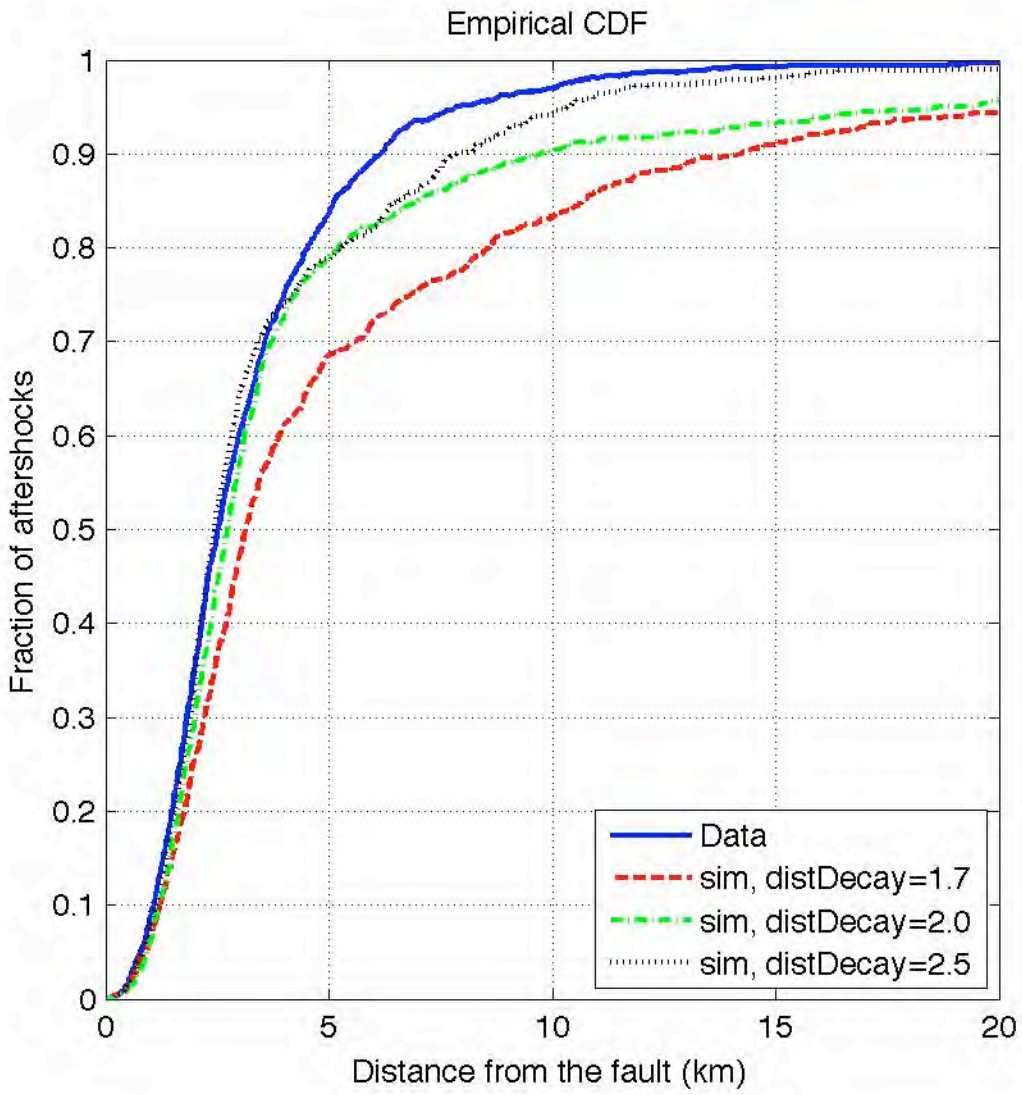


Figure 4b: Near field close up of Figure 4a.

The opposite results seen for Landers and Northridge might be artifacts, due to the variable completeness in the two aftershock data sets. The different focal mechanisms might also come into play – thrust earthquakes (Northridge) might produce fundamentally different aftershock patterns than strike slip earthquakes on vertical planes (Landers). The fact that aftershock depths tend to be more poorly constrained in earthquake catalogs than map locations might also be a factor. We need to investigate the matter further and repeat the analysis with relocated aftershock catalogs.

It is also very instructive to look at additional mainshocks. Below we aftershock distance CDFs from the M 6.3 Joshua Tree and M 7.1 Hector Mine earthquakes along with Northridge and Landers. It can be seen that for all of the sequences ~50% of the aftershocks occur within 3 km of the mainshock fault. In the far field Joshua Tree behaves more like the Northridge earthquake, while Hector Mine behaves more like the Landers

earthquake. The analysis of more earthquakes, and more ETAS simulations are required to converge on the best choices for the ETAS modeling, and to determine whether most aftershock sequences tend towards one set of parameters or whether sequence-specific parameters might be needed.

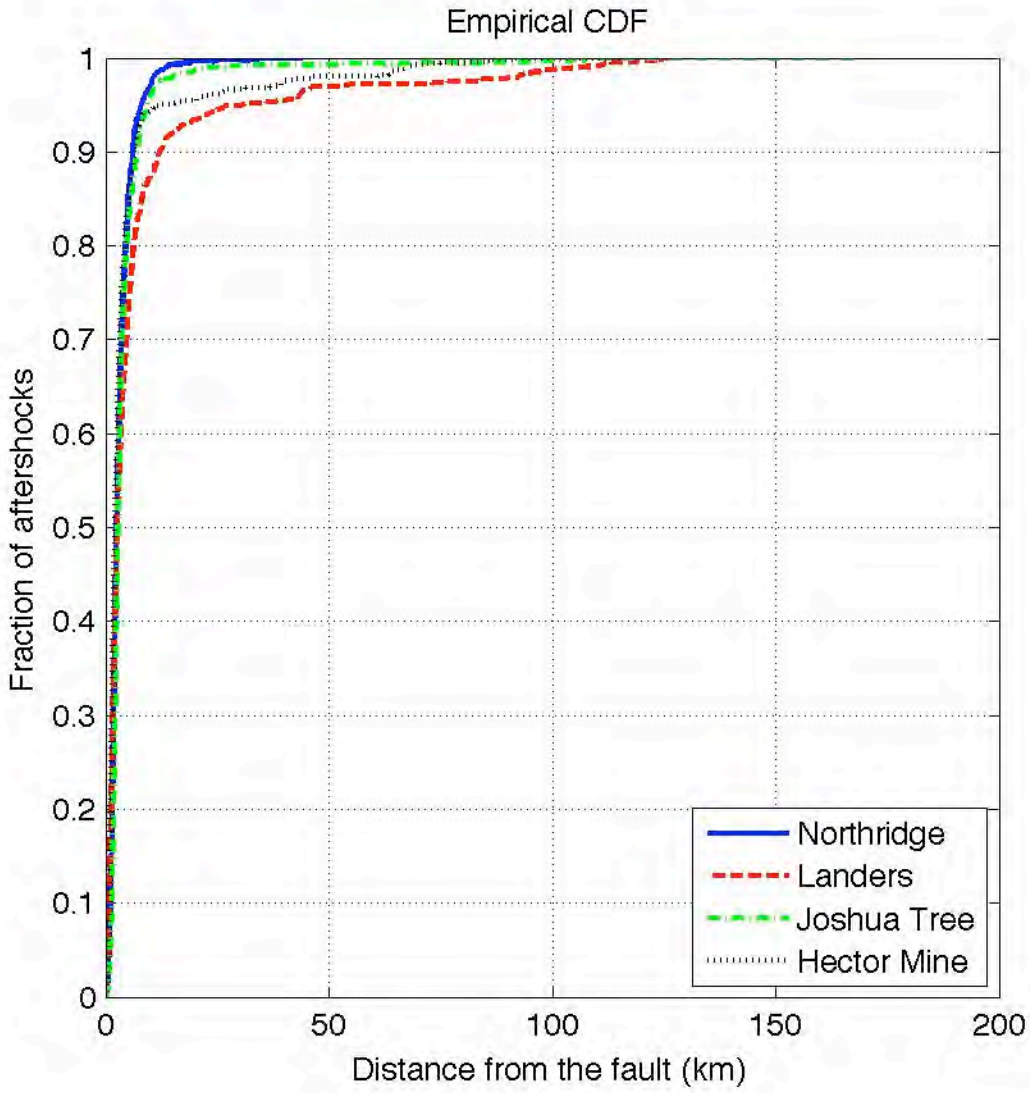
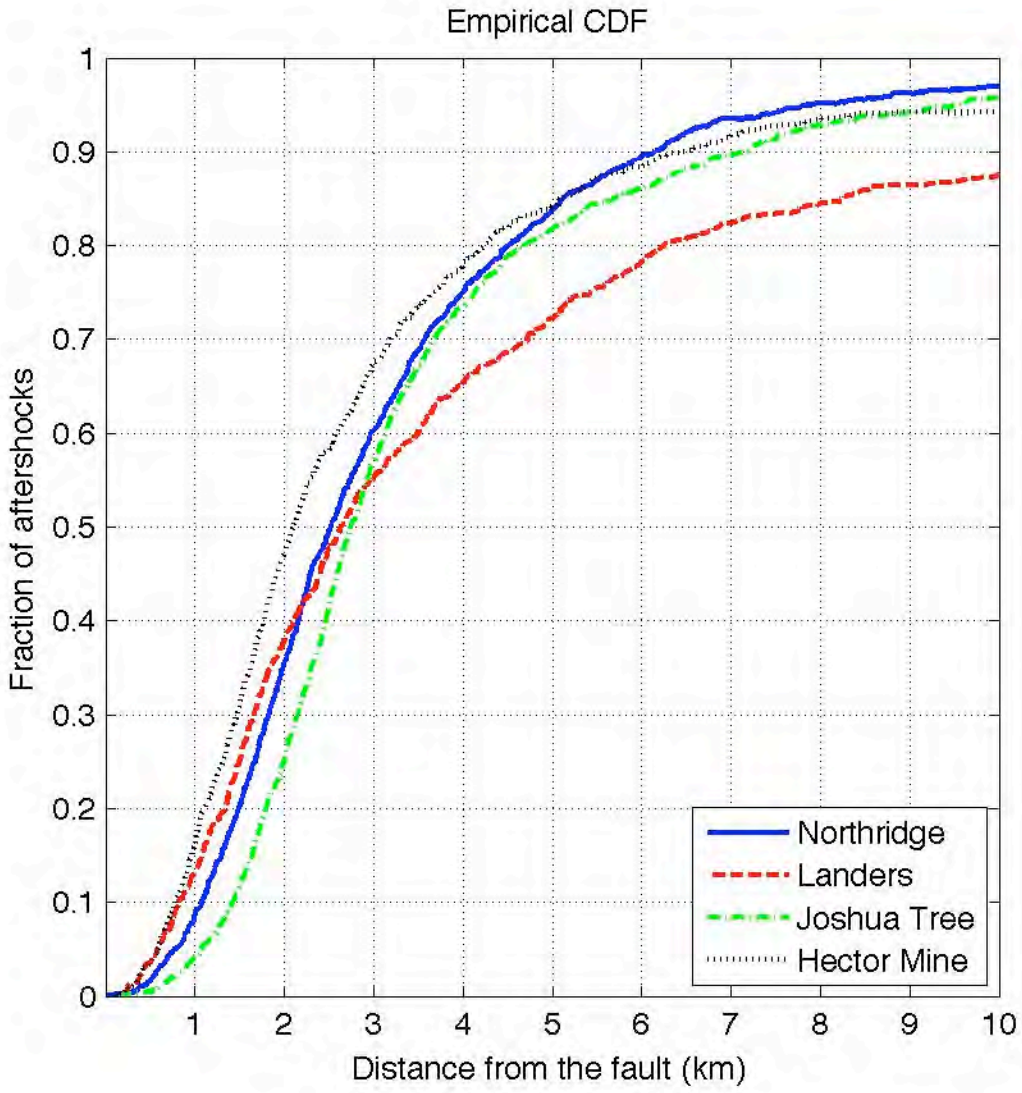


Figure 5a: CDF plots of the distances between aftershock hypocenters and mainshock fault points for the real first week aftershock data of the Northridge, Landers, Joshua Tree, and Hector Mine earthquakes.


Empirical CDF
Fraction of aftershocks
Distance from the fault (km)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1
2
3
4
5
6
7
8
9
10
Northridge
Landers
Joshua Tree
Hector Mine

Task P6 - Evaluate the Agnew & Jones method

Task Leader: Andy Michael (USGS)

Short-term probabilities based on earthquake clustering in California have been determined with two competing models that we find are more compatible than previously understood. The Reasenberg and Jones (1989) model combines the Gutenberg-Richter magnitude-frequency distribution with the modified-Omori law to produce the rate and probability of aftershocks following an event. By extrapolating to magnitudes above that of the initial event they also produce the probability that the initial event will be a foreshock. The Agnew and Jones (1991) model evaluates the probability that an event is a foreshock to a larger event by combining separate magnitude-frequency distribution for small to moderate background events with a long-term rate for the large events they attempt to forecast. In areas where the long-term rates are based on a characteristic earthquake model the long-term rates of the large events are significantly greater than the rates extrapolated from the Gutenberg-Richter distribution used for the smaller events. We call the combination of these two distributions a characteristic magnitude-frequency distribution. In such areas, the Agnew and Jones probabilities of the large events are significantly greater (up to two orders of magnitude larger) than those obtained from the Reasenberg and Jones model.

We have demonstrated that combining a characteristic earthquake magnitude-frequency distribution with the modified-Omori law creates a version of the Reasenberg and Jones clustering model that produces probabilities of large earthquakes that are equal to those obtained from the Agnew and Jones approach while also producing probabilities for small to moderate events that are equal to those from the Reasenberg and Jones model (Michael, 2011, submitted, BSSA). Thus, such a model combines the two existing models into one.

Our approach of using an ETAS model with the magnitude-frequency distribution drawn from the long-term model is equivalent to using a characteristic magnitude-frequency distribution near faults with high long-term rates of large earthquakes. Thus, in those areas, our approach will produce probabilities of large earthquakes that are similar to the Agnew and Jones approach. This is because in both our new approach and the Agnew and Jones model the long-term rates are reflected in the short-term clustering probabilities. This was not true of the Reasenberg and Jones approach where each cluster was considered in isolation from the long-term forecasts. It is also true that uncertainties in the long-term rates of the large events will propagate proportionally into the short-term forecasts. Thus, one of the most important issues in both the long-term and short-term forecasts is the topic of characteristic earthquake behavior.

Task P8 - Evaluation Other Time Dependencies

Task Leader: Jeanne Hardebeck

Large Earthquake Clustering:

A key question for operational earthquake forecasting in California is: If we build a spatiotemporal clustering model based on a stationary ETAS model determined primarily from smaller earthquakes, will it successfully forecast the larger (M≥6.5) potentially damaging earthquakes that we are most concerned about? Work on the global catalog of large earthquakes has proposed that there is long-range, long-term clustering of large earthquakes that is not captured by the ETAS model alone, and that a second layer of clustering on ~100 km length-scales and ~30 year time scales is necessary. To this end, the double-branching model has been proposed (*Marzocchi and Lombardi*, 2008).

Preliminary work has shown, however, that the double-branching model is not needed in California, and that ETAS alone does a good job of fitting the California catalog. *Hardebeck* (2010) fit ETAS parameters to the UCERF2 instrumental catalog, which is dominated by smaller earthquakes. The combined historical and instrumental UCERF2 catalog of M≥6.5 earthquake was declustered using the same ETAS parameters, and no significant residual spatiotemporal patterns were found. Warner Marzocchi (personal communication, 2010) independently confirmed this result, showing that the double-branching model does no better than the ETAS model in explaining the California earthquake catalog.

The ETAS model for the UCERF2 M≥6.5 earthquake catalog has one unsatisfying feature. Some earthquakes that intuition suggests are connected - for example Landers following Joshua Tree and Hector Mine following Landers - are assigned a low probability of being triggered. The ETAS model is not modeling long-term, long-range triggering between large earthquakes; rather, these large earthquakes appear Poissonian. Catalogs of large earthquakes in general tend to contain little clustering (e.g. *Sornette and Werner*, 2005) and hence have limited usefulness for operational earthquake forecasting. Including smaller earthquakes in the ETAS model, and therefore more secondary triggering of large earthquakes, increases the clustering. If all earthquakes with M≥3 are included in the ETAS model, the Landers and Hector Mine events are identified as triggered. Similar results were found by *Felzer et al.* (2003).

Another concern is that the California catalog of large (M≥6.5) earthquakes is quite short, and may contain too few earthquakes to identify subtle non-ETAS-like behavior with any statistical significance. Therefore, the failure to reject the ETAS null hypothesis is not a disproof of the double branching model. More work is required to further evaluate the double-branching model and its possible application to California. The double-branching model was developed on global datasets that encompass many different tectonic regions. A single set of ETAS parameters probably isn't appropriate everywhere, in fact *Zhuang* (2010) finds that even Japan can't be fit by a single set of ETAS parameter values. Forcing a single set of parameter values onto the entire globe may lead to poor fits in some locations, which appear as non-ETAS clustering. Additionally, the global catalog is dominated by subduction-zone earthquakes. Assuming that double-branching reflects a real

phenomenon, it is possible that it is related to the unique geometry of subduction zones and not applicable to most of California.

Paleoearthquake studies (e.g. *Rockwell et al.*, 2000; *Oskin et al.*, 2008) have shown that the earthquakes in the Eastern California Shear Zone (ECSZ) tend to cluster in time, with several fault systems producing large earthquakes within 1000-2000 years of each other, separated by >2000 years of few or no large earthquakes. Similar clustering has been identified in the Los Angeles area, out of phase with the ECSZ (e.g. *Dolan et al.*, 2007). Different hypothesis have been put forth, including mode-switching between the ECSZ and the Los Angeles area (*Dolan et al.*, 2007) and synchronization within the ECSZ due to stress-transfer coupling (*Scholz*, 2010).

These are intriguing results that should continue to be investigated. However, the mode-switching and coupling models currently exist as qualitative models, and have not yet been quantitatively tested. The lack of a quantitative framework also prevents these models from producing quantitative forecasts that could be incorporated into UCERF3. Additionally, given the very long time scales of this phenomenon, it is not clear that including it would alter the UCERF3 forecasts. For example, if we were to forecast the next 30 years based on the seismicity of the last 100 years and the geodetic deformation of the last 10 years, a ~1000 year cycle in the earthquake behavior would not be a large effect. The issue of mode-switching and coupling is part of the larger issue of reconciling geologic data, which samples long time-periods, with seismological and geodetic data, which samples the more recent past, and accounting for any differences in behavior at these different time scales.

Swarms:

Although swarms are typically dominated by small-magnitude earthquakes, and UCERF is focused on the larger more damaging earthquakes, swarm forecasts could be beneficial to society in two different contexts:

(1) An ongoing swarm in or near an urban area would be felt by many people and may cause property damage. A felt swarm would be a source of great concern for the affected community, and the public would be anxious to know what may happen as the swarm continues. An example is the 2008 Mogul swarm near Reno, Nevada (e.g. *Powers and Maugh*, 2008).
(2) Swarms tend to occur near the Salton Sea, in proximity to the southernmost San Andreas Fault, which the UCERF2 report identified as the fault section most likely to fail in the next 30 years. Assuming that each earthquake in a swarm has some probability of triggering a San Andreas earthquake, the rate of earthquakes during the swarm must be forecast accurately in order to properly forecast the probability of triggering a southern San Andreas earthquake.

The temporal evolution of swarms differs from regular seismicity primarily in the greatly increased rate of spontaneous earthquakes, sometimes at rates thousands of times higher

than usual (*Llenos et al.*, 2009). This increase in the rate of spontaneous earthquakes is thought to be due to increased loading rate due to slow slip events or fluid movement. To properly implement swarm forecasts in a spatiotemporal clustering model would require identifying the existence of a swarm, quantifying the rate of spontaneous swarm events, and temporarily updating the background rate accordingly.

A spatiotemporal clustering model based on a stationary ETAS model will provide some forecasting ability during a swarm, since it would accurately model the number of aftershocks spawned by the swarm events that have already occurred. Although a stationary ETAS model would underestimate the number of new spontaneous events occurring during the forecast period, it would still forecast a large number of events to occur in the swarm area. This is adequate for the first implementation of the ETAS-based spatiotemporal clustering model, so the priority should be to build a functioning ETAS-based forecast. A swarm module could be built into the ETAS framework at a later date.

References:

Dolan, J. F., D. D. Bowman, and C. G. Sammis (2007). Long-range and long-term fault interactions in Southern California, Geology 35, no. 9, 855–858.

Felzer, K.R., Abercrombie, R.E., Ekström, G. (2003). Secondary aftershocks and their importance for aftershock forecasting, Bulletin of the Seismological Society of America, 93 (4), pp. 1433-1448.

Hardebeck, J. L. (2010). No difference in the spatial-temporal clustering of large and small earthquakes in California, in preparation.

Llenos, A. L., J. J. McGuire, Y. Ogata (2009). Modeling seismic swarms triggered by aseismic transients, Earth and Planetary Science Letters 281, 59-69.

Powers, A. and T. H. Maugh (2008). Strange series of quakes sets Reno on edge, *Los Angeles Times* (http://articles.latimes.com/2008/may/01/nation/na-quakes1).

Marzocchi, W., Lombardi, A.M. (2008). A double branching model for earthquake occurrence, Journal of Geophysical Research B: Solid Earth, 113 (8), B08317, DOI: 10.1029/2007JB005472.

Oskin, M., L. Perg, E. Shelef, M. Strane, E. Gurney, B. Singer, and X. Zhang (2008). Elevated shear zone loading rate during an earthquake cluster in eastern California, Geology 36, no. 6, 507–510.

Rockwell, T.K., Lindvall, S., Herzberg, M., Murbach, D., Dawson, T., and Berger, G., (2000). Paleoseismology of the Johnson Valley, Kickapoo, and Homestead Valley faults: Clustering of earthquakes in the eastern California shear zone: Bulletin of the Seismological Society of America, v. 90, p. 1200–1236.

Scholz, C. H. (2010). Large Earthquake Triggering, Clustering, and the Synchronization of Faults, Bulletin of the Seismological Society of America, Vol. 100, No. 3, pp. 901-909.

Sornette, D., Werner, M.J. (2005). Apparent clustering and apparent background earthquakes biased by undetected seismicity, Journal of Geophysical Research B: Solid Earth, 110 (9), B09303, DOI: 10.1029/2005JB003621.

Zhuang, J. (2010). Detecting spatiotemporal seismicity variation by using the maximizing weighted likelihood method, in preparation.

Task P10 - Compile Slip in Last Event Data

Task Leader: Ray Weldon (UO)

Slip for the last event along a fault, in conjunction with the application of appropriate recurrence models, can be used to estimate the timing and/or size of future ground-rupturing earthquakes. Toward this goal, we will compile historic, paleoseismic and geomorphic data on the **timing** and **slip** for the last event on high-rate California faults. Slip measurements are relatively easy to acquire along highly active faults because offsets from the last event is usually well preserved on the landscape. For strike slip faults, slip is determined from offset meter-scale landforms, whereas for dip-slip faults, slip is derived from topographic profiles across recent scarps, or from the youngest offset units exposed in paleoseismic trenches. All faults that have ruptured historically, and many faults where the approximate (^{14}C) timing of the most recent event is known from paleoseismic studies, have estimates of the associated slip documented in the literature, or our own paleoevents database. Furthermore, recent analyses of newly available, high-resolution LIDaR topography for micro-geomorphic offsets have substantially increased the number of slip measurements available for the F4 task. For faults or fault sections with LIDaR coverage, but limited, poor, or unavailable offset data, we will identify promising reaches from which to extract slip measurements ourselves.

We have developed a four-step process to focus and streamline this data collection. Step one prioritizes faults with the highest slip rates, and longest time since the last event relative to average recurrence interval (Figure 1). Slip rate, timing of the last event, and recurrence interval are used from past working group reports, unless more recent data are available.

Google docs UCERF3 F4 Fault priority list Private to only me Saved se

File Edit View Insert Format Form Tools Help

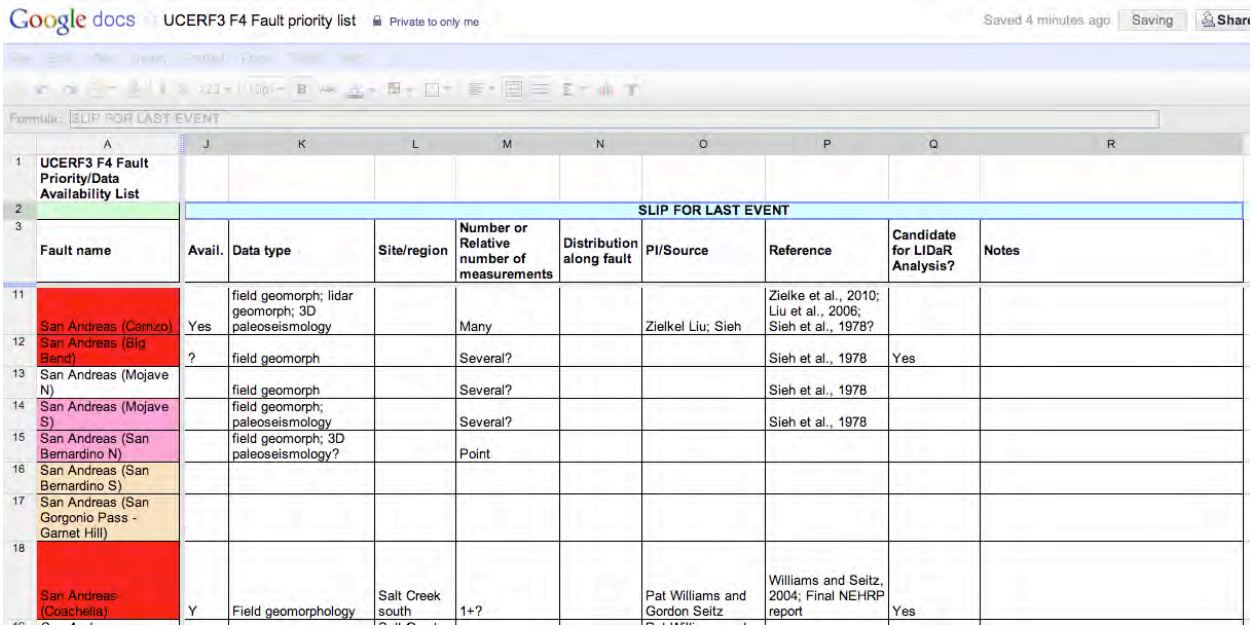
Formula: WG 2,007

	A	B	C	D	E	F	G	H	I
1	**UCERF3 F4 Fault Priority/Data Availability List**								
2		**FAULT INFORMATION**							
3	**Fault name**	**Slip rate**	**+/-**	**Reference**	**Timing of last rupture**	**Reference**	**Avg Recurrence [mean (min-max)]**	**Reference**	**Years since last rupture**
11	San Andreas (Carrizo)	34	3	Slip rate based on Sieh and Jahns (1984).	1857	Historical	88 (47-129)	Sinan et al., 2010	154
12	San Andreas (Big Bend)	34		Slip rate based on Sieh and Jahns (1984).	1857	Historical	75; 83 (not calculated)	Scharer et al., 2009; SCEC abstract	154
13	San Andreas (Mojave N)	27	7	Slip rate based on Sieh (1984), Salyards et al. (1992), and WGCEP (1995)	1857	Historical	190 (155-252)	WG 2007	154
14	San Andreas (Mojave S)	29	7	Slip rate based on Sieh (1984), Salyards et al. (1992), and WGCEP (1995)	1857	Historical	155 (129-208)	WG 2007	154
15	San Andreas (San Bernardino N)	22	6	Slip rate reported by Weldon and Sieh (1985)	1812		175 (145-208)	WG 2,007	199
16	San Andreas (San Bernardino S)	16	6	Slip rate reported by Weldon and Sieh (1985)	1812		244 (199-405)	WG 2,007	199
17	San Andreas (San Gorgonio Pass - Garnet Hill)	10	6	Slip rate reported by Weldon and Sieh (1985)	c. 1680		398 (224-1002)	WG 2,007	331
18	San Andreas (Coachella)	20	6	Slip rate based on Sieh and Williams (1990); Sieh (1986); Keller et al. (1982); Bronkowski (1981)	c. 1680	Philibosian et al., 2011; Fumal et al., 2002; Sieh et al., 1978)	156 (61-214); 246 years	WG 2007; Dawson et al., 2008	331

Figure 1*. Example of F4 fault priority list. Higher priority faults or fault sections (red colors) have had more time elapse since the last rupture than their average recurrence interval.*

This effort is coordinated with Tim Dawson and Ned Field. All tabular data is compiled in a Google Docs spreadsheet to aid the exchange of data between team members.

In step two, a literature search determines the availability of offset data for the highest priority faults from step one (Figure 2). Wherever possible, we directly contact PIs of published slip studies to ascertain whether other data exist in unpublished archives, proposals, gray literature, abstracts or publications in preparation. Step 2 also identifies faults or fault reaches with limited data that are potential candidates for microgeomorphic analysis.



Google docs UCERF3 F4 Fault priority list Private to only me — Saved 4 minutes ago — Saving

	Fault name	Avail.	Data type	Site/region	Number or Relative number of measurements	Distribution along fault	PI/Source	Reference	Candidate for LIDaR Analysis?	Notes
1	UCERF3 F4 Fault Priority/Data Availability List									
2		SLIP FOR LAST EVENT								
11	San Andreas (Carrizo)	Yes	field geomorph; lidar geomorph; 3D paleoseismology		Many		Zielkel Liu; Sieh	Zielke et al., 2010; Liu et al., 2006; Sieh et al., 1978?		
12	San Andreas (Big Bend)	?	field geomorph		Several?			Sieh et al., 1978	Yes	
13	San Andreas (Mojave N)		field geomorph		Several?			Sieh et al., 1978		
14	San Andreas (Mojave S)		field geomorph; paleoseismology		Several?			Sieh et al., 1978		
15	San Andreas (San Bernardino N)		field geomorph; 3D paleoseismology?		Point					
16	San Andreas (San Bernardino S)									
17	San Andreas (San Gorgonio Pass - Garnet Hill)									
18	San Andreas (Coachella)	Y	Field geomorphology	Salt Creek south	1+?		Pat Williams and Gordon Seitz	Williams and Seitz, 2004; Final NEHRP report	Yes	

***Figure 2**. Information compiled during the literature search in step two.*

In step three, we will attempt to fill data gaps identified in step two by conducting microgeomorphic analysis of faults with available LIDaR data, and high preservation potential of offset geomorphic landforms, Methodology for these analyses has been developed and implemented successfully in recent studies by Middleton, Salisbury, and Zielke.

Lastly, step four compiles data collected during steps two and three into a composite table of slip measurements. Table attributes are based on the working group paleo-sites database and a comprehensive reporting approach developed by Zielke and others for offset channels. All data are evaluated using a five-tiered quality index based on semi-quantitative rating methodology developed by Zielke, Arrowsmith and others to characterize the reliability of offset geomorphic features to preserve coseismic slip. For the this task we expand their rating system to include measurements from dip-slip faults.

To date we are finishing the priority list in step one, and have begun contacting PIs in step two. **PIs with knowledge of unpublished slip data, or suggestions for reaches that may yield good results from microgeomophic analyses are encouraged to contact us.**