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

by the Working Group on California Earthquake Probabilities (WGCEP)

Notes:

• This version of the project plan will constitute our Dec 31 report to CEA titled “Methodology Assessment – Proposed Solutions to Issues (Report #2)”

• However, this is more than just proposed solutions to issues, but rather a relatively complete description of the model we envision. As such, future reports can hopefully iterate on this one.

• We hope to compile a “Key Assumptions” list before the Nov 11th-12th meeting and to add it as a section shortly thereafter.

• This report assumes some familiarity with both previous WGCEPs and the literature on time-dependent probabilities (e.g., BPT and ETAS are not defined in detail). References are also a bit incomplete on some of these topics, which we will certainly rectify in the future.

• Not all core WGCEP participants necessarily agree with all aspects of the model outlined here.

• Due to the extensive nature of revisions, changes have not been tracked, although you could always use the “compare documents” function on our previous report.
# Table of Contents

Introduction ........................................................................................................................................... 4  
Background ........................................................................................................................................ 4  
Implementation Plan .......................................................................................................................... 6  
  Epistemic Uncertainties and Logic Trees ......................................................................................... 6  
  Participants ........................................................................................................................................ 6  
  Consensus Building ......................................................................................................................... 6  
  Coordination with NSHMP ............................................................................................................. 7  
  Time Dependencies, Operational Aspects, and Potential Users .................................................. 7  
  Contingency Plans .......................................................................................................................... 8  
CEA Delivery Schedule ..................................................................................................................... 10  
Main Model Components ............................................................................................................... 11  
  Fault Models .................................................................................................................................. 11  
    Tasks: ....................................................................................................................................... 12  
  Deformation Models ...................................................................................................................... 13  
    Tasks: ....................................................................................................................................... 14  
  Earthquake Rate Models .............................................................................................................. 15  
    Methodology: .............................................................................................................................. 16  
    Implementation Considerations: ................................................................................................. 22  
    Tasks: ....................................................................................................................................... 24  
  Earthquake Probability Models: .................................................................................................. 27  
    Resolve Interpretation of Empirical Model: ............................................................................... 27  
    Develop Self-Consistent Elastic Rebound Models: .................................................................... 28  
    Apply Spatial-Temporal Clustering Models: ............................................................................... 35  
    Tasks: ....................................................................................................................................... 40  
Model Implementation ...................................................................................................................... 41  
  Tasks: ....................................................................................................................................... 41  
  Other Possible Tasks: .................................................................................................................... 42  
Planned Workshops & Review Meetings ......................................................................................... 43  
References ......................................................................................................................................... 45  
Appendix – Detailed Task Descriptions ......................................................................................... 47  
  Task F1 - Update Faults ................................................................................................................ 47  
  Task F2 – Reevaluate Fault Endpoints ......................................................................................... 49  
  Task F3: Database Issues ............................................................................................................. 51  
  Task F4 - Update Faults ................................................................................................................ 53  
  Task D1 – Evaluate New Deformation Models ............................................................................ 54  
  Task D2 – B-Fault Bias .................................................................................................................... 55  
  Task D3 – Line Integral Tools ..................................................................................................... 55  
  Task R1 - Evaluate Along-Strike Distribution of Average Slip ................................................... 62  
  Task R2 – Evaluate Magnitude-Scaling Relationships and Depth of Rupture ......................... 64  
  Task R3 - Paleoseismic Recurrence Interval Estimates ............................................................... 68  
  Task R4: Probability of Seeing Events in a Paleo Trench ............................................................ 74  
  Task R5: Solve the Large Inversion Problem .............................................................................. 75
Task R6 - Fault-to-Fault Jumping Probabilities

Task R7 - Reassess Historical Earthquake Catalog

Task R8 - Reevaluate Earthquake Catalog

Task R9: Smoothed Seismicity Model

Task R11: Focal Mechanisms of Off-Fault Seismicity

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

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

Task R15 – Cascadia Subduction Zone

Task P1 – Address Empirical Model

Task P2 – ETAS explains Empirical Model?

Task P3 - Coulomb Stress Explains Empirical Model?

Tasks P5 & P6 - Implement ETAS for spatial-temporal clustering, and Evaluate the Agnew & Jones method

Task P8 - Evaluation Other Time Dependencies
Introduction

On June 25, 2009 the California Earthquake Authority (CEA) governing board approved $2 million for the development of a Uniform California Earthquake Rupture Forecast, version 3 (UCERF3), and in late October 2009 they approved the final contract. The project started on January 1, 2010, and will last for 30 months.

Our primary goals with UCERF3 are to include multi-fault ruptures and spatiotemporal clustering. The latter will require robust interoperability with real-time seismicity information, and as such, UCERF3 will bring us into the realm of operational earthquake forecasting.

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

On April 14, 2008, the Working Group on California Earthquake Probabilities (http://www.WGCEP.org) publicly released the Uniform California Earthquake Rupture Forecast, version 2 (UCERF2). The development of this model was a joint effort between SCEC, the USGS, and CGS, with considerable support from the California Earthquake Authority (CEA), a provider of residential earthquake insurance. The main report, 16 appendices, executive summary, supplemental data, press release, and a fact sheet are all available at http://www.SCEC.org/ucerf.

Perhaps the most important accomplishment represented by UCERF2 was the development of a statewide model that uses consistent methodologies, data-handling standards, and uncertainty treatment in all regions. Also noteworthy is the coordination and consistency between UCERF2 and the model used in the 2008 USGS national seismic hazard maps (the latter using a time-independent version of UCERF2).

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 more 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 ≥ 7) earthquakes, which could have a significant impact on both hazard and loss estimates.
The working group identified the following shortcomings and/or issues with UCERF2 (which represents opportunities for improvement in UCERF3):

- **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).

All of the above issues are discussed extensively in the UCERF2 report. Each of these problems motivates aspects of the UCERF3 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 an extensive and extensible IT infrastructure upon which we can build, which will save much time and money compared to starting from scratch.
Implementation Plan

Epistemic Uncertainties and Logic Trees
Because there will be no single consensus model for UCERF3, it will be important that the modeling framework adequately represent epistemic uncertainties (our lack of understanding of how nature works, as opposed to “aleatory” uncertainty which represents the inherent randomness assumed in a given model). As with UCERF2, we anticipate representing epistemic uncertainties in UCERF3 using logic-tree branches that account for multiple models constructed under different assumptions and constraints.

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) (see “Participants at http://www.WGCEP.org for current members). Other WGCEP participants include research scientists, resource experts, model advocates, and IT professionals.

The ExCom is responsible for convening experts, reviewing options and making decisions about model components, and orchestrating implementation of the model and supporting databases. One role of the ExCom is to ensure that the models incorporated into UCERF3 span the range of model viability. The MOC is in charge of resource allocation and approving project plans, budgets, and schedules; it is also responsible for seeing that the models are properly reviewed and delivered. The SRP 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’s 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 preliminary 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/or getting feedback from user communities.
Decisions with respect to logic-tree branches and weights are the responsibility of the ExCom. The ExCom must 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 likely need to rely on expert opinion in establishing some logic-tree branch weights, it is our explicit goal to base these decisions on criteria that are as quantitative, reproducible, and testable as possible. Take the likelihood of a rupture jumping from one fault to another as an example. Ideally we would have a formula that provides a jumping probability as a function of fault separation, relative geometry, sense of slip, slip rates, hypocenter, etc. Because no such formula yet exists, however, we may instead be forced to rely on expert judgment based on a case-by-case analysis of each neighboring fault combination in California. Clearly, a formula giving a fault jumping probability is more testable than relatively subjective expert opinion, as the former could be formally tested against events occurring worldwide. This is not to denigrate expert opinion however, as it can be a powerful way of assimilating complex information and diverse information on how nature operates.

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 catastrophe bonds. Long-term forecasts could (and do) influence 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 almost 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 phenomena, it has been
demonstrated that they can produce significant probability changes over periods of years to decades. By considering all time dependencies within one model framework, we will be able to develop a consistent set of forecasts.

The utility of UCERF3 will be dictated by not only what the user community is interested in applying, but also by the confidence we have in the forecast given uncertainties. Therefore, it will be important to have an ongoing dialogue between potential users and model developers throughout the duration of the project. This will help to clarify both their needs and our ability to deliver something meaningful given present knowledge. Use in earthquake insurance will certainly be a priority given CEA’s financial support. However, there are other potential uses as well, like as a resource to help the California and/or National Earthquake Prediction Evaluation Councils advise on earthquake threats following significant events. The USGS currently makes short-term forecasts during earthquake clusters for users including the California Emergency Management Agency, other emergency responders, and utilities. Our operational model will improve these short-term forecasts by making them consistent with the long-term model.

It is also important to note that perceived needs of the user community should not be the sole driver of our priorities. The USGS ShakeMaps are a good example of a product whose usefulness was not fully anticipated in advance. In fact, until we have both a time-dependent, operational model and the tools with which to explore loss implications, no one will really know what’s important for users. This again emphasizes the need for ongoing dialogue from the very beginning.

Feedback from potential users may very well focus our efforts on either shorter-term or longer-term time dependencies. Our scientific understanding may also point us in a particular direction, as we would not want to spend a lot of time building a model with such large uncertainties that it’s rendered useless. That being said, there are also good scientific reasons for attempting to construct a “broadband” time-dependent model. As discussed above, there is not a physically meaningful division between short-term and long-term forecasts, and attempting to draw such a line may present more problems than it solves. Another is that there’s no better way to highlight the important, deeper scientific issues than to actually attempt to build a system-level model. An example of such a question in the context of our present goals is “what’s the difference between a multi-fault rupture and a separate earthquake that happened to be triggered quickly?” Our attempting to build a broadband model will both identify and stimulate progress on resolving fundamental scientific issues.

Finally, it’s important to emphasize that we endeavor to build a model that could be used in an operational sense. This should not be interpreted as a commitment or promise to build a model that will be maintained operationally, as we do not yet know the feasibility or what resources will be required for such ongoing maintenance (let alone the full scope of legal issues).

**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. The previous WGCEP efforts have repeatedly reinforced the truthfulness of “it’s easier said than done” and “the devil is in the details”. In the worst case, the WGCEP may conclude that the best available science has not yet provided a representation of
multi-fault ruptures and/or spatial-temporal clustering that is adequate for operational purposes. A project plan has been developed to deal with these uncertainties. In particular, the UCERF3 logic-tree structure will be capable of handling this situation by using model branches developed for UCERF2 as appropriate fallbacks.
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

As with UCERF2, UCERF3 will be constructed from the four main model components shown and defined in Figure 1. We acknowledge that dividing any complex interactive system into separate components has some degree of artificiality and arbitrariness. Nevertheless, we believe those established here are both meaningful and necessary, at least for the time being. Where the distinction may become problematic 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 does have both physical meaning (in terms of being conceivably measurable) and practical use (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? Furthermore, physics-based earthquake simulators, which are discussed more 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 Earthquake Rate and Probability models in what follows may dissolve at some point as UCERF3 is developed.

Fault Models

Fault models give the spatial geometry of the larger, known, and active faults throughout the region, with alternative models representing epistemic uncertainties. The faults are divided into “Fault Sections”, each of which is composed of:

- 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

where distinct faults section are defined only to the extent that these attributes vary along strike, so some fault sections can be quite long (e.g., the northern San Andreas Fault has only four sections).

For UCERF3 we will update and revise the fault models developed for UCERF2 (going from fault model versions 2.1 and 2.2 to versions 3.x).
Because one of the goals in UCERF3 is to include multi-fault ruptures, reconsideration of fault endpoints is particularly important, especially since most faults were not originally mapped with this issue in mind. This and other important tasks associated with the development of the new fault models are described in Table 1.

**Tasks:**

**Table 1. Task for the development of Fault Models.**

Please note the following: Task leaders are listed in bold typeface and the “participants” represent current possibilities with no contractual obligation implied; those tasks that warrant further discussion are also described in an appendix (but only if further discussion is in order at this time); some tasks may change and others may be added as the project evolves; some tasks could just as appropriately been associated with one of the three other model components (the choice was arbitrary); and finally, USGS Western Region participants are listed in red and USGS Central Region participants are listed in Blue (not including Jones or Parsons) for accounting purposes.

<table>
<thead>
<tr>
<th>Fault Model(s)</th>
<th>Task</th>
<th>Description</th>
<th>Leader &amp; Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1) Update Faults</td>
<td>a) Revise or add any new, important faults (even if slip rates are not well constrained because deformation modeling might help define these). Also remove any faults that are no longer considered viable. This will include results from PG&amp;E supported work and will be coordinated with the SCEC statewide community fault model development. See appendix for further discussion of this task.</td>
<td>Dawson, J. Shaw, Wills, Weldon, Haller, Grant, Powers, Parsons, and Murray.</td>
</tr>
<tr>
<td></td>
<td>F2) Reevaluate Fault Endpoints</td>
<td>b) Reconsider fault endpoints given importance of this for multi-fault ruptures. Endpoints should be evaluated using a synthesis of available geologic, geophysical, and seismological data and this uncertainty will be characterized in the fault database. One question is how to represent these uncertainties (e.g., for storage in the database). This task is potentially endless, so it will be important to prioritize and maintain realistic expectations. See appendix for further discussion of this task.</td>
<td>Dawson, Parsons, Powers, J. Shaw, Plesch, L. Grant, A. Michael</td>
</tr>
<tr>
<td></td>
<td>F3) Database Issues</td>
<td>d) Maintain and update the California Reference Fault Parameter Database in order to serve the needs of UCERF3. The Fault-Section database will be updated with new and revised faults as described in Task F1. The database will also be expanded to accommodate the results of Task F2 (Re-evaluation of Fault Endpoints). The Paleo Sites database will be maintained and updated, with an emphasis on storing geologic slip rates, in order to provide a database for comparison to the geodetically-based deformation models, as well as a resource for the effort to examine the data for potential biases in slip rates used in the UCERF2. Coordination between UCERF3 and the USGS NSHMP will be necessary in order to ensure revisions to the UCERF3 database are integrated into the NSHMP database. See appendix for further discussion of this task.</td>
<td>Dawson, Weldon, Field, Haller, Petersen, Wills, Jordan, McCarthy, Powers, Biasi, Parsons.</td>
</tr>
</tbody>
</table>
Deformation Models

Each deformation model gives slip-rate estimates at various locations within 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, and the off-fault deformation was represented with a set of geographic polygons with an associated residual slip rate (e.g., Figure 2). 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.
Figure 2. Faults and slip rates for Deformation Model 2.1, plus the polygons (green) representing significant deformation elsewhere in the region.

For 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. One big question for UCERF3 is whether the more quantitative models, such as Peter Bird’s NeoKinema (e.g., Bird, 2009), can now be used in place of expert opinion (the topic of a SCEC workshop on April 1-2, 2010).

Another important question is the overall amount of seismic deformation occurring off the explicitly modeled faults. 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). One question is whether we should reduce such fault slip rates by some fraction in order to avoid double counting with respect to the off-fault seismicity that will get added to the Earthquake Rate Model. In fact, NeoKinema suggests that the off-fault deformation is 30% (which implies we may have over-estimated some fault slip rates in UCERF2).

Tasks:

This and other important tasks associated with the development of the new deformation models are described in Table 2.

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
<th>Leader &amp; Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1) Evaluate New Deformation Models</td>
<td>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 (&amp; 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 improve the Earthquake Rate Models.</td>
<td>Thatcher, Zeng, Hearn, Johnson, and Sandwell.</td>
</tr>
</tbody>
</table>
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 2)?

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)?

See the report from the April 1-2, 201 Workshop for a broader discussion, or the appendix here for an outline of next steps.

D2) B-fault bias?

e) Evaluate whether B-fault geologic slip rates are biased (always rounded up, or always given some minimum/default value?). See appendix for further discussion of this task.

Weldon, Dawson

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

The final product here will be slip-rate estimates at various locations within each fault model (ideally at 5-10 km intervals along the faults), plus perhaps the spatial distribution of off-fault deformation (and hopefully more refined than the large polygons used in UCERF2). Again, it will be important to have a range of deformation models in order to represent uncertainties with respect to the above issues (versions 3.1, 3.2, ...). 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). It will also be important to state whether the slip rates represent deformation on the main fault surface or whether they apply to a zone surrounding the fault. These deformation models will also be of direct use to the physics based earthquake simulators, so we are coordinating accordingly.

**Earthquake Rate Models**

The goal here 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). For now these models will include aftershocks (which will get removed later in the Earthquake Probability Models as necessary). As stated above, our primary aim with UCERF3 is to relax
segmentation and include multi-fault ruptures. 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 in UCERF2, here we endeavor to solve for the rates of all events simultaneously. For this purpose we will attempt to use the formal inverse approach outlined by Field and Page (2010), which builds on the work of Andrews and Schwerer (2000). We will also explore the methodology outlined by Parsons and Geist (2009), which uses a Monte Carlo technique rather than a formal inversion. David Jackson’s group is also pursuing and approach that we will consider, although we don’t yet have enough details to include their approach in this report.

We start with a semi concise statement of how we intend to solve this problem, and then follow with a discussion of several important issues.

**Methodology:**

First, let's consider only those ruptures that occur on the faults defined in our Fault and Deformation Models (we’ll add the background/off-fault seismicity below). Here we will only strive to model events that have a rupture length greater than or equal to the seismogenic thickness (relegating smaller events to the off-fault seismicity discussed below). In order to relax segmentation, we first subdivide each fault section into a number of equal-length subsections, where the subsection lengths are just less than or equal to half the seismogenic thickness (Figure 3). We now have $S$ subsections (e.g., $S=1598$ for the example shown at the bottom of Figure 3).
Figure 3. Top: UCERF2 fault sections for Deformation Model 2.1. Bottom: results of dividing sections into an integer number of equal-length subsections (that have lengths equal to, or just less than, half the section’s seismogenic thickness). Note that all subsections in green are connected to all others in green without jumping more than 5 km between faults.
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 exact distance being an adjustable parameter). The minimum number of two subsections ensures that rupture lengths are at least about the seismogenic thickness (since subsection lengths are about half that). Note in Figure 3 that nearly all the subsections in southern California are connection to all others without jumping more than 5 km between subsections (illustrating the high connectivity of the fault network). Here we further filter the set of ruptures by requiring that:

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

(see the discussion of Task R5 in the appendix for further details). Note that by “viable” we mean within the realm of possibility (or passing the “laugh test”). This does not imply that all ruptures are equally likely, which will be accounted for later. We now have the total set of \( R \) viable fault ruptures, and we now want to solve the long-term rate or frequency \( (f_r) \) of each \( r^{th} \) rupture.

We can define the participation matrix \( G_{sr} \) as specifying whether the \( r^{th} \) rupture utilizes the \( s^{th} \) subsection (having a value of 1 if so and 0 if not). We can also define the slip matrix \( D_{sr} \) as giving the average amount of slip on the \( s^{th} \) subsection in the \( r^{th} \) event, and by average we mean over multiple occurrences of the rupture (how \( D_{sr} \) is actually computed is discussed below). We can now define the first two equation sets in our inverse problem as:

\[
\sum_{r=1}^{R} D_{sr} f_r = v_s, \quad \text{Equation Set (1)}
\]

where \( v_s \) is the subsection slip rate (where known), and

\[
\sum_{r=1}^{R} G_{sr} P_{sr}^{\text{paleo}} f_r = f_s^{\text{paleo}}, \quad \text{Equation Set (2)}
\]
where \( f_s^{\text{paleo}} \) is a paleoseismically inferred event-rate estimate (where known) and \( P_r^{\text{paleo}} \) is the probability that the \( r \)th rupture would be seen in a paleoseismic trench. The first equation set simply ensures that the model satisfies slip-rate data (classic “moment balancing”), and the second ensures that the model is consistent with paleoseismic observations.

We can also add a-priori constraints on any known rupture rates using

\[
 f_r = f_r^{a-\text{priori}}. \quad \text{Equation Set (3)}
\]

This allows us to force, for example, the rate of Parkfield events to be the historically observed average of one every \( \sim 25 \) years (Bakun et al., 2005).

To further condition the inversion we can also add smoothing constraints via

\[
 f_r - f_r' = 0. \quad \text{Equation Set (4)}
\]

This forces the rate of rupture \( r \) to be the same as that of \( r' \) (unless the data constraints above dictate otherwise). For example, this would allow us to prevent abrupt changes in the rate of an M 6.8 event along the S SAF (unless, again, slip rates and/or paleoseismic event rates say otherwise). We tentatively plan to apply this constraint only along (and within) well-defined faults such as the SAF, SJF, and Calaveras (and not for ruptures that involve multiple faults). One could argue whether such smoothness exists in nature. However, even if it doesn’t, smoothness may still be appropriate in terms of generating an average model for policy purposes (until data can resolve where actual variability exists).

We now introduce equations that couple the fault-based rupture model to the “off-fault” or “background” seismicity. Explicit goals here are to prevent an over-prediction of M 6.5 to 7.0 events (the UCERF2 bulge problem) and to more seamlessly integrate the fault model and background seismicity. The basic idea is to force the combined models to match a Gutenberg Richter distribution in different sub-regions of the state. In what follows we will abbreviate “magnitude frequency distribution” with “MFD” and “Gutenberg Richter” with “GR”.

For a given sub-region defined by a polygon, we can obtain the a-value for the GR distribution from historical and instrumental seismicity. Assuming a b-value of 1.0 (or some other reasonable value) we can then define the sub-region’s GR distribution out to arbitrarily large magnitude. The constraint in the inversion is simply that the total MFD for all the fault-based ruptures that nucleate in this region must be less than or equal to this GR distribution. We write this equation set as
where each element in the 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 the $r^{th}$ rupture that falls within the $g^{th}$ sub-region, and $GR_m^g$ represents the GR rate of the $m^{th}$ magnitude bin in the $g^{th}$ sub-region.

Once the inversion has been solved, the MFD for off-fault seismicity is just the 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 4a). However, there is nothing stopping the rates of off-fault seismicity from being negative, which would occur if the inversion just couldn’t get the fault-based MFD below the regional GR constraint (implying, for example, there aren’t enough large, multi-fault ruptures to push moment onto). In this case we could set the negative rates in the off-fault MFD to zero, and let the consequent bulge remain. Or better yet, we could examine whether there might be something wrong with the inventory of faults or slip rates in the deformation model, or whether the sub-region’s a-value might be biased due to temporal variability.

If the resultant maximum magnitude of the background-seismicity is uncomfortably low, as illustrated in Figure 4b, 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 could be achieved by replacing $GR_m^g$ in Equation Set (6) 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 4b 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 4c). This would be consistent with UCERF2 generally applying an $M_{max}$ of 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 could then be partitioned among the 0.1 by 0.1 degree background-seismicity grid cells according a spatially smoothed distribution of a-values (as in UCERF2).

As noted, the GR constraint of Equation Set (6) can be applied over different sub-regions and at different scales. Applying it only to the statewide region would not prevent, for example, northern California from having a strong bulge as long as southern California had a commensurate anti-bulge. Applying the constraint at a 0.1 by 0.1 degree grid-cell resolution would effectively force all points in space to be perfectly GR (which adherents of the characteristic MFD on faults would object to). Therefore, we tentatively plan to apply this constraint to 1.0 by 1.0 degree sub-regions (e.g., the grids lines in Figure 3), but the optimal choice will certainly be a subject of exploration.
Figure 4. 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 can’t 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 $M_{\text{max}}$ (arbitrarily set at M7.2 here).
Implementation Considerations:

The equations in the inversion can be weighted differently based on data uncertainties and/or how precisely the constraint needs to be met.

It will be important to consider whether the slip rates in Equation Set (1) should be reduced to account for the moment released in sub-seismogenic thickness ruptures, or because the slip-rates apply to a broad deformation zone rather than the main fault surface. Note also that the inversion could conceivably handle slip rates that apply to more than one fault (e.g., from larger scale block modeling, or even between GPS station pairs). The exact transition of slip rates at the ends of faults that neighbor each other will likely be important with respect to multi-fault rupture rates.

The description of Equation Set (1) above glossed over challenging details in defining the slip on the $s^{th}$ subsection in the $r^{th}$ rupture ($D_{sr}$). The traditional approach (e.g., WGCEP, 2008; Field and Page, 2010) is to compute the magnitude from the total rupture area, get the average slip from the moment-magnitude relationship, and then distribute this average slip over the subsections involved using a model of how slip varies, on average, along rupture (e.g., the “tapered” square-root-of-sine model of Weldon et al. (2007)). One important question is how average slip varies along strike for multi-fault ruptures (i.e., does slip pinch out at the intersection or does a single taper apply to the entire rupture?). This question is being addressed as part of Task R1, listed in Table 3 below and discussed in the appendix. The answer, coupled with slip-rate transitions at the ends of faults, will have a huge impact on the rate of multi-fault events obtained by the inversion.

Another question is whether larger events actually penetrate below the depth of microseismicity, and more generally, how average slip varies with depth? Our previous models have assumed that slip does not penetrate below the depth of microseismicity (shown schematically in Figure 5a below, which is from King and Wesnousky (2007)). If slip actually penetrates deeper for larger events, but also remains constant as a function of depth above, then we have been overestimating $D_{sr}$ and thereby under-estimating the rates of such events. Here the usual practice of moment-rate balancing becomes problematic because depth of rupture is magnitude dependent. Things are more complicated if slip both penetrates deeper and also exhibits a depth-dependence like that shown in Figure 5b. In this case we may be better off slip-rate balancing at an intermediate depth of say 6km (where aseismic processes are at a minimum). The exact influence of this on event rates will depend on the distribution of slip with depth. Issues related to this are being addressed in Task R2, listed in Table 3 and discussed in the appendix.
Figure 5. From King and Wesnousky (2007). a) Average fault slip is constant down to the lower seismogenic depth (defined by the depth of microseismicity). b) an alternative model where slip both penetrates deeper in larger earthquakes, and is depth dependent.

As presented above, the inversion treats all “viable” multi-fault ruptures equally (apart from how slip-rate variations at the ends of faults, and the distribution of average slip along the rupture, influences the final rates). In other words, a multi-fault rupture that has to jump 5 km is not penalized relative to one that has to jump only 2 km. If we had some model that gave us the relative probability of such ruptures, then these could be added to the inversion (e.g., by setting the a-priori rates to 0 in Equation Set (3) & adjusting their relative weights accordingly). Task R6, listed in Table 3 and discussed in the appendix, is exploring different ways such multi-fault rupture probabilities could be defined.

We intend to solve the mathematical inversion using a simulated annealing approach being developed by Morgan Page (Task R5) rather than using the Non-Negative Least Squares solution (Lawson and Hanson, 1974) as applied by Field and Page (2010). The primary reason is that simulated annealing gives a range of models that sample the solution space, which will be important since our problem will at best be mixed-determined. The range of models obtained, each of which will be more or less segmented or characteristic on specific faults, will hopefully form a rational basis for constructing logic-tree branches. We think this will allow us to represent a broader range of models than has been included in the past.

It should be noted that several other constraints could be added to the inversion. For example, we could prevent the rates of events on specific faults from exceeding the historically observed rate (e.g., if the model would otherwise imply a very high probability that we should have seen more events than we have). We could also impose a smoothness constraint on each of the
various MFDs (on faults or in sub-regions) to avoid jagged peaks and valleys. Other types of smoothness constraints could also be added. However, even if warranted conceptually, adding these will increase the computational demands, so a more efficient approach may be to leave these out and achieve the same smoothing effects by averaging over the different models generated by the simulated annealing algorithm.

We don’t yet know whether this inversion is computationally feasible. It will presumably help to have a good starting model, which we can construct by assigning each rupture an initial rate equal the total regional rate (in the associated magnitude bin) divided by the total number of ruptures in that bin. These initial rates could also be weighted according to some average slip rate over the fault sections utilized by each rupture. If computation demands are still an impediment, then we can either downsize the problem or explore the Monte Carlo approach of Parsons and Geist (2009).

Conceptually, the approach outlined here is much simpler, more objective, more reproducible, and more unified than that adopted in UCERF2. Note in particular that we no longer have type-C zones, nor Type-A versus Type-B faults. Granted, the solution space is more unwieldy here, and will somehow have to be reduced to a finite number of logic-tree branches (e.g., by averaging separate families of models).

**Tasks:**

Tasks on the above, and other issues related to the development of UCERF3 earthquake rate models are described in Table 3.

<table>
<thead>
<tr>
<th>Task Description</th>
<th>Leader &amp; Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1) Evaluate Along-Strike Distribution of Average Slip</td>
<td>Biasi, Weldon, Dawson, &amp; Wesnousky.</td>
</tr>
<tr>
<td>R2) Evaluate Magnitude-Scaling Relationships and Depth of Rupture</td>
<td>B. Shaw, Hanks, Somerville, Page, Beeler, Wenousky, Seok Goo Song</td>
</tr>
</tbody>
</table>

Table 3. Task for the development of Earthquake Rate Models.

Please see the notes given for Table 1, as they apply here too.
<p>| R3) Paleoseismic Recurrence Interval Estimates | Update and/or add to Tom Parsons’ compilation of mean recurrence interval (MRI) estimates for paleoseismic sites. We also need to consider independence of these from slip rate estimates at the same locations. How gray can the literature be for sites in this compilation? <strong>LEADERSHIP WILL BE RESOLVED WHEN WE DECIDE ON EXACTLY WHAT WE NEED. See appendix for further discussion of this task.</strong> | Parsons or Biasi, Weldon, Dawson, &amp; Grant-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. Should these models be defined on trench-by-trench basis (to account for the unique attributes of each paleo site)? | Weldon, Petersen, Biasi, Dawson. |
| R5) Solve the Large Inversion Problem | Determine whether we can solve the large inverse problem (solving for the rate of all events), including the ability to adequately sample the solution space. Here “moment-balancing” may be replaced with slip-rate balancing at a specific depth (e.g., at a depth where overall aseismic slip is a minimum) if we can get the model mentioned above for average slip as a function of depth and magnitude. <strong>See appendix for further discussion of this task.</strong> | Page, Field, Parsons, Jordan |
| R6) Fault-to-Fault Jumping Probabilities | Develop models of fault-to-fault jumping probabilities (relative to the likelihood of through-going rupture if there were no separation, no change in strike or faulting style, or no change in whatever metric is used to define the probabilities). What’s needed here is a review of the literature (both observational and theoretical), a recommended applicable model based on this review, and a research agenda for making further progress. 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 how we quantify and utilize uncertainties in fault endpoints? Finally, do we develop generic rules or do we consider each possible connection in our fault model on a case-by-case basis. Three approaches are currently being taken: 1) develop rules based on observations; 2) explore whether coulomb calcs can define relative likelihoods (based on mechanical self-compatibility); and 3) use full dynamic rupture modeling on carefully defined examples. <strong>See appendix for further discussion of this task.</strong> | Harris, Jackson, Wesnousky, Dawson, &amp; Biasi, J. Shaw? |
| 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. <strong>See appendix for further discussion of this task.</strong> | Parsons, Bakun? |
| R8) Reevaluate Earthquake Catalog | Reevaluate 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 decluster 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? Another question is whether we can estimate the magnitude frequency distribution for “off fault” events. <strong>See appendix for further discussion of this task.</strong> | Michael, Felzer, Bakun, Parsons, Schorlemmer, Hardebeck |
| R9) Smoothed | Reevaluate procedures for smoothing instrumental seismicity in | Felzer, Mueller, Biasi, |</p>
<table>
<thead>
<tr>
<th>Task ID</th>
<th>Task Description</th>
<th>Responsible Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>R10) Mmax for off-fault seismicity</td>
<td>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’s needed to satisfy the extra deformation in our previously defined type-C zones (or any new zones defined by the more sophisticated 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.</td>
<td>Parsons, Field, Michael</td>
</tr>
<tr>
<td>R11) Focal mechanisms of off-fault seismicity</td>
<td>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.</td>
<td>Jackson, Hauksson</td>
</tr>
<tr>
<td>R12) Distribution of Slips in Paleo Trench</td>
<td>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); they keep promising something but have not yet delivered. See appendix for further discussion of this task.</td>
<td>Weldon, Hecker, Dawson, &amp; Biasi</td>
</tr>
<tr>
<td>R13) Evaluate Physics Based Earthquake Simulators (for rate estimates)</td>
<td>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.</td>
<td>Field, Michael, Tullis, Dieterich, Richards-Dinger, Ward, Rundle, Pollitz, Beeler</td>
</tr>
<tr>
<td>R14) Reconsider aleatory uncert. in Mag from given Area</td>
<td>Currently we give 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 is a very important issue for SCEC’s CyberShake project, and may contribute to the precarious-rock problem as well. We need to have a cooperative workshop CyberShakers, NGA developers, and other ground-motion modelers to address this issue.</td>
<td>Field, Campbell, Graves, others?</td>
</tr>
<tr>
<td>R15 Cascadia subduction zone</td>
<td>Develop a complete, revised model for Cascadia. Note that this component will be developed somewhat 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 is hosting a meeting on Cascadia in Corvallis, OR on Nov 18-19, 2010. See appendix for further discussion of this task.</td>
<td>Frankel, Weldon, Petersen</td>
</tr>
</tbody>
</table>
Earthquake Probability Models:

The earthquake probability models give the probability that any one of the events in the long-term Earthquake Rate Model will occur over a specified time span (using, for example, information on the date of the last event). Our main goals here for UCERF3 are the following: 1) try to resolve the interpretation of the “Empirical Model” (the apparent rate change in microseismicity); 2) develop self-consistent elastic rebound models; and 3) apply spatial-temporal clustering models.

Resolve Interpretation of Empirical Model:

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, and for the central, and southern parts of the state. The rate decrease cannot be documented in the Mojave Desert because of lack of historic data, but recent high seismicity rates in that region are suggestive that in the Mojave rates have increased rather than decreased.

The decrease in seismicity rate in the San Francisco Bay Area has long been recognized, and has traditionally been attributed to a static stress shadow imposed by 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. This has lead to some weight being placed on the “Empirical Model”, which involved applying the observed seismicity rate changes to the fault-based sources in the long-term Earthquake Rate Model.

Two main issues with respect to the empirical model need to be addressed by UCERF3. The first is verifying the size of the seismicity rate change by re-evaluating the historic catalog. Since re-evaluation of all of the historic intensities would be a monumental task, the effort might focus on re-evaluating the errors associated with historic 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 historic 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-Hayward-San Jacinto fault system; the contrast in seismic activity here between the historic and instrumental catalogs is striking. 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 and/or time dependent changes in the background rate can allow the ETAS model to work, and whether such changes are physically reasonable and applicable to modeling future behavior.
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 relatively straightforward when a fault is assumed to obey strict segmentation in the sense that no multi-segment ruptures occur (e.g., WGCEP, 1988, 1990). However, the calculation is not straightforward when multi-segment ruptures are included, in essence because we are attempting to apply a point-process model to what is clearly non a point process.

The methodology of WGCEP (2003) was applied by WGCEP (2008) in computing elastic-rebound probabilities for UCERF2. One way to describe the WGCEP (2003) approach is that they first computed the probability that each segment will rupture (from the long term-rate and date of last event, and assuming a BPT distribution) and then partitioned these probabilities among all ruptures that could be triggered by the segment. However, and as discussed extensively in Appendix N of the UCERF2 report, this methodology has a self-consistency problem that manifests in a variety of ways. One manifestation is that final segment probabilities (aggregated over all ruptures) are not equal to the segment probabilities computed and applied in the first place. Another manifestation, as revealed by Monte-Carlo simulations, is that the distribution of segment recurrence intervals implied by the model is not the same as that initially assumed (exemplified in Figure 6, taken from Appendix N of the UCERF2 report). 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 the short recurrence intervals on that segment that are inconsistent with the distribution assumed in the first place (Figure 6). This also leads to a bias in the simulated rate of events compared to the long-term rate (about 3% for the UCERF2 example in Figure 6a). These problems increase with the number of segments a fault is divided into.

WGCEP (2008) applied the WGCEP (2003) methodology in spite of these shortcomings because 1) they lacked an alternative; 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 a Bayesian perspective that probability is a statement of “the degree of belief that an event will occur” (D’Agostini, 2003)).
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 SSAF as modeled for UCERF2. b) An example for an 80 km fault with 5-km segments (essentially un-segmented) with a Gutenberg-Richter distribution of events. Both examples are taken from Appendix N of the UCERF2 report.

Unfortunately these problems worsen as the fault is divided into more and more segments, and especially if we relax segmentation altogether (as we plan to do for UCERF3). Figure 6b shows the results for a simple “un segmented” 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-base simulators (Ward, 2000; Rundel et el. 2006; and Dieterich and Richards-Dinger, 2010), as these models both embody elastic rebound and make no assumptions regarding segmentation. Of course we don’t know whether any simulator is correct, so it will be important to test any such statistical behavior we infer for robustness against the range of simulator results, as well as against actual observations to the extent possible.

If a fault does not obey segmentation, then it would seem impossible for all points on that fault to honor a perfect renewal-model distribution such as BPT or Lognormal, 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 the Dieterich and Richards-Dinger (2010) simulator, where the entire northern California fault system has been modeled. This plot looks nothing like the usual renewal-model distributions.
Figure 7. The distribution of M$\geq$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 (details available upon request).

Stated another way, even if we had perfect knowledge of the recurrence-interval distribution at one or more points on a fault (like in Figure 7), it’s not clear how to turn this information around into rupture probabilities in 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 now present an alternative procedure that so far looks promising, although more work is needed to justify its application. 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. One sensible approach would be to apply an average time-predictable model:

$$T_r^{pred} = \sum_{s=1}^{S} \left( \frac{D_s^{last}}{V_s} + T_s^{last} \right) = \sum_{s=1}^{S} \frac{D_s^{last}}{V_s} + \sum_{s=1}^{S} T_s^{last} = \Delta T_r^{pred} + T_r^{last}$$
This states that the predicted time of this \( r \)th rupture (\( T_{r}^{pred} \)) is simply the average time at which the slip rate (\( v_r \)) 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 as such, this represents a straightforward generalization of the “time-predictable” model introduced by Shimazaki and Nakata (1980).

The above equation can be rewritten as

\[
T_{r}^{pred} = \Delta T_{r}^{pred} + T_{r}^{last}
\]

where

\[
\Delta T_{r}^{pred} = \frac{\sum_{s=1}^{N} D_{s}^{last}}{v_r} = \frac{D_{r}^{last}}{\bar{v}_r}
\]

where \( D_{r}^{last} \) and \( \bar{v}_r \) are the slip-in-last-event and slip rate, respectively, averaged over the subsections involved in the \( r \)th rupture, and

\[
T_{r}^{last} = \frac{\sum_{s=1}^{N} T_{s}^{last}}{S}.
\]

Ideally we would have enough observational data to examine the extent to which these predicted intervals, \( \Delta T_{r}^{pred} = T_{r}^{pred} - T_{r}^{last} \), agree with observed intervals: \( \Delta T_{r}^{obs} = T_{r}^{obs} - T_{r}^{last} \) (where \( T_{r}^{obs} \) is the occurrence time of an actual event). What we do have, however, are synthetic catalogs produced by physics-based simulators. From these we can examine the distribution of the ratio of observed to predicted intervals (\( \Delta T_{r}^{obs} / \Delta T_{r}^{pred} \)), where “observed” here means that produced by the simulation, and predicted is that computed from the above equations.

Before showing results, we need to address an important question of what events reset the clock with respect 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 also makes sense in that the fault-based ruptures in our long-term Earthquake Rate Model are restricted to events 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 \( M \geq 6.5 \) 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
Lognormal 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, leading to Poisson recurrence-interval statistics.

Figures 8a and 8c show the same results as in Figure 8a, but where the Ward (2000) and Rundel et al. (2006) simulator results, respectively, have been used. The agreement between simulators in Figures 8 is encouraging, as all seem consistent with a BPT or Lognormal distribution with a COV between 0.23 and 0.3. Note in particular that there are zero short recurrence intervals in Figure 8, implying that M≥6.5 “aftershocks” or triggered events never occur within the rupture surface of larger main shocks (at least not in these simulator results).

The preceding discussion assumes we know exactly where the next rupture will occur (leaving only the question of when). Of course we don’t know which of the many ruptures in our long-term model will be the next to occur, so we must now consider that it could be any event. We do this by using the following formula:

\[ P_r = P_r^{\text{Pois}}(\Delta T_{\text{longTerm}}) \cdot \frac{P_r^{\text{BPT}}(\Delta T_{r_{\text{pred}}}, \overline{T}_r_{\text{last}}, \text{COV})}{P_r^{\text{Pois}}(\Delta T_{\text{pred}})} \]

where \( P_r^{\text{Pois}}(\Delta T_{\text{longTerm}}) \) is the Poisson probability given the long-term rate of the \( r^{th} \) rupture, and \( P_r^{\text{BPT}}(\Delta T_{r_{\text{pred}}}, \overline{T}_r_{\text{last}}, \text{COV}) \) and \( P_r^{\text{Pois}}(\Delta T_{\text{pred}}) \) are the probabilities computed from the BPT and Poisson models, respectively, assuming the \( r^{th} \) rupture is the next to go as described above. Note that the long-term Poisson probability of a rupture goes toward zero as the fault is represented with a greater and greater number of smaller and 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 will give lower probabilities for events in areas that have recently ruptured, and higher probabilities where they haven’t (essentially filling the gaps). It will also permit some spatial overlap for events that occur in close temporal proximity (e.g., for cascading sequences like on North Anatolian fault).
Figure 8. Distribution of \( \frac{\Delta T_{\text{obs}}}{\Delta T_{\text{pred}}} \) (described in the text) obtained from the physic-based simulators as labeled. The input model used in these simulator runs is the same as in Figure 7. Shown with a black and blue lines are best-fit BPT and Lognormal 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.
Monte Carlo simulations conducted previously (with the model shown in Figure 6b) demonstrate that this methodology is relatively unbiased in terms of event rates and moment rates, but it will be important to reconfirm this once we’ve implemented the methodology for the full UCERF3 model.

A seemingly reasonable alternative implementation would be the equivalent average slip-predictable model, where $D_{\text{last}}$ above is replace with $D_{\text{next}}$ (the slip in the next event). However, Monte Carlo simulations reveal a significant bias with this approach. Essentially the procedure preferentially chooses smaller events earlier in the cycle, thereby skewing both 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 $D_{\text{last}}$. Where this is unknown, we may be able to quantify the associated epistemic uncertainties for $D_{\text{last}}$ from the long-term model, the assumed time dependence, and the observed open interval. Whether such uncertainty bounds provide any value added compared to a default Poission model remains to be seen.

At this point we want to clearly state that more work will be required over the next year to further justify this approach. This will include close scrutiny of the physics-based simulators we use to support the methodology. To this end we are currently working with the SCEC Simulator Working Group being led by Terry Tullis, and we plan to have a community-wide workshop to discuss simulator applicability in Spring 2011. Some of the ways the simulators are being evaluated are discussed in the appendix entry for Tasks R13 and P9.

Some might express a lack of surprise that simulators imply elastic-rebound predictability given it’s built into these models. The salient question seems to be, however, whether the interactive
complexity of the fault system can mask any such elastic-rebound predictability. The preliminary answer implied by Figure 8 is no, but it will be important to see whether simulators can be tuned to get rid of it (while still honoring other observational constraints such as regional Gutenberg Richter). Even if all simulators agree that elastic rebound is apparent under all reasonable parameter settings, it’s of course possible that all those models are wrong. If so, it will be interesting to hear an articulation of how they are getting the physics wrong.

Other things that remain to be explored with the methodology presented here include:

- What conditions most effectively reset the clock on each subsection (e.g., magnitude threshold, down-dip width of rupture)?
- Predictability differences between different faults or fault sections?
- How can we formally test this methodology, either against real observations or against the simulators themselves?
- 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 present in UCERF3, then so far this is the most promising approach we know of.

**Apply Spatial-Temporal Clustering Models:**

Our goal here is to include spatial-temporal clustering, in essence to acknowledge that “aftershocks” can be large and damaging. A good example of such triggering is the Joshua Tree, Landers, Big Bear, and Hector Mine sequence that occurred in/near the Mojave Desert in the 1990s. According to UCERF2, this chain of events was 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 for exploratory purposes.

Because the exact physical process responsible for earthquake triggering remains controversial (e.g., Felzer and Brodsky, 2006; Richards-Dinger et al., 2010), we feel the most justifiable application for UCERF3 will be empirically based clustering models (e.g., Ogata, 1988; Reasenberg and Jones, 1989, 1994;). The Short Term Earthquake Probability (STEP) methodology of Gerstenberger et al. (1995) represents the most official earthquake-triggering
model developed to date. Their approach applies aftershock statistics to revise earthquake probabilities in real time for M≥3 events that unfold throughout the region. The model we propose here for UCERF3 is inspired by STEP, and attempts to build upon that methodology. Because a complete description of STEP is beyond the scope of this report, we instead list the attributes of STEP that will be contrasted with our proposed model:

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

2. Triggered events are sample from a Gutenberg Richter distribution between M 5 and 8 everywhere in the region, which is inconsistent with the underlying long-term model (which says, for example, that M 8 events can only occur on a few faults like the San Andreas).

3. There is nothing stopping an M 8 event from immediately triggering itself. In fact, the likelihood of any given an event in STEP will be the 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 combines different models based on a sophisticated analysis of generic, sequence-specific, and spatially variable parameters. While this may indeed improve predictability (RELM tests are thus far inconclusive according to D. Schorlemmer as of Nov 1, 2010), it comes at a cost of significant complexity. It also makes tracking aleatory versus epistemic uncertainties a challenge, especially since the combination of models is spatially variable.

We now outline the procedure we’re hereby proposing for UCERF3, and discuss how it addresses the issues with STEP mentioned above. One thing to keep in mind is that our earthquake Rate Model gives us 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 nucleations, the rates of finite-fault ruptures 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 lat-lon 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 spatial-temporal clustering model include:

a) For a given start-time and forecast duration, we collect all previously observed M≥2.5 events, plus randomly sample spontaneous (non-triggered) events from our long-term earthquake Rate Model (and including any elastic rebound event-probability modifications as described in the previous section). 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) using the generic CA parameters from Hardebeck et al. (2008):

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

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 shown to the left below), plus a spatial decay of \( R^n \) from the main shock fault surface (shown in the middle), 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 of 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:

```
Ruptures that nucleate in that bin
```

We now need to collect secondary aftershocks from 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 stable 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 M 8 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. Finally, updating the model based on ongoing seismicity will delineate any “blue” versus “red” lobes that would be present if static stress changes are important.

One thing to note is that this algorithm generates suites of synthetic catalogs, each of which represents a viable sequence of triggered events. This is good in that loss modeling is generally conducted using such synthetic catalogs (referred to as “stochastic event sets”) in order to account for the spatial correlation of ground motions across a portfolio of sites. What’s new for
loss modeling here is that UCERF3 event sets will include spatial-temporal clustering (as opposed to being sampled 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 doesn’t change with time). Then the losses for each synthetic catalog from UCERF3 can be easily (and quickly) aggregated, and statistics can then be compiled over the different viable synthetic catalogs. This efficiency means that operational earthquake loss forecasting could be well within reach.

Our creation of synthetic catalogs is in contrast to 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 also have an alternative formulation for achieving this as described in the appendix under Task P5.

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 certainly be useful in terms of deciding what further complexities are warranted.

One question is whether application of statewide generic parameters, which were derived assuming a Gutenberg Richter distribution, will produce problems if applied in sub-regions where the magnitude-frequency distribution is clearly non Gutenberg Richter. For example, could our approach lead to runaway sequences in areas that exhibit a characteristic magnitude-frequency-distribution (because larger events will be sampled at a relatively higher rate, which will in turn spawn more aftershocks)? If so, we may need to solve for and apply spatially variable aftershock parameters.

Another detail glossed over above is the need to decluster the long-term Earthquake Rate Model before sampling aftershocks from it (to avoid double counting). However, since the fraction of main shocks versus aftershocks is assumed to be magnitude independent, this should simply involve finding what fractional reduction to the long-term model produces final catalogs that are consistent with the long-term rates.

Other issues that will need consideration include:

- Computation time.
- How far back in time do we need to go for collecting observed, viable main shocks?
- Is the distribution of nucleation points on a rupture surface really uniform?
- Is the fraction of main shocks versus aftershocks really magnitude independent?
• The spatial-temporal 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.

Tasks:

Tasks on the time-dependent models described above, as well as other issues related to the
development of UCERF3 Earthquake Probability Models, are described in Table 4.

**Table 4. Task for the development of Earthquake Probability Models.**

Please see the notes given for Table 1, as they apply here too.

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1) Address “Empirical” model</td>
<td>Examine robustness of apparent rate changes given reevaluation of historical catalog (task above) and for different time and space slices. <em>See appendix for further discussion of this task.</em></td>
<td>Felzer, Parsons</td>
</tr>
<tr>
<td>P2) ETAS explains Empirical Model?</td>
<td>Investigate whether ETAS is sufficient to explain the observed rate changes in the empirical model. <em>See appendix for further discussion of this task.</em></td>
<td>Felzer, Page, Michael?</td>
</tr>
<tr>
<td>P3) Coulomb Stress explains Empirical Model?</td>
<td>Investigate whether static coulomb stress changes can explain the observed rate changes. <em>See appendix for further discussion of this task.</em></td>
<td>Parsons, Powers?, Pollitz</td>
</tr>
<tr>
<td>P4) Develop self-consistent renewal models</td>
<td>Develop self-consistent, elastic-rebound-motivated renewal models, which are currently lacking for anything but strictly segmented models. <em>This issue is already described above in this report and also in Appendix N of the UCERF2 Report.</em></td>
<td>Field &amp; Page</td>
</tr>
<tr>
<td>P5) Implement ETAS for spatial-temporal clustering</td>
<td>This task is described above in a section dedicated to the topic above. <em>See appendix for more discussion.</em></td>
<td>Michael, Felzer, Page, Field, Powers</td>
</tr>
<tr>
<td>P6) Evaluate Agnew and Jones</td>
<td>Does the Agnew and Jones (1991) approach constitute a unique and implementable model? <em>See appendix for more.</em></td>
<td>Michael</td>
</tr>
<tr>
<td>P7) Evaluate Static Stress Change Models</td>
<td>Do static-stress change models constitute unique and implementable models (from an operational perspective)? <em>See discussion of Task P3 in appendix for further discussion.</em></td>
<td>Parsons, Powers</td>
</tr>
<tr>
<td>P8) Evaluate other time dependencies</td>
<td>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? <em>See appendix for further discussion of this task.</em></td>
<td>Hardebeck, Dolan?</td>
</tr>
<tr>
<td>P9) Evaluate Physics-based simulators (for probabilities)</td>
<td>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. <em>See appendix entry fro related task Task R13 for discussion.</em></td>
<td>Field, Michael, Tullis, Dieterich, Richards-Dinger, Ward, Rundle, Pollitz, Beeler</td>
</tr>
</tbody>
</table>
## Model Implementation

This section lists some of the model implementation issues.

### Tasks:

Table 5. Tasks on Potential Implementation Issues.

Please see the notes given for Table 1, as they apply here too.

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>I1)</td>
<td>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 need to start thinking about solutions now.</td>
<td>Field, Haller, Petersen, McCarthy, Wills, Dawson, Jordan, Milner, Husband</td>
</tr>
<tr>
<td>I2)</td>
<td>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., spatial and temporal clustering). Note that this activity is not part of the CEA-sponsored scope of work. Funding for this activity remains in question.</td>
<td>Porter, Field, Luco</td>
</tr>
<tr>
<td>I3)</td>
<td>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)</td>
<td>Field, Luco, Petersen, Porter, Campbell</td>
</tr>
<tr>
<td>I4)</td>
<td>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 stomach for long-term operations given this will require dedicated resources that don’t currently exist?</td>
<td>Powers, Field, Milner, Jordan, Gerstenberger, Jones, Earle, Petersen, Buland, Michael</td>
</tr>
<tr>
<td>I5)</td>
<td>Outline a clear strategy for testing both model predictions and embedded assumptions via coordination with CSEP. The first step will be to list all the assumptions that are likely to be made in UCERF3. We will then conduct a workshop (listed below) to discuss how each might be formally tested.</td>
<td>Schorlemmer, Jackson, Jordan, Field, Felzer, Page, Michael, Weldon</td>
</tr>
</tbody>
</table>
Other Possible Tasks:

- Conduct a more comprehensive and multidisciplinary evaluation of the distribution of aseismicity. This could include consideration of: a) microseismicity hypocenter distributions; b) geodetic modeling (e.g., relationship of locking depths to seismogenic depths); c) relationship of any spatial variations in b-values to the distribution of creep (is the former a proxy for the latter?); d) kinematic inversions; e) dynamic rupture modeling; f) physics-based earthquake simulators; and g) the relationship of proposed segment boundaries to creeping patches. Each of these is being pursued somewhat separately, so the question is whether a more combined effort would be useful.
- Formalize rules for data inclusion to avoid some double standards that existed in developing UCERF2.
- Contribute to the compilation of a precarious-rock-constraint database, which might inform our maximum-magnitude estimates, procedures for smoothing background seismicity, or double counting of aleatory variability of magnitude given area.
- Compile a list of key assumptions that are likely to be made in UCERF3 in order to facilitate discussions and for more formal testing procedures (e.g., via CSEP).
Planned Workshops & Review Meetings

The following two tables list currently anticipate SPR review meetings and workshops, respectively. Workshops, which by definition here include participants from the broader community, are aimed at addressing one or more topical issues. The review meetings, on the other hand, will involve formal evaluations by the SRP and possibly members of CEPEC, NEPEC, and the CEA science evaluation team. The topics and/or dates are subject to change as plans evolve, and it is possible that some of the review meetings and workshops will be coordinated 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 requirements (e.g., they are currently organizing a workshop on Cascadia that is not listed below).

### Planned SRP Review Meetings

<table>
<thead>
<tr>
<th>Review Meetings</th>
<th>Description</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Methodology Assessment</td>
<td>An overview of both Report #1 (Issues and Research Plan) and Report #2</td>
<td>November, 2010 (about a month before Report #2 is due)</td>
</tr>
<tr>
<td></td>
<td>(Proposed Solutions to Issues)</td>
<td></td>
</tr>
<tr>
<td>2) Proposed UCERF3 Plan</td>
<td>A comprehensive overview of the UCERF3 implementation plan (Report #3)</td>
<td>Mid June, 2011 (~2 weeks before SRP Report (#4) is due)</td>
</tr>
<tr>
<td>3) Preliminary UCERF3 Model</td>
<td>A comprehensive overview of the preliminary UCERF3 model (Report #6).</td>
<td>Mid April, 2012 (~2 weeks before SRP Report (#7) is due)</td>
</tr>
</tbody>
</table>
### Planned Workshops

<table>
<thead>
<tr>
<th>Title</th>
<th>Description</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Past</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCERF3 Planning Meeting</td>
<td>This workshop, which had broad community participation, was to discuss the goals and anticipated issues with building UCERF3.</td>
<td>Feb. 17-18, 2010</td>
</tr>
<tr>
<td>Incorporating Geodetic Data into UCERF3</td>
<td>This workshop began a comprehensive scientific discussion of how to incorporate GPS constraints on strain rates and fault slip rates into UCERF3.</td>
<td>April 1-2, 2010</td>
</tr>
<tr>
<td><strong>Future</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statewide Fault-Model &amp; Paleoseismic Data</td>
<td>This workshop will address 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 will also address paleoseismic trench data and its interpretation.</td>
<td>Oct., 2010</td>
</tr>
<tr>
<td>Distribution of Slip in Large Earthquakes</td>
<td>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.</td>
<td>Oct., 2010</td>
</tr>
<tr>
<td>Instrumental &amp; Historical Seismicity</td>
<td>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).</td>
<td>Nov., 2010</td>
</tr>
<tr>
<td>Assumptions &amp; Model Testing</td>
<td>This workshop will review likely UCERF3 assumptions and discuss how these might be formally tested.</td>
<td>Nov., 2010</td>
</tr>
<tr>
<td>Time-Dependent Models</td>
<td>This workshop will address what represents both “best-available science” and implementable models with respect to time-dependent probabilities in our UCERF3 operational forecast. Of particular emphasis here will be the interpretation of the empirical model, how to apply elastic rebound in unsegmented fault models, and how to represent spatial-temporal clustering.</td>
<td>Feb., 2011</td>
</tr>
<tr>
<td>Use of Physics-based Simulators</td>
<td>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 short-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.</td>
<td>Feb, 2011 (coord. w/ the above?)</td>
</tr>
<tr>
<td>Possible Implementation &amp; User-Community Issues</td>
<td>This workshop among key stakeholders and general users will address anticipated issues associated with using UCERF3. 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.</td>
<td>March, 2011</td>
</tr>
<tr>
<td>UCERF3 Deformation Models</td>
<td>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.</td>
<td>April, 2011</td>
</tr>
<tr>
<td>Overview of UCERF3 Plan and Preliminary Model</td>
<td>This workshop will constitute a complete overview of the anticipated UCERF3 model to the broader community. The timing of this workshop is to enable feedback to influence the final product.</td>
<td>Feb., 2012</td>
</tr>
<tr>
<td>Overview of UCERF3 for user communities</td>
<td>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.</td>
<td>Sept., 2012</td>
</tr>
</tbody>
</table>
References


Appendix – Detailed Task Descriptions

**Task F1 - Update Faults**
Task Leader: Tim Dawson (CGS)

A key component of the UCERF3 study is the statewide fault model 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 will be the primary objective of this Task, and lead to an improved representation of the known active faults included in the fault model.

The task will include:

- **Integration of new faults and revision of existing faults from recent studies.** Recent studies defining the location, geometry, and deformation rates of faults in California are available for inclusion in the UCERF3 fault model. For example, the new Fault Activity Map of California (Jennings and Bryant, 2010) represents the most up to date compilation of active fault traces in California and includes newly mapped faults not included in the UCERF2 fault model. Defining the three dimensional geometry of these faults will be a component of this task. Ongoing work by the USGS and PG&E, synthesizing geology, geodesy, and geophysical studies, is leading to an improved understanding of faulting and the tectonics of the Central Coast Ranges. This work is being used to develop a fault model for the central coast for use in geodetic block modeling and an updated PSHA for the Diablo Canyon Nuclear Power Plant. Coordination with this work and its integration into the UCERF3 fault model will lead to a better representation of active faults within the Central Coast Ranges. New mapping of other faults, such as the Maacama fault (Sowers and others 2009) and Bartlett Springs fault (Lienkaemper and Brown, 2009) is leading to a better understanding of these faults at the surface, and an evaluation of this mapping will provide valuable insight in how these faults will be defined in the model. The full scope of work is not limited to the aforementioned faults, and a search of the
available literature will be conducted for the entire State as part of this effort. As in UCERF2, alternative viable fault models, where appropriate, will be included in order to capture the epistemic uncertainties associated with the possible fault geometries.

- **Integration of the Statewide Community Fault Model into the UCERF3 fault model.** An ongoing SCEC-funded effort is being led by John Shaw and Andreas Plesch (Harvard) to develop a Statewide Community Fault Model for California (SCFM). SCFM will likely be a major improvement to the fault geometries used in UCERF2, particularly for northern California where many of the fault geometries are legacy geometries from previous versions of the National Seismic Hazard Maps and have not been reevaluated systematically. Another potential contribution of SCFM to UCERF3 is a revised seismogenic depth surface, based on the Waldhauser and Schaff (2008) relocated catalog. This depth surface is used to define the lower boundary of the fault surfaces in the model and may lead to better estimates of the base of seismicity for many of the faults in the model. The integration of the SCFM faults into the UCERF3 will lead to a more complete active fault catalog and better representation of fault geometries in the model. Part of this effort will include a comparison between the SCFM and UCERF2 models in order to identify areas with significant changes in the fault model representations. The documentation of these changes will be important if there are large changes in the earthquake rupture forecast in different regions of the State and these changes need to be explained. We expect that some UCERF3 efforts, such as an evaluation of fault endpoints and fault junctions, will benefit SCFM as well. Because UCERF3 will be reexamining fault endpoints for a number of faults in California, new information developed from this effort can be incorporated into SCFM to improve the Statewide model. A draft version of SCFM is scheduled for released in Fall 2010. Coordination the UCERF3 and SCFM efforts will be essential in order to ensure consistency between the models and this cooperation is already underway.

- **UCERF3 fault model as a foundation for GPS block models.** A significant amount of UCERF3 resources are being dedicated to the evaluation of geodesy-based deformation models. In the UCERF2 fault model, many active faults with poorly constrained or unknown slip rates were not included in the hazard calculation. A major question that UCERF3 will address is whether GPS-based deformation models can provide slip rates on faults where there are no geologic slip rate estimates. One approach is the use of block models in order to constrain fault slip rates using geodesy. A full catalog of active and potentially active faults will need to be developed in order to provide the GPS block modelers a consistent set of faults and fault geometries from which they can build their block models. We expect this interface between the geologically defined faults and the geodetic modelers will be ongoing throughout the course of this project. For example, strain rate maps being developed by the geodetic community may highlight areas where a closer examination of the geology is warranted to look for features indicative of active tectonics and assess the current state of geologic understanding of these areas. This may
result in the addition of faults to the model that are defined as active based on geodetic evidence.

*Deliverables:* An updated fault model that includes the most up to date fault locations and geometries for California. The fault model will include documentation describing significant changes from the UCERF2 model as well as the rationale for these changes. This fault model (Version 3.X) will provide the basis for the UCERF3 fault-based earthquake rupture forecast as well as provide the basis for the block models used in geodetic modeling.

**References**


Jennings, C.W., and Bryant, W.A., 2010, Fault activity map of California: California Geological Survey Geologic Data Map No. 6, map scale 1:750,000.


**Task F2 – Reevaluate Fault Endpoints**

Task Leader: Tim Dawson (CGS)

One of the issues UCERF3 will address is how to relax segmentation and accommodate fault to fault ruptures within the UCERF3 model framework. A detailed inventory of fault traces, as well as three-dimensional representations are available for faults within the UCERF study region through compilations such as the Fault Activity Map of California (Jennings and Bryant, 2010) and the Community Fault Model for Southern California (Plesch and others, 2007). However, not explicitly included in these compilations, is a characterization of fault endpoints. Do the faults simply end on a map, or are there associated structures, such as connecting faults or folding, that accommodate deformation between one active fault to the next? Does an active mapped fault trace end because that is where the fault physically terminates, or is there a lack of suitable geology that records evidence of active faulting further along strike? What are the
uncertainties associated with each endpoint? What can geology, geophysical data, and seismicity tell us about how faults end and accommodate deformation from one fault to another?

The goal of this task is a systematic reevaluation of fault endpoints for the faults in the UCERF3 fault model with a focus on looking at fault pairs that may participate in multi-fault ruptures. This characterization will synthesize available geologic, geophysical, and seismicity data in order to characterize fault endpoints and search for possible connecting structures between faults. Such structures may be expressed as surface faults or folds between major faults, boundaries defined by gravity or magnetic anomalies, or alignments of seismicity. Available geologic mapping will provide information regarding the surface expression of deformation between fault pairs. Additional data, such as subsurface information from geophysical profiles, and double-difference relocated seismicity can provide additional information regarding the three-dimensional geometry of fault pairs and how they connect at depth.

A separate, but related issue, are the uncertainties in the location of fault endpoints. It is important to keep in mind that much of the original fault mapping in these compilations was not done with the perspective of multi-fault ruptures. Also, the surface manifestation of structures between faults can be subdued, or not preserved in the geologic record because areas between faults are often either extensional or compressive regimes, where geologic evidence of faulting is obscured by burial or erosion. Is there sufficient existing geologic information to adequately characterize these areas, or are there substantial gaps in our understanding of these areas that we need to consider them as highly uncertain? Figure 1 illustrates an example of this issue where no active faults are currently mapped between the Mohawk Valley and Polaris fault systems in Northern California. However, as shown on the map, there are bedrock faults that could provide a direct connection between these two faults. What is unknown is whether or not these faults are truly “inactive”, or if evidence of recent activity is obscured by surface processes, such as recent glaciation. Another possibility is that the available geologic mapping in this area is not adequate to say one way or the other. Assigning uncertainties to the fault endpoints may have important implications on what fault pairs are allowed to rupture together. In the example above, if the mapping was determined to be unresolved between these two faults, then the uncertainties assigned to the terminations of the two faults would be rather large, perhaps enough to allow a direct connection between them. If, on the other hand, the geologic evidence was sufficient to say that the gap in active faulting is real, then the uncertainties associated with each fault endpoint would be small, and an earthquake spanning the ~10 km gap between these two faults would be considered highly unlikely.

Another example of a fault system that will need to be reevaluated in the context of multi-fault ruptures is the Great Valley System (Sections 7 – 14) that bounds the west side of the San Joaquin Valley. As currently represented in the UCERF fault model, this fault system is nearly continuous and a candidate for allowing fault to fault jumps. However, these representations are adopted from the 2002 National Seismic Hazard Map (NSHM), which did not consider multi-fault ruptures in the model. Simply adopting the NSHM model may not be an appropriate approach if these “legacy” faults are generalized seismic sources and the endpoints, as well as subsurface geometries, are not well characterized.

An additional element of this task is to revisit the “Connect more B-faults option” that UCERF2 included in an attempt to model larger faults from sets of shorter faults. The faults were connected because their orientation, proximity, structural style, and slip rate are similar enough
that they are believed capable of rupturing together. However, this process was largely *ad hoc* and the specific reasons why faults were connected are not well documented. This task will develop a database for each potentially connected fault pair, describing features such as proximity (both surface and downdip), faulting styles, kinematic compatibility, and slip rate, as well as describe the rationale for connecting faults. This database will be necessary if UCERF3 adopts a “rules based approach” for Task R6 (Fault to Fault Jumping Probabilities), where a set of predetermined rules, specifying to criteria necessary for faults to rupture together, will be applied to faults in the model. If fault to fault jumps are considered on a case by case basis, this database will still be useful in helping to identify fault pairs capable of rupturing together.

References

Jennings, C.W., and Bryant, W.A., 2010, Fault activity map of California: California Geological Survey Geologic Data Map No. 6, map scale 1:750,000.


Figure 1. Area between the active traces of the Mohawk Valley and Polaris fault systems showing bedrock faults as possible connection. (Jennings and others, 2010).

**Task F3: Database Issues**

Task Leader: Tim Dawson (CGS)
The California Reference Fault Parameter Database (CRFPD), developed for UCERF2, will be maintained and updated in order to serve the needs of the UCERF3 project. Currently, CRFPD consists of two components: The first component is the Fault Section database, which stores the spatial geometry of each fault section, including fault section name, fault trace, dip, rake, upper and lower depths of seismicity, and aseismicity factors. The second component is the Paleosites Database, intended to store event ages, slip-per-event data, and geologic slip rates developed at paleoseismic sites. The results of Task F1 (Update Faults), such as the addition of new faults and modification of existing fault geometries based on new data, will feed directly into the Fault Section Database. The Fault Section Database will also need to be modified to accommodate the results of Task F2 (Reevaluate Fault Endpoints), likely with the addition of database fields that record the uncertainties associated with the fault endpoints.

The Paleosites database will be maintained and updated, with an emphasis on the compilation of geologic slip rates. This effort is important for two reasons: First, a database of geologic slip rates, for each fault where available, is needed in order to compare geologically-derived slip rates to slip/strain rates generated by the deformation models being developed using geodetic data. The other use of the compilation of geologic slip rates is to examine potential biases in the slip rates used in the UCERF2 study. In UCERF2, the slip rates used in the model were generally consensus-generated slip rates, often adopted from previous Working Groups and versions of the National Seismic Hazard Map (NSHM). One issue is the uncertainties assigned to the fault section slip rates are somewhat arbitrary, typically \( 1/4 \) of the slip rate for rates on faults considered well-constrained, and \( 1/2 \) of the slip rate for poorly constrained slip rates, as described in Appendix A of the UCERF2 report. Another potential issue is the reported slip rate uncertainties are symmetrical around a preferred value. However, many of the reported geologic slip rates are actually minimum or maximum values, with uncertainties that are asymmetric around the preferred value. Finally, as currently represented, there is no distinction between the intervals of time the slip rates represent. A geologic slip rate derived from an offset during one earthquake and the time since that earthquake is probably not an appropriate geologic slip rate to use, yet some of the slip rates used in UCERF2 include such rates. Similarly, a long-term geologic slip rate averaged over the past 1 Ma may also not be representative of the current tectonic regime. All of the above issues may potentially contribute to a bias, potentially to the higher side, of slip rates, particularly for the low \( \leq 1\text{mm/yr} \) slip rate faults.

An effort is underway to review the literature for each fault in the UCERF model and document the original published geologic slip rates, with attention given to the above issues in order to develop a database of “pure” geologic slip rates. The Paleosites Database is the obvious location to document the efforts of this evaluation, and the database can be modified as necessary to capture aspects of interest, such as the uncertainties associated with the offset feature and dating, as well as a qualitative rating of the slip rate based on the offset feature, dating constraints, and assumptions that went into deriving the slip rate. The Paleosites Database is well-suited to storing this information, and will be necessary if we want to generate slip rates using a standardized approach, such as the one advocated by Zechar and Frankel (2010), or if we want to be able to generate asymmetric probability density functions for slip rates that represent a minimum or maximum rate, or are believed to be weighted towards one end of the reported range rather than the other.

References

**Task F4 - Update Faults**
Task Leader: Ray Weldon (UO)

*(some text taken from email for Arrowsmith)*

*Statement of the Problem:* The slip in the last event on a fault can be used to estimate the timing and/or size of future ground-rupturing earthquakes, in conjunction with the application of appropriate recurrence models (like a time predictable model to estimate the time to the next earthquake given the last event and a slip rate, or the size of future events assuming characteristic slip). This data is relatively easy to acquire because the most recent event is usually well preserved on the landscape for highly active faults.

*Proposed Solution:* We propose to collect available historic, paleoseismic and microgeomorphic data to define this parameter along as many faults as possible in the model. All faults that have ruptured historically, and many faults for which the approximate (C-14) timing of the most recent event is known from paleoseismic studies have estimates of the associated slip that can be derived from the literature or our paleoevents database.

In addition, high resolution topography from LIDAR data is very powerful and it covers much of the high slip rate parts of the fault system (eg San Andreas, San Jacinto, Elsinore, Garlock, Death Valley/Fish Lake Valley, Owens Valley, Hayward, Calaveras, as well as a Blackwater, Calico, Helendale, Lenwood, Tin Mnts, San Cayetano, Panamint, Mud Hills, Hunter Mountain, San Gregorio, Little Salmon, Paicines, Green Valley, Rogers Creek, Maacama; [http://www.opentopography.org](http://www.opentopography.org)). We propose to compile the recently published results; encourage and possibly help (fund) some of the more mature activities so their results can be published and included easily; and then encourage groups to target certain reaches of faults where there is likely to be a good signal that can be measured and where the slip per event is otherwise not well known.

Coupled with data from Task R12 – Distribution of Repeated Slip at a Site on a Fault – this data also may be used to estimate reasonable ranges of average slip for faults where repeated slip has not been documented and coupled with scaling relationships it could be used to help constrain Mmax both for faults and regions of faults that are not individually characterized.
**Task D1 – Evaluate New Deformation Models**

Task Leader: Wayne Thatcher (USGS)

*(This came in an email to Ned Field on Nov. 1st, 2010)*

SCEC Proposal Plans:

1. Kaj (with substantial participation from Yuehua) will work on further block modeling, with a new, refined test case for community (Bird, McCaffrey, Meade-Loveless, etc) and other calculations. BIGGEST PIECE

2. David will refine strain rate models, collate new responses from community, and make comparisons with geologic slip rates, seismicity.

3. Liz with work on postseismic transient deformation, collaborating with Fred Pollitz on standard VE relaxation modeling and her own FEM approaches to constrain 'ghost transient' influence on GPS field.

4. Wayne will draft proposal seeking SCEC funding for final wrap-up workshop in May 2010 in Pomona. Goal is GPS community endorsement of conveners' plan for UCERF3 'numbers'. Similar size and attendee list as April 2010 workshop. Proposal needs to be submitted by one of academic participants (Liz & I worked on last year's and it was submitted by Liz).

This would involve 4 separate proposals--we concluded it'd be too complex on this short timescale to try putting together a single coordinated proposal. But each will have a similar up front introduction/rationale that I've agreed to write.

Next 6 Months:

1. Lots of separate work by Liz, David, Kaj and Yuehua with frequent email/phone interactions and encouragement from Wayne

2. Conveners' meeting for progress reports and mid-course adjustments, notionally scheduled for early March in Golden. We need this face-to-face meeting to really see where we stand, what remaining tasks need doing, and get your input.

3. Announce final wrap-up workshop, coordinate logistics with Tran, plan the sessions, and line up key participants and speakers.

4. Workshop itself.

...Then prepare final 'numbers' for you and begin final reports for UCERF3 appendices.
**Task D2 – B-Fault Bias**

Task Leader: Ray Weldon (UO)

*Statement of the Problem:* There is some evidence that poorly studied, generally low slip rate faults have rates that are systematically overestimated. This is due to several factors including the fact that once one recognizes that a fault is active, its minimum slip rate is always greater than zero, so if its upper limit is poorly defined, its total slip rate range is centered at a rate greater than the actual rate. For example, if a fault with a slip rate of 0.1 mm/yr is recognized as active, but its upper rate can only be limited to less than 1 mm/yr, its range is reported as 0-1 mm/yr, which is then modeled as 0.5 +/-0.5 mm/yr. Other problems include the fact that one often can only determine rates of low slip rate faults at places where their rate is greater than average, and thus most easily studied biasing the rate if it is used as an average, and finally, when presented with equivocal data geologists appear to unconsciously prefer larger offsets and higher slip rates, simply because it makes their results more exciting and relevant.

*Proposed Solution:* 1) We propose to review the evidence that goes into determining the slip rate of all B faults, as part of our larger effort to review the slip rates of all of the faults in our model, with an eye to detect and eliminate the upper bias. 2) We intend to explore probability distribution functions to better capture the asymmetrical distribution of error. 3) We intend to look at the cumulative strain (using line integrals and strain tensor analysis) in regions with large numbers of B faults to see if we cap the total slip rate for groups of faults and thus better define the upper limits of individual faults that contribute to the total.

**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:


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.


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


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


Merritts, 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.


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


retreat events in the Columbia River littoral cell, USA, Marine Geology 273 (1-4), pp. 44-61.


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.


Task R1 - Evaluate Along-Strike Distribution of Average Slip

Task Leader: Glenn Biasi (UNR)

In UCERF-2 rupture displacement along strike for faults was modeled using a rupture profile developed by averaging over multiple events (Biasi and Weldon, 2006; Biasi and Weldon, 2008). To average over multiple ruptures, ruptures were first normalized by length and by average displacement. The resulting average shape compares extremely well with a functional form, \( \sin(x/2L)^{1/2} \). This shape fits well across a range of subsets also, including only short (<30 km) ruptures and only events >200 km in length. The average shape is something of a statistical construct, and not expected for any particular earthquake.

We plan to assess more systematically the assumptions and consequences of assumptions behind this simple, empirically observed shape. First, normalizing by rupture length assumes that the spatial variability of displacement of short ruptures is similar to that of long ones. Second, by normalizing by displacement, real differences among ruptures can be removed. Using this shape can, in effect, create ruptures reflecting a constant stress drop (Shaw, 2009). Mechanically this may make sense, but either way, the consequences of normalizing by average displacement should be clearly understood. Available data may be sufficient to evaluate both of these assumptions.

When multiple fault segments are involved, it is not known whether the sub-fault sections follow this general shape. Also, rupture step-overs are characterized by greater degrees of distributed displacement, rotation and extension at the ends, and distortion of displacement gradients. We plan geologic and geometric assessments of step-overs, which should give clues to the mechanical linkage of fault sections. We have identified 12 events in the data of Wesnousky (2008) likely to be of immediate use. Understanding the Landers earthquake will be important for UCERF-3 efforts to link B-faults because multiple faults were involved, and because the rupture did not follow the full lengths of the available B faults that it did use.

We also find that the average displacements of sub-fault ruptures of the Landers event are almost all too large to be predicted using standard scaling from their lengths. This suggests that the sub-faults were not individually loaded and accidentally ruptured together, but rather that they are resolving strain buildup of a larger, integrated system. We will investigate whether this a general feature when faults link. How do short faults “know” what displacement to have? If displacements on short faults reflect the final event magnitude, and not some more natural size scaled by their length, what is the mechanism? Can it be identified and used in a predictive understanding of fault linkages?

We will be using the data and descriptions of Wesnousky (2008) as a primary resource. To update this list we have identified the following earthquakes as either postdating his compilation, or that were not included for other reasons but that might be useful in the present work:
2010 Sierra El Mayor, Mexico  
2009 L’Alquila, Italy  
2008 Wenchuan, China  
2006 Machaze, Mozambique  
2005 Pakistan  
2004 Parkfield, California  
2004 Mid-Niigata, Japan  
1995 Kobe, Japan  
1976 Montagua, Guatemala  
1973 Luhuo, China  
1931 Funyun, China  
1905 Bulnay, Mongolia

As another source of displacement data, we plan to examine results from finite-source rupture modeling. Finite-source models provide a coarse and normally under-constrained estimate of the distribution of displacement on the fault surface. Because of the non-uniqueness of model results, an assessment is required of which features are robust and likely to be useful to understand general properties of earthquake rupture. A primary resource for this will be the compilation of 152 models for 80 earthquakes compiled by Martin Mai. Earthquakes from M4.1 to 8.9 are included. These data are presented in a standardized format that facilitates analysis of metrics such as maximum and average displacement, rupture displacement near complexities, and displacement gradients near the ends of ruptures. A comparison of slip metrics is planned between surface and subsurface slip for events where both are available. A recently developed surface slip gradient analysis method by Shaw (in press) will be applied, after extending it to the two-dimensional sub-surface case. We expect this comparison to constrain how slip, especially at long wavelengths, should be distributed in scenario ruptures.

References


Biasi, G. P. and R. J. Weldon (2008). Exploration of the impact of ground rupture shape on selection of rupture scenarios developed from paleoseismic data on the southern San Andreas Fault.”, Southern California Earthquake Center Annual Meeting, Southern California Earthquake Center, Palm Springs, California.

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

Task Leader: Bruce Shaw (LDEO)

Statement of the Problem

In UCERF2 Magnitude-Area scaling relations contributed one of the main sources of uncertainty in the final hazard estimates. In this task, we seek ways of finding additional constraints or alternative methods of using scaling which will reduce these uncertainties.

Proposed Solution

In UCERF2, Magnitude-Area scaling way used in two different ways. One way was to estimate the sizes of events that occurred on given sections of faults and the associated shaking through empirical attenuation relations. A second way was through deriving effective average slip in events to do moment or slip-rate balancing which then set overall rates. While this second use is a valid methodology, it introduces a few difficulties which are sources of significant uncertainty in the final hazard estimates. Prominently, differences in proposed empirical magnitude-area scaling laws imply substantial differences in hazard estimates, mainly due to the fact that the moment sum is dominated by the largest events, and differences in magnitude estimates of the largest events in the different empirical relationships end up affecting rates across the whole spectrum of sizes of events. An additional complication in this pathway is the recently raised question of whether or not a significant fraction of the moment may be coming from deeper slip below the seismogenic layer, thereby complicating ways of moment-balancing to ensure long term slip rates are matched.

In UCERF3, we propose to use an alternative complementary scaling law, slip versus length scaling, in combination with magnitude-area scaling. The slip-length scaling would be used in two different ways. First, and most significantly, we propose to use slip length scaling instead of magnitude-area scaling to do the slip-rate (or `moment') balancing on faults. This alternative pathway has a number of advantages. Importantly, because it is observable at the surface, it substitutes a directly observable relation for one that requires assumptions about how slip is distributed in depth. In this way, it reduces epistemic uncertainties. Moreover, since it feeds back on the rate of smaller events through the rate balancing, it contributes to reducing what is probably the largest contribution to uncertainty associated with the magnitude-area scaling relations.


The second proposed use of slip-length scaling observations is as an additional constraint on magnitude-area scaling in determining the sizes of events. While the down-dip distribution of slip matters in trying to use magnitude-area scaling to balance slip rates, it does not matter in trying to use them to determine the size of events and the associated shaking when using empirical attenuation relationships. In this latter case, there is a self-consistency in how seismogenic area is defined so that the resulting magnitudes and inferred shaking do not depend on uncertain depth dependencies of slip. At the same time, however, we can use the implied slip-length scaling inherent in definitions of area to see whether these implied scaling are consistent with the direct observations. In this second use, we project the magnitude-area scaling onto slip-length scaling, as was done in UCERF2 to do moment or slip rate balancing, but now use this to see whether the empirical scaling laws in that projection look consistent with the direct slip-length scaling observations. Figure 1 shows this approach. We learn a few things by doing this.

First, this transformation to slip-length viewed in linear-linear space, as opposed to the log-log space where the empirical fits were originally made, is a better reference frame to try to do slip-rate balancing from since the slips will be added linearly in the sum. Transforming to linear-linear space emphasizes trying to fit the largest events, whereas log-log space emphasizes trying to fit the widest range of events. To insure proper weighting of different information contained in the different events, and given the structure of the uncertainties in the data, it is still the case that it makes sense to do the fitting of the curves in log-log space; but it is important to make sure that the resulting fits do a good job in linear-linear space, since the largest events are most important to the sum, and where the most ambiguity lies.

Second, the scaling laws developed for magnitude-area seem fairly reasonable compared with the magnitude-area data rescaled to implied slip-length values (here assuming a constant seismogenic thickness of H=15km for all events, and constant down-dip average slip). However, when compared with a database of geologically observed slip-length data there are a number of significant issues that arise. One issue, seen in Figure 1b, is that the scaling laws appear to over predict the observations of geological surface slip. Part of the discrepancy here is due to the different events in the two databases; this is an issue we will return to. Understanding the sources of these differences is one aspect of this task that will be pursued. Another issue is we see one of the scaling laws used in UCERF2, the Hanks-Bakun scaling law [Hanks and Bakun, 2008], just does not appear viable even when rescaled in amplitude as all of the curves are done in Figure 1c.

One proposed solution is therefore to replace the Hanks-Bakun relation with other scaling laws which fit both the magnitude-area scaling data and slip-length data. A full review of all the candidate scaling laws is being carried out as part of this task. Further work on this is continuing. But at least one candidate scaling laws exists which matches, at least in functional form, both the magnitude-area and slip-length scaling data, that of Shaw [2009], which generalizes the Hanks-Bakun two-regime scaling to a three-regime scaling. In this way, as a second use of slip-length scaling, we propose to constrain magnitude-area scaling laws to also be consistent with slip-length scaling laws.

One final issue concerns the role of different focal mechanisms. This discussion has focused on strike-slip events since they are the predominant focal mechanism and the source of greatest
uncertainty. We also need to deal with thrust and normal faulting events. Figure 2 shows the slip-length scaling for different focal mechanisms adapted from the Wesnousky [2008] database. There are much less data for the normal and thrust events, and no very long aspect ratio events. From the figure, the thrust events appear to have somewhat higher slips than either the strike-slip or normal faulting events. If we fit a slope to the data, the thrust events have about a factor of two more slip for a given length. One potential solution is to use this sparse data to rescale the general functional form found for the strike-slip data. Since scale lengths in the problem are with respect to the down-dip seismogenic width, there would also be a factor of $H/cos(\theta)$ for seismogenic width $H$, where $\theta$ is the dip angle. For normal faults $\theta$ would have a default value of 60 degrees and for thrust faults 30 degrees. This would delay crossovers in scalings to larger lengths than they appear in the strike-slip case, further extending the increases we see at small lengths to larger lengths before any saturation effects started.

These proposed solutions raise a number of questions that need further discussion and study.

1. To help determine the impact of using slip-length scaling as opposed to magnitude-area scaling in the rate-balancing, it would be helpful to disaggregate the contributions to the uncertainty in hazard of the two different ways magnitude-area scaling was used in UCERF2. How much does the rate-balancing contribute? How much does the size-shaking contribute? This would help quantify the utility of different pathways and prioritize efforts at reducing uncertainty.

2. Some of the difference between the magnitude-area data rescaled onto implied slip-length data and the direct slip-length data is due to a difference in events considered in the two databases. In particular, there are a number of older Tibetan Plateau events that have implied high slip values. How do we want to deal with these older less constrained events? The globally determined moments are less certain from earlier events. And recent work with LIDAR on the use of channel offsets to infer penultimate slip has raised questions of the interplay between climatic and seismic processes (specifically the apparent overprediction of slip in the Carrizo Plain in the original paleoseismic estimates of the 1857 Ft Tejon event [Zielke, et al, 2010]). Given the significant role these outlying events are playing in the regressions and discrepancies, we suggest three courses of action. One, these older less constrained events be given larger relative error bars in regressions when they are included. Two, a separate regression using just better constrained recent events be made as an end member case. Three, a concerted effort be made to reexamine the older events in light of recent LIDAR results.

3. Implicit in the use of surface slip data is the question that it is representative of deeper seismogenic slip during the large events that dominate the net slip. There are a few pieces of evidence to support this, but it deserves scrutiny. One important point is that surface creep is rarely observed. It does occur on some faults, and in some cases, but it is rare, not common. Typically, then, most of the slip in the stable-sliding upper layer is occurring during large events. Since slip on average must keep up at all depths, this appears to be strong evidence that average surface slip is representative of average seismogenic slip. However, the possibility that unconsolidated materials at the surface may be masking offsets by accommodating some deformation and reducing observed surface slip values is one issue that deserves further scrutiny. Examining consistency of slip estimates from geodetic, seismological, and surface-slip
observations from better-constrained recent events is one way of examining this issue. Dynamic models have shown support for the argument that average slip behavior in the stable sliding surface is representative of average slip behavior down into the upper half of the seismogenic layer. While epistemic uncertainties in physical parameterizations and processes make these insights from modeling secondary rather than primary evidence, further study to examine questions of robustness seem worthwhile.

4. The use of slip-length scaling to do the slip-rate balancing and magnitude-area and attenuation relationships to do shaking estimates avoids the need to determine detailed depth dependence of slip behavior. Attempts to go beyond attenuation relationships to do more physics-based estimates of shaking, such as cybershake, however, will again make demands on knowledge of the depth-dependence of slip. This has shown up already in discrepancies between estimates of shaking from parameterized kinematic ruptures in cybershake when using unmodified UCERF2 magnitude-area scaling relations. The need for this depth-dependent slip behavior is therefore tied to the type of modeling used to estimate shaking. Developing dynamic modeling capabilities for use as guides in parameterizing ruptures thus remains a pathway of importance to further develop physics-based seismic hazard estimates. In terms of coordinating developments, is this a pathway being considered for UCERF3, or farther in the future? 5. Further work on focal mechanisms issues is called for, particularly given the sparsity of data for surface slip. Other measures of events relative to strike-slip events would help extrapolate the proper scaling to use.

References


Figure 1: Slip-length scaling. (a) Data (red circles) derived from [Hanks and Bakun, 2008] magnitude-area scaled by assumed seismogenic depth H=15km. Different color lines represent different magnitude-area scaling laws rescaled the same way the data has been; Ellsworth-B [WGCEP, 2003] (black), Wells-Coppersmith [1994] (yellow), Hanks-Bakun [2008] (green), Shaw [2009] (blue). (b) Data derived from Wesnousky [2008]. Colored lines same as in (a). (c) Same data as (b), but lines rescaled by factor of $.6$.

Figure 2: Slip-length scaling for different focal mechanisms. Strike-slip (red), normal (blue), thrust (magenta). Note that thrust mechanism data lies on average somewhat above the other two focal mechanisms. Data adapted from Wesnousky [2008].

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 palaeosite. 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](image.png)

**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 paleoseismic 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 $\mu \sim 210$ yr, and coefficient of variation $\alpha \sim 0.6$. 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

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 (e.g., 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 Large Inversion Problem

Task Leader: Morgan Page (USGS)

The purpose of this task is no less than to solve for the long-term rate of all possible (“on-fault”) ruptures. As described in the main report, we have an inversion methodology that solves for the rates of ruptures that are consistent with a) slip-rate constraints, b) paleoseismic event rates, c) a-priori rupture rate estimates, d) smoothness constraints, and e) constraints on the magnitude distribution (Field and Page, BSSA, in press). These constraints are linear and can be described via a matrix equation of the form \( Ax = d \). Our task is to set up this matrix equation and solve for \( x \), given \( A \) and \( d \).

We will solve the inverse problem via a simulated annealing algorithm. 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. It is designed to efficiently search a large parameter space without getting stuck in local minima. Next, quite importantly, 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.

Defining all possible ruptures
In order to set up the inversion, all ruptures must be defined \textit{a priori}. This task depends heavily on task R6 (fault-to-fault jumping probabilities) and requires generating simple rules to determine a given set of fault sections can rupture together in a single earthquake. We are pursuing two approaches in parallel to define the on-fault ruptures: 1) empirical rules for the probability of fault-to-fault jumps and bends based on an analysis of past surface-rupturing earthquakes, 2) a Coulomb model that will compute the kinematic consistency of possible rupture scenarios. Within the inversion methodology it is possible to weight individual ruptures by their likelihood based on fault geometry; alternatively, a binary approach could be implemented that would allow all fault ruptures within certain viability criteria to be included in the inversion without penalty. Reasonable rules, such as those described in the main body of the report, lead to on the order of 100,000 possible on-fault ruptures for all of California.

\textit{Memory constraints}

In the absence of regularization, the size of the A matrix in the inverse problem scales as the number of subsections multiplied by the number of ruptures. Given approximately 100,000 possible ruptures, the A matrix stored at double-precision (16-bit) elements will be at least several hundred megabytes in size. This is a feasible matrix size for modern computers that typically have 4-8 GB of memory. In addition, the A matrix is sparse, so if memory constraints become a problem, we plan to use sparse matrix techniques.

\textit{Computational time}

As the problem size increases, each iteration of the simulated annealing algorithm takes longer to run; this increase scales with the time required to do the forward problem (matrix multiplication). In addition the number of simulated annealing iterations required to reach a given level of misfit increases. Based on initial tests, it appears that the computational time to reach a given level of misfit scales as $N^{1.5}$, where $N$ is the number of matrix elements. Extrapolating this relationship from smaller problems, we estimate that the large inverse problem could be solved via a simulated annealing algorithm in approximately 16 hours. Improved agreement with data could be achieved with longer annealing times.

The simulated annealing algorithm is easily parallelizable to multiple processors. One such algorithm that could be used in a parallel algorithm of the simulated annealing component is parallel tempering, which searches multiple portions of the parameter space at the same time. Alternatively, multiple processors could be used to run the inversion on different iterations of data; this would allow the data resolution to be thoroughly explored.

\textit{An alternative earthquake simulator}

The inversion method is quite computationally intensive. For this reason we are also exploring an alternative method being developed by Tom Parsons. This method is a spatial seismic gap model that Monte Carlo samples earthquakes from a Gutenberg-Richter distribution (Parsons and Geist, 2009). Hypocenters are placed randomly at locations on the fault (with more preference given to locations that have a higher slip rate). Ruptures then grow onto neighboring patches, or onto other faults within a specified distance, in the directions that have the least accumulated slip. Earthquakes continue to be placed until a predefined slip budget is satisfied; at the end of each simulation a set of earthquakes has been defined that satisfies the \textit{a priori} slip rates and
follows a Gutenberg-Richter magnitude distribution. One current problem with this methodology is that best-fit solutions have a maximum magnitude of 7.7, which is in disagreement with historical earthquakes in California. Possible solutions to this problem, such as incorporating more sophisticated fault-jumping rules, are being explored.

**References**


**Task R6 - Fault-to-Fault Jumping Probabilities**

Task Leader: Ruth Harris (USGS)

One of the primary goals for UCERF3 is to include fault-to-fault earthquake ruptures, which will require having some kind of estimate of the likelihood of such events. Ideally we would have some model giving the probability of fault jumping given some information:

\[
Prob(\text{jump} | \text{information})
\]

where that information might be one or more of the following: distance between faults, relative orientation or change in strike, style of faulting, hypocenter, overall size of the event, slip rate, etc. Unfortunately no such model exists, and it’s not clear exactly how to develop one.

Another and-member approach would be to rely on expert opinion on a case-by-case basis (evaluating each possible fault combination separately). At the very least we would want this to provide a Boolean answer (yes vs. no) to the question of whether or not a given rupture jump is possible (or better yet, a probability). This is less desirable than using an objective formula due to reproducibility and testability issues.

A third option would be to compute a metric based on coulomb stress change calculations [Reasenberg and Simpson, Science, 1992; Stein et al., Science, 1992; Harris and Simpson, Nature, 1992; Parsons et al., JGR, 1999], that would be a static or quasi-static Coulomb failure stress-change simplification of the comprehensive dynamic rupture calculations presented, in the work by Harris et al. [GRL, 1991], and Harris and Day [JGR, 1993], among others. An example of this is a study by Harris et al. [BSSA, 2002] where the likelihood of fault-fault jumping by the 1999 M7.4 Izmit, Turkey earthquake was estimated using a simple static stress change calculation near the eastern end of the ruptured faults. Although the static stress change calculation couldn’t allow for the dynamic two-way interaction with the fault bend at the eastern end, the simpler calculation provided some guidance as to why rupture was possible to a seemingly unlikely fault-orientation. Using a static or quasi-static strategy, faults that are calculated to have positive Coulomb stress changes over their entire surface would presumably imply a greater mechanical compatibility (and therefore likelihood) than those that are a mix of
stress increases and stress decreases. One question is how to normalize these calculations to give
the relative likelihood for all the various rupture possibilities.

We plan to pursue all of these approaches in developing UCERF3.

Specifically, we will convene a meeting of experts on this topic to go through the actual fault-
jumping candidate pairs in California on a case-by-case basis, with the goal of defining an expert
opinion probability or Boolean for each pair (we may need to prioritize according to important
faults if time is an issue). By taking notes on each scenario addressed, we would hopefully build
a body of reasoning that could then be used for establishing more generic rules, or even a
formula as articulated above. One question is how to handle uncertainties in the faults
themselves; perhaps we assume perfect accuracy for this exercise?

We may need to build on Wesnousky’s fault rupture database to add new events and or new
parameters that might be used for predictive purposes (fault orientation etc.). We also need to
consider the fact that Wesnousky’s database represents measurements taken after large events
rather than from the more limited information available before an event (the latter is what we
will be dealing with).

Participants of this workshop could include: Ruth Harris, Steve Wesnousky, David Oglesby,
David Jackson, Ray Weldon, Tom Parsons, Peter Powers, Morgan Page, Bruce Shaw, Tim
Dawson, John Shaw (the latter two in order to inform the needs for future fault models), and
anyone else that’s interested.

The hope would be for this workshop to lead to a report defining a proposed usable model, a
review of the literature, and a path forward in terms of future research.

A broader community workshop could then be conducted to review the proposed solution.

Finally, it should be noted that we are concurrently proposing to address this research topic by
using a SCEC-community-tested computer-code to simulate 3D dynamic (spontaneous) rupture
simulations of earthquake ruptures encountering fault stepovers. These simulations include the
known physics of stress-waves and fault-friction, and intertia, and allow for rigorous treatment of
off-fault behavior [e.g., Duan and Day, JGR, 2008; Ma and Andrews, JGR, 2010; Harris et al,
manuscript in preparation]. In the future, multi-earthquake-cycle simulations may also be
brought to bear on the problem. Simulations of multiple events with simple non-linear fault
geometry have already been completed by Duan and Oglesby [JGR, 2005; 2006], in addition to
work by B. Shaw and others. An upcoming challenge will be to include the entire known fault
geometry of California [CGS, John Shaw, Waldhauser, etc.] along with realistic friction [e.g.,
Lapusta and Liu, JGR, 2009; Kaneko et al., Nature Geoscience, 2010] and appropriate
formulations for the short-term and long-term responses of the surrounding rocks that will allow
for viable suites of dynamic and quasi-static earthquake simulations to be run.

**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. Current practice uses contours of likely intensity centers to identify approximate allowable earthquake centroids. These are then assigned to likely faults [e.g., Bakun and Wentworth, 1997]. 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 proposal is to modify the Bakun 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


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.

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.

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 historic catalogs. These changes would affect all of the historical earthquake catalogs.

B Different historic 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.

F) Finally, 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.

References:


**Task R9: Smoothed Seismicity Model**
Task Leader: Karen Felzer (USGS)

We have programmed an implementation of Helmstetter et al. (2007) (and the follow up paper, Werner et al. (2010)) to produce a smoothed seismicity model. We are using this model because it has been most successful in the CSEP forecasting test to date. Helmstetter et al. (2007) gives an equation for the gain of a forecast, which is a metric that compares the log likelihood scores of a forecast to the log-likelihood score of a uniform probability map. We use this metric below to evaluate different ways of implementing the Helmstetter method. We also plan to use the gain to compare the optimal forecast using the Helmstetter method with the smoothed seismicity map used for UCERF2.

Our current implementation of the Helmstetter algorithm uses a power law kernel around each earthquake, which is slightly preferred over the Gaussian by Helmstetter et al. (2007). The equation for the power law kernel is given by,

\[ K_d(r) = \frac{C(d)}{(r^2 + d^2)^{1.5}} \]

where \(C(d)\) normalizes the kernel, \(r\) is distance from the earthquake being smoothed, and \(d\) is the smoothing constant. A key innovation of Helmstetter (2007) is that \(d\) is not a set distance but rather 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 is clearly visible even though the kernel around each individual earthquake is circular. In the past kernels around earthquakes along major faults often needed to be elongated by hand.

Helmstetter et al. (2007) uses the catalog down to M 2. Since the catalog is not uniformly complete to this magnitude, this requires complex calculations and corrections. Werner et al. (2010) showed that little gain was lost when the minimum magnitude was moved up to M 4, which is the minimum magnitude traditionally used for the National Hazard Maps. In our implementation we have used a minimum magnitude of M 4 and forgo the completeness magnitude corrections.

Helmstetter et al. (2007) used the earthquake catalog back to 1981, when the modern and much more complete catalog begins. We find that using data all the way back to 1850 actually slightly improves gains despite the fact that the older part of the catalog is poorly located. Presumably this is because despite the poor locations the old earthquakes help to
highlight longer term active areas of the state. We also note that because the older part of the catalog is incomplete it contains on the order of 10 to 100 times fewer earthquakes/year then the modern catalog, so even when data back to 1850 is included earthquakes occurring after the initiation of the instrumental catalog in 1932 still dominate the results.

Using data from 1850 and M≥4 we ran three five year trial forecasts (over 1993-1997, 1997-2001, and 2002-2006, respectively). For the smoothing parameter we find an optimal value of n=3, which compares to the n=3.75 found by Werner et al. (2010). We presume that the different result is because Werner et al. (2010) tested over a wider range of minimum magnitudes. As the minimum magnitude cutoff drops earthquakes become closer together and hence we might expect that the optimal n may increase.

Helmstetter et al. (2007) used a catalog that was declustered using the method of Reasenberg (1985). We ran trials with a catalog that has been declustered with the Gardner and Knopoff (1974) method, which removes slightly more earthquakes than Reasenberg (1985) and has traditionally been used by the National Hazard Maps. We also ran trials using the complete catalog and catalogs in which earthquakes were weighted with an inverse power law from their time of occurrence, which models aftershock decay.

In several additional short trials set right after major earthquakes, weighting earthquakes with an inverse power law performed much better, as would be expected since this method focuses attention on the most recent aftershock zone. For the five year forecasts, however, using the entire catalog, without any weighting or declustering, actually performed best, although using the declustered and the weighted catalog gave nearly the same results. This is in contrast to Werner et al (2010), who stated that declustering improves gain. The difference may result from the different magnitude cutoffs and catalog duration used. The difference in declustering method may also be important. It is important to note, however, that a probability map based on a thoroughly declustered catalog is really not desired, as large aftershock zones remain sources of active seismicity for years. Indeed, Wang et al. (2010) demonstrates that the Reasenberg (1985) method does not remove all aftershocks, and this is borne out by the highlighting of the Landers and Hector Mine aftershock zones in the Werner et al. (2010) smoothed map.

Teasing out the declustering issue remains an issue for further work. We also need to evaluate whether smoothed maps made with the Helmstetter method are more consistent with precarious rock positions than the smoothed seismicity map produced for UCERF 2.

References


**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 normalization 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 1. 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


**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 historic 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? 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?)
- Spatial-temporal 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).
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 updip 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 2011. The CSZ issues noted above will be discussed at this workshop, so this workshop will also be important for UCERF 3.
Task P1 – Address Empirical Model
Task Leader: Karen Felzer (USGS)

Status of Model Components

The empirical model is the application of a reduction of the short term expected number of earthquakes based on observations that the recent, instrumental (post-1932) California catalog contains a lower seismicity rate than the 1850-2007 average. In UCERF2 the empirical correction was applied in region-specific amounts to all regions in the state for which sufficient data was available. Regions were drawn to encompass areas of similar levels of catalog completeness, where the latter was determined from the locations of population centers and newspapers and later seismic instruments, based on the methodology of Schorlemmer and Woessner (2008). The state as a whole is complete to only ~M 7.5 from 1850 (UCERF2, Appendix I), thus the use of different completeness regions is essential to calculating catalog-based seismicity rates.

Over the whole state the seismicity rate in the modern instrumental era is about 75% of the 1850-2006 average. Given difficulties in estimating the magnitudes and locations of historic earthquakes, there is significant question regarding the robustness of the rate decrease. Recent seismicity rates are also low in comparison to GPS measurements, however. Ward (1998) estimated that the 1850-1996 rate of seismic moment release in California was only 75% - 86% of the long term geodetic-based rate. The average annual seismic moment release from 1932-2006 is about 60% of the long term geodetic rate estimated by Ward (1998). Thus re-evaluation of the catalog may allow for more precise determination of the rate difference, but there are multiple lines of evidence that the current seismicity rate is lower than the long term average. In fact, it can be shown with simulations that the clustering of most earthquakes in aftershock sequences means that the majority of the time the seismicity rate is expected to be below average. Important questions remain, however, about the cause of the rate reduction (see task 26 below). If it cannot be shown definitively that the cause is anything other than aftershock clustering then something like the ETAS model rather than a straight linear projection should be used to estimate how this rate reduction might play out in the future.

Research Plan

Re-calculate the rate change between the instrumental and full catalog after reassessment of the historic catalog (Task 17) using the methods, regions, and completeness thresholds given in UCERF2, Appendix I.

Error bars on the rate change for most regions were very high in UCERF2, and given the limited data are likely to remain high. We will investigate the implications of these large errors on the best course of action.

References


**Task P2 – ETAS explains Empirical Model?**

Task Leader: Karen Felzer (USGS)

**Current Status**

UCERF2 demonstrated that a decrease in rate between the instrumental and historic catalog exists in all regions of California with the possible exception of the Mojave, which may have experienced a rate increase. The change appears strongest in the San Francisco Bay Area, where the rate decrease is on the order of 50% for 1906-2006 vs. 1850-1906. An iconic figure by Ross Stein (Figure 1, below) shows 14 M≥6 earthquakes in the 75 years preceding 1906 and 1 in the 75 years following 1906, an apparently dramatic shift in the seismicity rate. A similar plot, known colloquially as the “tombstone plot”, shows 33 M≥5.5 earthquakes for 1850-1906 and 10 M≥5.5 earthquakes for 1906-2002.

Variations in seismicity rate are normally expected as a consequence of 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 modeled with the stochastic ETAS model (*Ogata*, 1988), although the model is limited because aftershocks of earthquakes too old, distant, or small to be in the earthquake catalog are not included, and because the magnitude of the smallest earthquake that can produce aftershocks is not known. As a result the model uses a steady background rate that is higher than the true rate, and outputs a lower limit on aftershock-related variability. Nonetheless, preliminary trials show that ETAS simulations can randomly produce M≥6 rate changes similar to that in the Stein figure, over similar times and areas (Figure 2). Whether ETAS can produce a coordinated seismicity rate decrease over a large part of the state, as found in UCERF2, still needs to be investigated.

An additional complication is that on closer inspection much of the statewide rate decrease is concentrated along the San Andreas system. Many more earthquakes occurred on or near the length of the SAF from 1855-1927 than from 1927-2000 (*Felzer and Brodsky*, 2005). This sharp localization of the rate decrease may not be reproducible by current ETAS modeling without the introduction of variation in local background rates. We note that such variability may not be real but may be needed because of the limitations of the ETAS model noted above.

**Research Plan**

We plan to run the ETAS model, in the form described by *Hardebeck et al.*, (2008), to test how well it reproduces the seismicity rate changes statewide and in the UCERF2 defined regions.
After the regional testing we plan to look at the San Andreas system and any other locations found to experience severe rate changes. If these changes cannot be reproduced with normal ETAS we plan to experiment with adding in localized background rate changes to see if we can better reproduce the catalog.

References


Figure 1. Iconic San Francisco Bay Area Shadow figure by Ross Stein
Figure 2: ETAS simulation of 150 year catalog. Note strong temporal rate changes in the San Francisco Bay Area and Mojave Desert.

Task P3 - Coulomb Stress Explains Empirical Model?
(also includes description of Task P7)

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 mainshock. 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.

Tasks P5 & P6 - Implement ETAS for spatial-temporal clustering, and Evaluate the Agnew & Jones method
Task Leader: Andy Michael (USGS)

Earthquake clustering, in both space and time, provides a rare opportunity to forecast earthquake behavior over short time intervals with high probability gains compared to the probabilities of earthquakes occurring as independent events. Due to the high probability gains, these forecasts can be useful to society [Michael et al., 1996]. Estimating earthquake probabilities within clusters has been addressed by two substantially different approaches. Reasenberg and Jones [1989] combined two fundamental laws of statistical seismology, the Gutenberg-Richter distribution of earthquakes with respect to magnitude [Gutenberg and Richter, 1944] and the Modified-Omori law which describes the temporal behavior of aftershock sequences [Utsu, 1961] in order to estimate the probability of different magnitude earthquakes occurring during time-windows of an aftershock sequence. By extending their relationship above the magnitude of the initial event they also provided a model of foreshock behavior. The calculations in Reasenberg and Jones [1989] were corrected in Reasenberg and Jones [1994] and hereafter we refer to both of these papers as RJ89. The ETAS approach [Ogata, 1988] captures the same two fundamental laws and can be used in a similar way and has been applied to California by Felzer et al. [2003] among others. And the STEP model added spatial information to the RJ89 approach while also extending the results from probabilities to hazard [Gerstenberger et al., 2005] Agnew and Jones [1991] (hereafter AJ91) produced a model only of foreshock behavior by artificially separating earthquakes into four classes: aftershocks, mainshocks, foreshocks, and background events. The AJ91 method does not provide probabilities of aftershocks but provides a way to combine long-term estimates of earthquake probabilities into short-term estimates based on clustering.
As described in the main body of the report, our primary effort will be to develop an ETAS-based alternative to STEP that properly incorporates the earthquake rates from the UCERF models. In addition to the simulation approach discussed in the main body, Morgan Page has also developed an analytic approach (available upon request) to solve for the average total triggering in a sequence based on the direct Omori law and including subsequent generations of aftershocks. These analytic solutions include both a specific solution for a Gutenberg-Richter magnitude frequency relationship with a strict upper cutoff magnitude and a general solution for an arbitrary magnitude-frequency relationship. For some magnitude-frequency relationships and direct Omori parameters the rate of earthquakes could diverge in time and so the analytic solution could be used to check that the average total rate decreases with time as observed in real sequences.

The AJ91 model does consider characteristic earthquakes and can include other magnitude-frequency distributions for the main shocks. However, it is limited because it does not include aftershocks. In the case of simple magnitude-frequency distributions that cover the full range of earthquakes it has been shown that the AJ91 model reduces to the RJ89 model [Michael, 2010]. Thus, the AJ91 model does not fulfill all of UCERF3’s needs and may be redundant with the ETAS model under development. The AJ91 model does include a triggering potential term that depends on the long-term seismicity rate in an area and we are currently investigating if that term is an accurate depiction of real sequences.

Depending on the results from these initial efforts we may need to combine the ETAS, STEP, and AJ91 models using a logic-tree. Given that STEP is based on the RJ89 model we do not expect to include the RJ89 model as a separate entity.


**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 spatial-temporal 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 spatial-temporal 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 spatial-temporal 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 spatial-temporal 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 spatial-temporal 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:


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


