

## **Report and recommendations of the 1 November 2006 Magnitude-Area relation summit convened by the Working Group for California Earthquake Probabilities**

**Participants.** Present at the meeting were Bill Ellsworth, Jessica Murray, Donald Wells, Tom Hanks, Bill Bakun, Paul Somerville, Ned Field, Ken Hudnut, Colin Williams, Egill Haukkson and Ross Stein. In addition, Paul Segall, Roland Bürgmann, and Jim Savage contributed references or written comments. Paul Somerville provided a written review of magnitude-area relations for the Working Group, and Tom Hanks wrote a comment on the Somerville report, which is appended to the end of this one.

**Presentations.** The agenda and the PowerPoint presentations given at the meeting by Ellsworth, Field, Somerville, Hanks, Murray, and Williams are available from:

<http://gravity.usc.edu/WGCEP/actionItems/meetings/110106/index.html>

### **Needs of WGCEP**

The Working Group for California Earthquake Probabilities (WGCEP) must be able to assign earthquake magnitudes from inferred fault geometry. The length of past fault ruptures and the length of continuous fault traces are the best observed parameters. The down-dip fault dimension  $W$  is least well resolved, resulting in considerable uncertainty in fault area,  $A$ . Empirical relations between fault area and moment-magnitude suggest that the static earthquake shear stress drop is roughly constant over a large range of magnitudes; this may transition to fault-length scaling for continental strike-slip earthquakes over  $M_w=7$ . The frequency-magnitude distribution for California is partly dependent on the magnitude-area relation, and so these two efforts must be self-consistent in method and in the definition in fault area.

### **Estimation of downdip fault dimension, $W$**

Limited geodetic sampling and large ambiguities in the inferred maximum depth of coseismic slip and the interseismic locking depth, render geodetic estimates of the down-dip fault dimension,  $W$ , inadequate for the purposes of WGCEP. There is wide consensus on

this conclusion by the geodesists, including Hudnut, Murray, Savage, Bürgmann, and Segall.

Heat flow observations are generally consistent with the maximum depth of seismicity along California faults, with high heat flow corresponding to a shallow depth of seismicity. This means that the lower depth of seismicity may trace the brittle-ductile transition on most faults. But heat flow coverage is spatially limited and the observations are variable, making thermal measurements and models more suitable as a test of other approaches, rather than as one that can furnish a lower depth of faulting for all California faults.

The WGCEP will therefore use the method of Nazareth and Hauksson (2004) to estimate the lower fault depth from background seismicity. The seismicity coverage is much more complete and uniform than is the geodetic coverage. Nazareth and Hauksson demonstrate, with an albeit limited southern California dataset, that the depth above which the 99.9% of the moment release of background seismicity occurs reasonably estimates the maximum depth of rupture in moderate to large earthquakes. This method suffers where background seismicity is sparse, such as on parts of the San Andreas. There is also the possibility that the lower depth extent of  $M \geq 7.4$  shocks will prove to exceed the lower depth of background seismicity. We nevertheless believe this to be the most sound and consistent method to estimate the lower depth available. The upper depth of faulting is left unresolved by this method, but we will arbitrarily set this to be 0 km except in special cases. This might overestimate the area, but few alternative assumptions exist.

### **Estimation of aseismic slip parameter, $R$**

The aseismic slip parameter,  $R$ , is better estimated from interseismic geodetic and creepmeter data than is the down-dip slip dimension,  $W$ . There are nevertheless few California faults known to undergo significant interseismic creep. So, except for creeping or partially creeping faults,  $R$  will not be used to adjust the modeled slip rate. All but the known creeping fault will be assumed to be fully locked between earthquakes.

### **Magnitude-Area relations used to infer earthquake size**

Equations relating  $M_w$  to rupture area,  $A$ , are derived from empirical earthquake datasets, most notably Wells & Coppersmith (1994). Of greatest importance to WGCEP are large strike-slip earthquakes, for

which the sample is small. Although essential, these datasets suffer from uncertainty in the down-dip dimension,  $W$ , and for many historical earthquakes there are also large uncertainties in  $M_w$  and in some cases, even the length of the rupture,  $L$ . Published equations by Wells & Coppersmith (1994), Hanks & Bakun (2002), Ellsworth (Working Group, 2002), as well as Paul Somerville's unpublished 2006 report, present alternatives.

Although Somerville uses a more uniform, modern dataset, it is as yet unclear that his method to infer  $W$  is compatible with how  $W$  is treated elsewhere in the WGCEP fault section database and deformation model. Although Hanks & Bakun (2002) provide a physical basis for the bilinear scaling in their equation (fault area stress drop for  $M_w < 7$  and scaling by fault length for  $M_w > 7$ ), there are too few observations at  $M_w \geq 7.5$  to be confident in the departure from a single slope. Thus for the immediate needs of the WGCEP, we recommend giving equal weight to the Working Group (2002) and Hanks & Bakun (2002) equations, cognizant that the former is also consistent with Wells & Coppersmith (1994). For the longerterm needs of the WGCEP, we would like these databases enhanced to include recent  $M_w > 7$  strike-slip events.

### **Impact on the magnitude-frequency distribution for California**

Although many parameter selections are needed to produce an all-California magnitude-frequency distribution, the WGCEP has found that most realizations of this distribution suffer from too high a rate of  $M \sim 6.75$  ( $6.25 < M < 6.75$ ) earthquakes in comparison to the inferred historical rate, a cause of considerable concern. Although the Somerville (2006) relation exacerbates this 'mid-magnitude bulge' in earthquake frequency, none of the magnitude-area relations removes it. This indicates that the mag-area relation is not the largest contributing factor in the disagreement between model and data.

### **Requests for additional analysis**

*For Donald Wells and Kevin Coppersmith:*

We would like you to include the 12 years of earthquakes that have struck since his study was published. In addition, there are new studies of many of the pre-1994 earthquakes that should cause these values to be reassessed. Please review, for example, the 1973  $M_w = 7.5$  Luhuo earthquake on the Xian Shiehe fault, for which the fault length

may need modification [Zhou, Allen, and Kanamori, BSSA, 73, 1585-1597, 1983].

*For Paul Somerville:*

An example of how you calculate the untrimmed area and trimming process would be very helpful, since this is so central to how you define area,  $A$ . How are the often large areas of low slip or no slip treated, for example, in calculating fault area? In contrast to your assertions, the shear modulus layering effects shown in Fig. 1 are subtle. Depth is unfortunately unlabeled in the figure. It looks to me that  $W$  for the layered models changes very little. The larger change appears to be in an increase in slip for the multiple-layered models.

*For Tom Hanks and Bill Bakun:*

Please add the large strike-slip Eqs that postdate your publication. These, at the very least, include 1997 Mw=7.5 Manyi (Tibet), 1999 Mw=7.6 Izmit & M=7.1 Düzce, 2001 Mw=7.9 Kokoxili (Kunlunshan), 2002 Mw=7.9 Denali. Another possible entry is the 27 May 1995 Ms=7.6 Northern Sakhalin earthquake, which was strike-slip [BSSA, 94, 117-130; doi: 10.1785/0120020175]. These additional or corrected events will roughly double the  $M \geq 7.5$  dataset on which you based the high-magnitude part of the relation. There will be large  $W$  uncertainties for some of these events, which we would like you to represent by error bounds on  $A$ , but your 1857, 1905, 1920, 1939 and 1957 shocks already have large uncertainty in  $W$ , and I doubt these will be worse.

### **Concluding recommendations**

For the immediate needs to the WGCEP, which is to deliver a Poisson probability model to the National Seismic Hazard Mapping Project, we will use the Nazareth & Hauksson method to estimate  $W$ . We will give equal weight to the Ellsworth B (Working Group 02) and Hanks & Bakun magnitude-area relations. We will limit the use of aseismic slip factor  $R$  to faults with indisputable and substantial creep. For the 2007 products of WGCEP, including a time-dependent model and revised Poisson models, we would like to revisit the magnitude-area relations after the requested updates and revisions have been made.

*Ross S. Stein*

Member, Executive Committee

Working Group for California Earthquake Probabilities

- Hanks, T. C., and W. H. Bakun (2002), A bilinear source-scaling model for M-log A Observations of continental earthquakes, *Bull. Seismol. Soc. Amer.*, 92, 1841-1846.
- Nazareth, J. J., and E. Hauksson (2004), The seismogenic thickness of the southern California crust, *Bull. Seismol. Soc. Amer.*, 94, 940-960.
- Wells, D. L., and K. J. Coppersmith (1994), New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seismol. Soc. Amer.*, 84, 974-1002.

## Appendix: 8 November 2006 letter of Tom Hanks

Ross,

While I found the  $M - \log A$  meeting of November 1 interesting and occasionally captivating, after a long time away from this business, it covered too much ground too quickly, and only at later times did I appreciate several matters described below. For the benefit of those who are new to this line of work, I'll start with a little background from Working Group (2003).

WG03 dealt not with one model, as implied in Paul's June 20 manuscript, but with three sets of them: (1) the Wells and Coppersmith (1994) relation,  $M = \log A + 4.0$  (to two significant figures); three relations developed by Bill Ellsworth,  $M = \log A + 4.1$ , 4.2, and 4.3; and (3) two bilinear relations developed by myself and Bill Bakun. All of these relations are presented in Chapter 4, pages 4-6, of *Earthquake Probabilities in the San Francisco Bay Region: 2002-2031*, U.S. Geological Survey Open-File Report 03-214, 2003, and they are plotted together with several data sets, including WC94, in Figure 4.2. The Ellsworth relations are based on the analysis of Appendix D of the same OFR. The basis for the bilinear relations is presented in HB02 in the June 2002 issue of the BSSA.

The essential problem for WG03 was that WC94 significantly underestimated  $M$  for the 1906 earthquake, given WG03's best estimates of its  $L$ ,  $W$ , and  $A$ , which in turn drove up the rate of occurrence of our 1906 facsimiles ( $M \sim 7.6$ , not  $\sim 7.8$ ) to something like once every 90 years. We knew that back in 1999 (see OFR 99-517, Appendix C) which provides Bill E's early analysis of  $M - \log A$  relations. In  $M - \log A$  space a line of unit slope is a line of constant stress drop, with the intercept value being proportional to this stress drop. Bill E was able to reach the right magnitudes for events of large  $A$  by assuming larger stress drops for them, but at the price of overestimating magnitudes at  $M \sim 6.7$  and smaller. This was an acceptable solution for WG03, because we were dealing with  $M$  of 6.7 and larger. The Hanks and Bakun bilinear forms, however, match the WC94 data at  $M$  both above and below 6.7, and, quite remarkably, return the average earthquake stress drop of 30 bars (26.7 bars in HB02) for  $M < 6.7$ . But this is a more complicated model, and the empirical validity of the  $L$ -model part depends mostly on the few data for  $M \geq 7.5$ ; fortunately, 5+ years later, there are more of these to work with.

Paul has evidently been commissioned by the CEA project to revisit/review the WG03  $M - \log A$  relations, and his analysis is presented in his June 20, 2006 document. While assembling his data in a very different way from WC94, he came to essentially the same functional relationship for strike-slip earthquakes as did WC94,  $M = 1.02 (1.05) + 3.98 (3.87)$ , where the first number in each pair is from WC94 and the second, in parentheses, is Paul's June 20 estimate (Table 3 of his manuscript). Surprisingly, however, the use of this  $M - \log A$  relation did not give rise to an anomalous rate of 7.6-ish earthquakes relative to the Ellsworth B (4.2 intercept value) model that Ned showed as slides 5 and 4 in His November 1 ppt, a matter I will finish on.

There are many things that I do not understand about Paul's document, and one of them is the "this review weights" in his Table 7. WC94 has been dismissed from class, and it seems that zero weight has been placed on the  $M > 6.71$  (the  $L$ -model part) of Hanks and Bakun, but these weights don't add up to 1, either. I must be missing something here, but the present appearances are that Paul is dismissing the two published papers in favor of two unpublished ones, a curious turn of events for a project of this visibility. How did this happen?

In part because of the situation displayed in Figure 3 of the June 20 document. Here we see Paul's data set, together with five  $M - \log A$  relations. The best-fitting model, presumably, is Somerville (2006), assuming it has been fit to his own data set. Which is the worst-fitting model? Without the benefit of numbers to deal with, I will hazard a guess that it is the WGCEP (2002) model, which is the Ellsworth model with intercept of 4.2 (see second paragraph). It sits below *all* of the data. These models, remember, are getting most/all of the weight recommended by Paul (I think, see previous paragraph).

Bill E used very different area estimates for his earthquakes, at least some of which are the same earthquakes that Paul used. But the only way for both of these models to be admissible is for the uncertainty in the area estimates to be at least a factor of 2, maybe 3. So my suggestion is that Bill E and Paul spend some time trying to resolve their differences on the area estimates for these 16 earthquakes. I will guess that this is much more important than what you have asked Bill B and me to do, since HB02 seems to be a marginal player in the CEA analysis. And if the uncertainty in area estimates is really this large, reliably estimating  $M$  from  $\log A$  may be a false dream.

This is a good place to note that the situation above illustrates why it is so dangerous to choose a single model in an exercise of this sort. Sure, aggregating several models is unlikely to provide the "right" answer, but the flip side of this coin is that you are much less likely to be really wrong, as is always possible when only one model is in play for a problem fraught with so much uncertainty.

This is also the place to emphasize again the importance of compiling a database that includes *everybody's* estimates for  $L$ ,  $W$ , and  $A$ , which should include the Wesnousky data set that Steve presented at the SCEC meeting in September. We need to find out objectively what kind of uncertainties we are really dealing with. Depending on what comes out of this, we may have to move on to weighting of these estimates on the basis of the reliability/resolution of the technique(s) used to obtain them. Bill E's graphic comparing depth ranges for the Denali earthquake determined from aftershocks, geodetic observations, and fault-slip inversions nicely illustrates the problems we have to deal with; inconsistencies among the various methodologies, non-uniquenesses, and resolving power issues. Both Jessica M and Ken H contributed important information on geodetic resolution that I have not fully assimilated. But we won't get all this done by the CEA deadlines.

In the meantime, we are left with the following situation according to Paul. WC94 is dismissed because "This model is represented by Model 7 described below" (p. 16 of the

June 20 document). Model 7 is Somerville (2006b) in Table 7, essentially the same as the WC94 *predictive model*. Left unsaid here, however, is that the WC94 *database*, the gold standard of estimating  $M$  from  $A$ , has also been dismissed, in favor of 16 earthquakes (Table 4), the seismologically determined, fault-slip inversions for which have been personally selected and trimmed by Paul. If the CEA project is to follow along these lines, it will need to provide better justification for this decision than has Paul.

Paul also made it very clear that he doesn't like the  $L$ -model because of the ground motions he calculates for it at  $M = 8.2$ . This is 0.3 magnitude unit larger than any earthquake likely to happen in California in its present circumstances—and larger than any moment magnitude reported for any continental strike-slip earthquake anywhere on the planet. And who knows what ground motions will arise from such an event even if it does occur, given the comparisons that Steve Day is making for dynamic models vis a vis kinematic models for great strike-slip earthquakes that we saw at the SCEC meeting and again at the ExGM meeting two months ago.

HB02 was fully cognizant of the implications of the  $L$ -model scaling for earthquake ground motion, namely the static stress drop increasing in magnitude ranges where the dynamic stress drops seem to be flat if not decreasing. I agree that ground motion predictions should be made for all of the (static)  $M - \log A$  relations, but they should be made in the presence of both median values and the intrinsic variability of the observed ground motions. We also know that high-frequency ground motion conforms to an  $a_{rms}$  stress drop of  $\sim 100$  bars, no matter what the static stress drop we may have obtained. Ken summarized this situation well: large earthquakes are/may be just different from the smaller ones, realized perhaps in their  $M - \log A$  relations, perhaps in their ground motion excitation, and/or perhaps in ways we can only guess at. But we won't decide this matter before the CEA deadlines, either.

To return to Slides 4 (Ellsworth 4.2) and 5 (Somerville, 2006) magnitude/frequency-of-occurrence plots that Ned showed. These plots, I would have thought, should have expressed greater differences, for the reasons given in the third paragraph. The principal difference between them, however, exists only above  $M > 7.8$ , at which point Ellsworth B begins to stand noticeably above Somerville. At  $M = 7.9$ , for example, the Ellsworth calculation yields  $\sim 4.5$ /ka events and the Somerville calculation has just 1.8/ka events of this  $M$  or greater; so these calculations do not seem to be moment (rate) balanced to the same number. In talking to Ned on November 3 about this matter, he mentioned that there is also a recurrence-interval constraint. This seems like a good point to sign off, so to let Ned explain how that is/may be affecting these calculations and to let Paul—and any other of the recipients of this communication—to respond to what they see here.

Tom  
November 8, 2006

