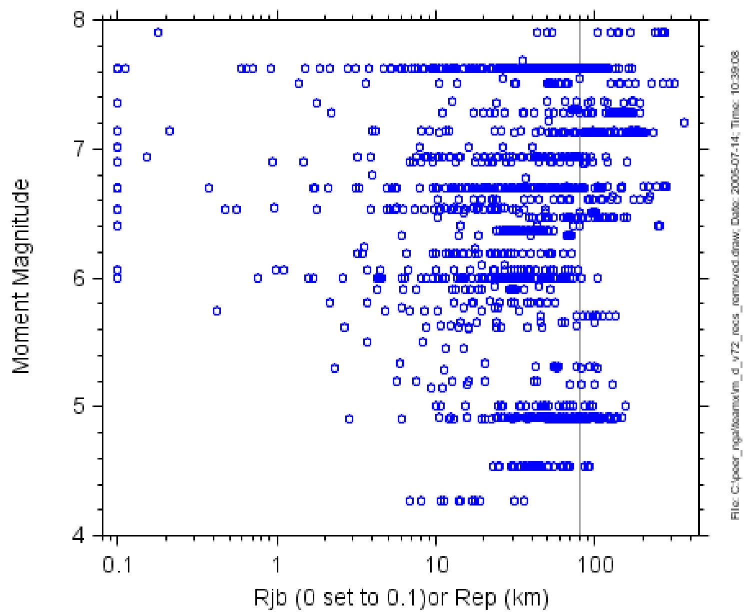
25 July 2005: Dave's notes on work done for PEER NGA project
File: \peer_nga\teamx\daves_notes_mechrake_c3_stage1_stage2_pga_25july2005.doc

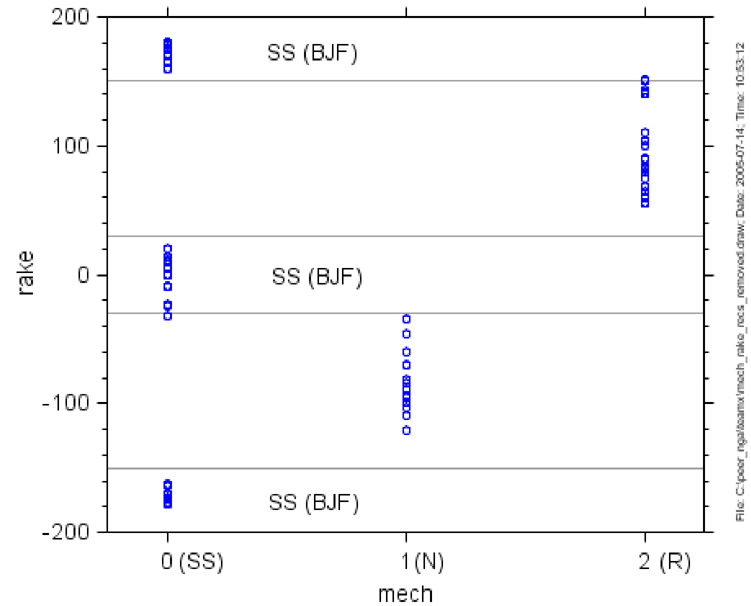
These notes cover 1) the choice of the mechanism determination compared to the BJF criteria (brief discussion), 2) the choice of the "anelastic" term (c3), and 3) the development of a "rock" pga prediction equation for use in the nonlinear corrections (to be used in Gail's scheme, which uses a modification of the Choi and Stewart nonlinear amps along with the BJF amps).

First, here is the M-D space we are working with (where in this and all other plots I have used only "flag = 0" records).



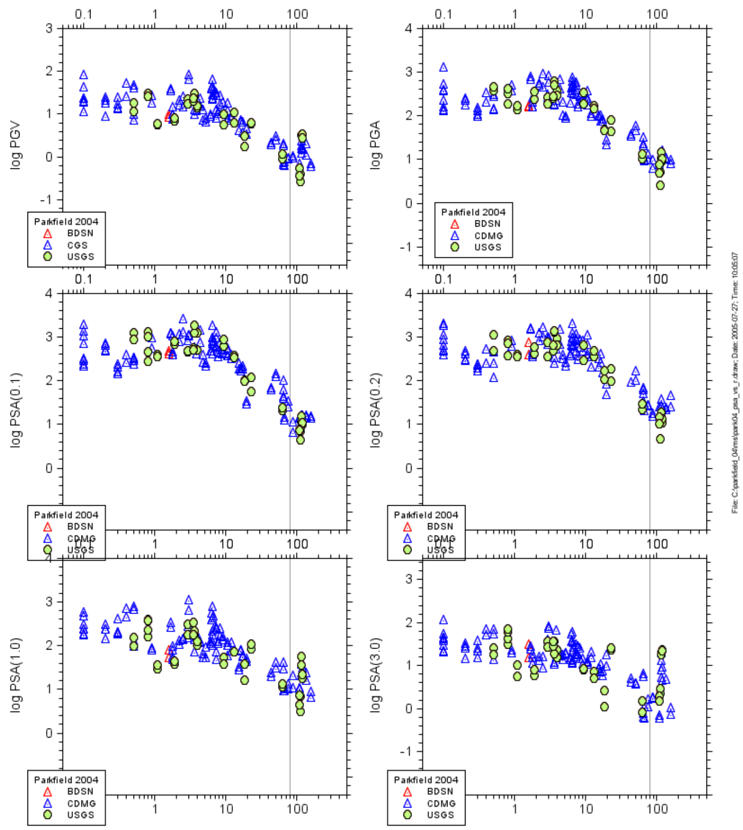
Mechanism Class:

The separation of quakes into strike slip, normal, and reverse slip is based on the P and T axis plunges, as proposed in the previously distributed report C:\peer_nga\database\fault_classification_using_p_t_axes.pdf. I wanted to see how the groupings compare to classification based on rake angle alone, using the BJF criteria that strike slip events have rake angles within 30 degrees of horizontal. The comparison below shows that with one or two exceptions, the groupings would be the same:



Anelastic coefficient:

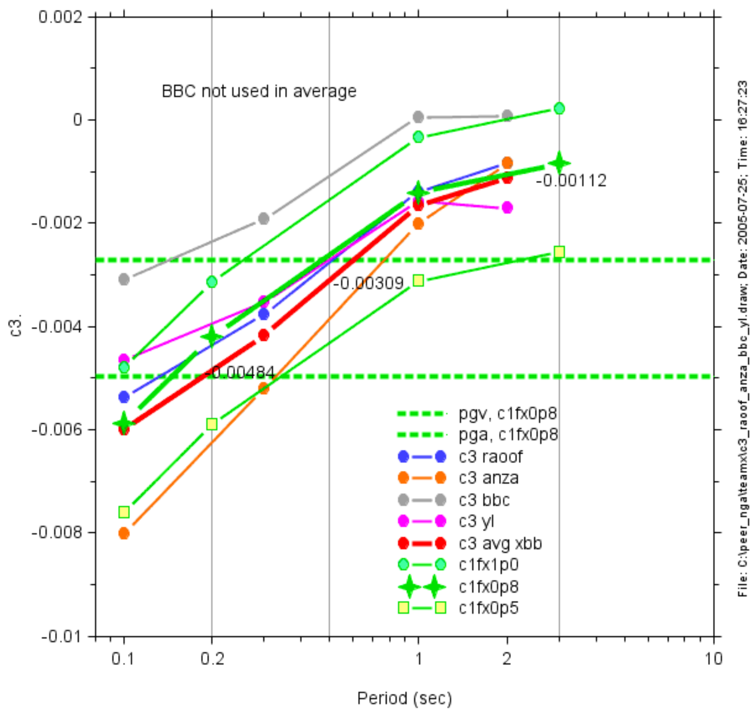
I used data gathered by Linda Seekins and Jack Boatwright for the Anza, Big Bear City, and Yorba Linda events (all M<5). The data were corrected to V30 760 m/s using the median values of site classes as assigned by Seekins and Boatwright. I computed pga, pgv, and psa for these events. I did NOT conbine both horizontal components. I also collected all of the 2004 Parkfield data that I could find (from the Berkeley Broadband Digital Network station close to the fault rupture, the CGS data, and the USGS data (which includes Borcherdt's GEOS data). I computed rjb, using a slip model Ji from his Caltech web site, as indicated on the CISN web site, for all stations. I made NO correction for site amps and I did NOT combine components. I used pga, pgv, and psa from data reporting agencies, if available, and otherwise computed them myself. Here is a plot of the data with different symbols for the different data providers:



I first ran the stage1 regression on the NGA dataset, restricted to V30 > 360 and r < 80 (using BJF to correct to V30=760 m/s), with only the c1 coefficient free to vary. This is an extension of work reported in C:\peer_nga\teamx\Dave_NGA_work_13June05.doc. This gave c1 coefficients as follows:

per	ndata	ncoeff	h	res_sdev	c01
-1.00	546	58	7.00	2.128E-01	-9.494E-01
0.00	546	58	6.40	2.108E-01	-9.041E-01
0.10	546	58	7.80	2.300E-01	-1.017E+00
0.20	546	58	8.00	2.227E-01	-9.369E-01
1.00	535	55	4.60	2.450E-01	-8.017E-01
3.00	403	42	5.80	2.427E-01	-8.290E-01

In the above, ncoeff-3 gives the number of events used in the analysis, and res_sdev gives the standard deviation of the residuals; this is σ_1 of BJF. I then ran the stage 1 regression using the data for the four events (Anza, Big Bear City, Yorba Linda, Parkfield). The regression was done several times, fixing the "geometrical spreading" term c1 to values of -1.0, -0.8, and -0.5 (Note: I put "geometrical spreading" and "anelastic" in quotes because someone complained when I used these terms as if they have a strict physical meaning--- we have to remember to be careful about terminology in reports and papers so as not to mislead people unintentionally). The choices of -1.0 and -0.8 were motivated by the NGA stage 1 regression. I found that all choices of c1 gave equally good fits to the data. Here is a confusing plot of the c3 factors for the various c1 choices:



The “c3 raoof” values came from plotting their combination of geometrical spreading and Q vs. r and using CoPlot to fit regression curves, solving for both the c1 and c3 coefficients. After considerable thought and looking at plots, I decided to use the c3 values for the c1=-0.8 regression of the four events, given by large green crosses (actually, I first decided on the c3 values for c1=-0.5 and did a lot of analysis, then reversed myself because the distance decay is about the same --- a bit steeper for the -0.5 case, which is why I chose it initially, but the geometrical spreading term is smaller than I would like). Here are the coefficients for the small earthquake stage 1 regression analysis for the three fixed values of c1:

per	ndata	h	res_ave	res_sdev	c01	c03
-1.00	1232	6.40	2.492E-07	3.135E-01	-1.000E+00	-1.629E-03
0.00	1232	7.20	-3.544E-07	3.363E-01	-1.000E+00	-3.890E-03
0.10	1232	6.80	2.843E-07	3.506E-01	-1.000E+00	-4.794E-03
0.20	1232	7.80	5.948E-07	3.581E-01	-1.000E+00	-3.150E-03
1.00	1232	6.60	9.086E-08	3.256E-01	-1.000E+00	-3.283E-04
3.00	1232	7.20	5.602E-08	3.116E-01	-1.000E+00	2.188E-04

per	ndata	h	res_ave	res_sdev	c01	c03
-1.00	1232	4.80	-1.296E-07	3.152E-01	-8.000E-01	-2.726E-03
0.00	1232	5.60	-2.119E-07	3.371E-01	-8.000E-01	-4.969E-03
0.10	1232	5.20	2.046E-07	3.512E-01	-8.000E-01	-5.875E-03
0.20	1232	6.00	-1.180E-07	3.585E-01	-8.000E-01	-4.209E-03
1.00	1232	4.80	1.605E-07	3.274E-01	-8.000E-01	-1.412E-03
3.00	1232	5.20	-6.038E-08	3.139E-01	-8.000E-01	-8.474E-04

per	ndata	h	res_ave	res_sdev	c01	c03
-1.00	1232	2.40	1.822E-07	3.205E-01	-5.000E-01	-4.452E-03
0.00	1232	3.20	2.105E-07	3.406E-01	-5.000E-01	-6.672E-03
0.10	1232	3.20	2.475E-07	3.546E-01	-5.000E-01	-7.606E-03
0.20	1232	3.60	-2.161E-07	3.612E-01	-5.000E-01	-5.907E-03

1.00	1232	2.20	-8.970E-08	3.323E-01	-5.000E-01	-3.132E-03
3.00	1232	2.60	-3.904E-08	3.197E-01	-5.000E-01	-2.564E-03

And here are the regression coefficients for the NGA dataset, using the small earthquake c3 values (rounded to two significant digits):

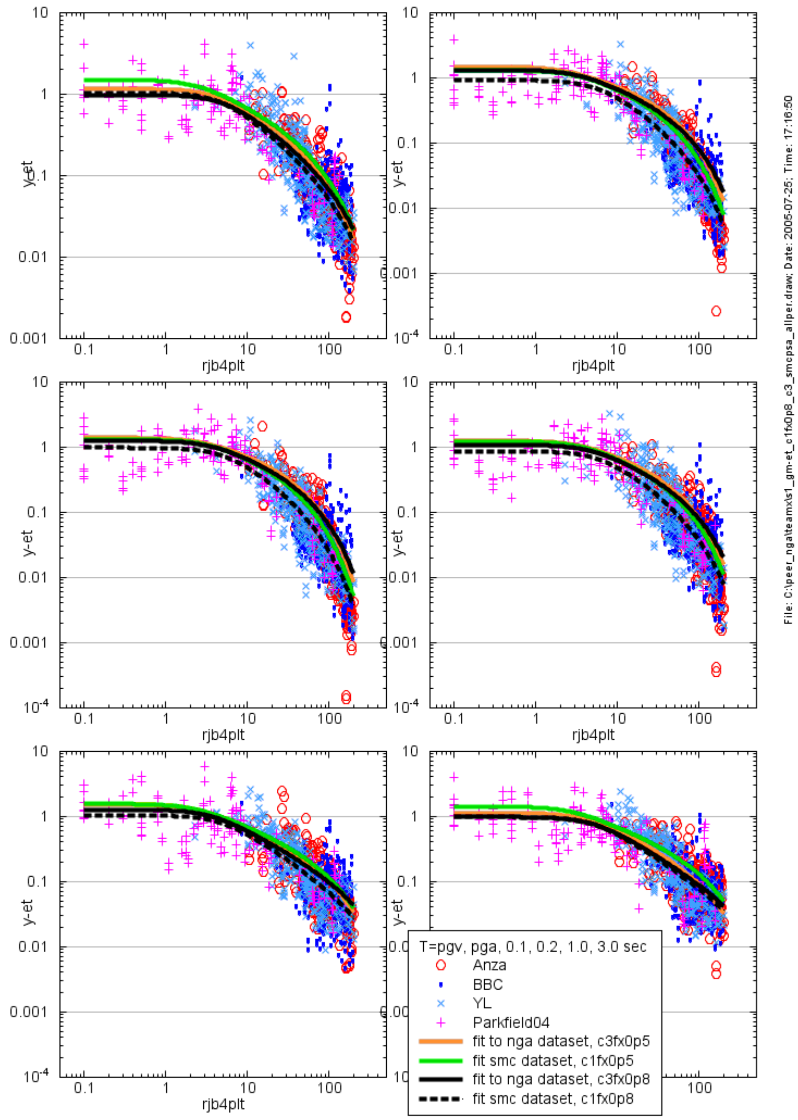
per	ndata	ncoeff	h	res_ave	res_sdev	c01	c03
-1.00	546	58	5.20	3.100E-08	2.135E-01	-7.094E-01	-2.700E-03
0.00	546	58	3.00	3.278E-07	2.104E-01	-4.868E-01	-5.000E-03
0.10	546	58	3.40	2.677E-07	2.290E-01	-4.911E-01	-5.900E-03
0.20	546	58	4.40	2.856E-07	2.224E-01	-5.447E-01	-4.200E-03
1.00	535	55	3.60	-2.223E-07	2.447E-01	-6.810E-01	-1.400E-03
3.00	403	42	5.20	-4.785E-08	2.428E-01	-7.540E-01	-8.500E-04

Note that the h values and the geometrical spreading terms are smaller than for the corresponding regression on the four events.   There are tradeoffs in h(T), c1, and c3--- the three factors were not completely consistent to the small event and NGA stage 1 regression runs--- very frustrating. I used a function in CoPlot to play with equations, varying terms to see if I could find consistency, but I failed to do so.

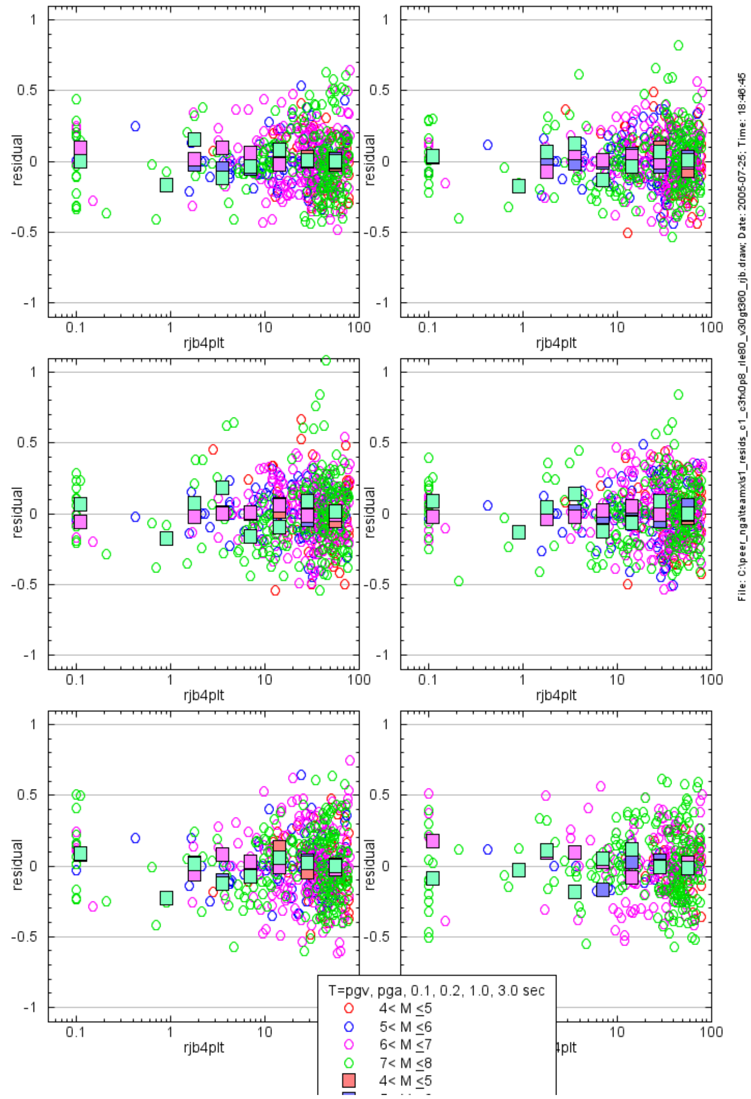
Here is a plot of the four-event data, after removing the event term, with the predicted ground motion as a function of distance. Note that the stage 1 equation is

Log Y = c1*log(r/rref) + c3*(r-rref) + event terms,

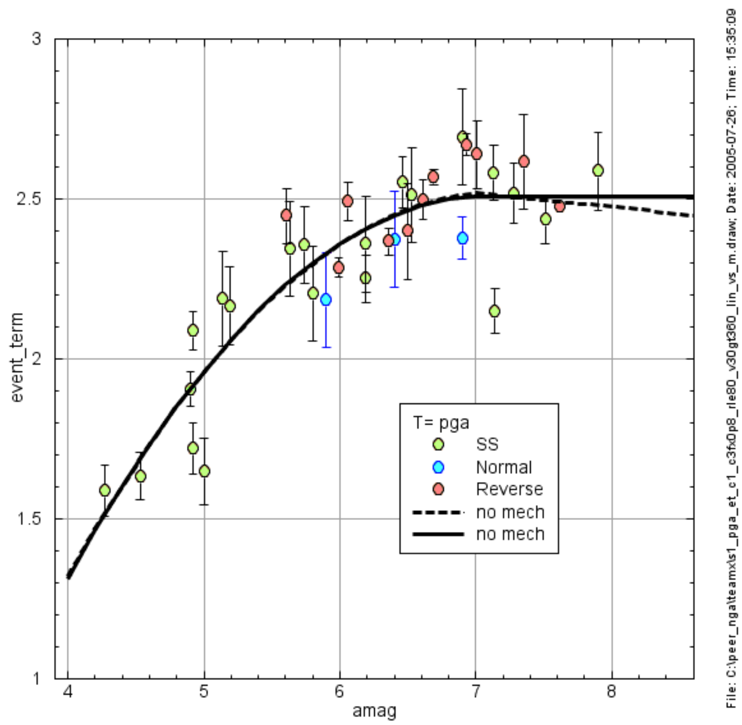
where r = sqrt(rjb^2+h(T)^2).   I include “rref” so that the event terms are just the value of log Y for r = rref. The problem is that because h(T) varies with period and differs for the small event and NGA datasets, the predicted ground motion from the NGA analysis in the graph below are not always centered on the data. The plot uses event terms from the regression on the four-event dataset assuming c1 = -0.8; the curves are for regression on both the NGA and four-event datasets, using the -0.5 and -0.8 values. Recall that for the four-event dataset, the c1 coefficient was fixed at -0.5 and -0.8, so the curves are labeled “c1fx0p5” and “c1fx0p8”. For the NGA dataset, the c3 coefficients are fixed as those from the four-event data, so the curves are labeled “c3fx0p5” and “c3fx0p8”.

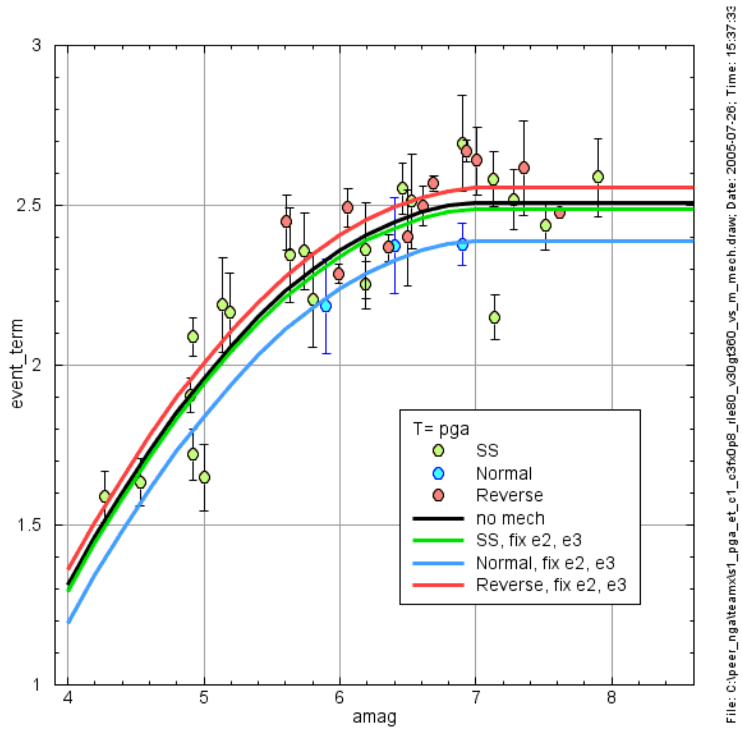


Note that I should have determined c3 from FAS decay, but used psa instead. Will this make a difference? I recall that simulations suggest that larger M decay less rapidly. If so, the attenuation obtained from analyzing small earthquake PSA may not be representative of that from larger earthquakes --- in other words, there is a magnitude dependence in the attenuation. I looked at this by plotting residuals and binned residuals as a function of distance. The results are shown below. I do not see any magnitude dependence in the residuals.



Having decided on the stage1 coefficients, I ran the stage2 analysis for pga, using the NGA dataset with r < 80 and V30 > 360 (pga's were corrected using BJF site amps). I did the stage 2 regression assuming no magnitude dependence for M above 7.0, and with a linear M dependence above M 7.0. Only events with two or more observations were used. In the regressions shown in the plot below I lumped all mechanisms together, but I've indicated the mechanism for each event term by the color. Following this is a plot in which I fixed the magnitude dependence to be equal to that for the lumped-mechanism run, and I solved for mechanism factor.



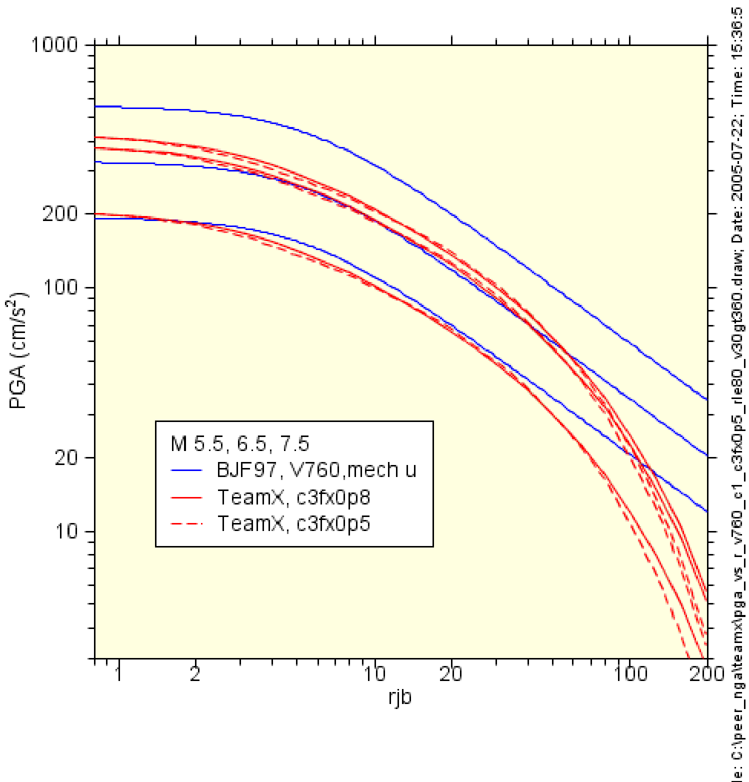


Given the small number of normal mechanism data and the small difference in strikeslip vs reverse, I decided to use the combined mechanism results for purposes of deriving an equation for pga that can be used in the nonlinear site amplification analysis.

Here is a comparison of pga predictions using

1. the equation I derived from my stage1 and stage2 analysis of the PEER dataset, constraining c3, and using data for r < 80 and V30 > 360 (pga's were corrected using BJF site amps), and a stage2 regression with no magnitude dependence for M above 7.0, and with a linear M dependence above M 7.0.

2. the BJF equations for V30=760 and mechanism unspecified.



This comparison can also be used to judge whether the c3fx0p5 or c3fx0p8 results should be used (note that the c3fx0p5 curves decay more rapidly than the c3fx0p8 curve fat greater distances).

Here is the result of the stage 2 regression for pga:

per	amhinge	nevents	sig2	e01	e02	e03
0.00	7.00	35	1.116E-01	2.506E+00	2.200E-02	-1.254E-01

Based on all of the above,

pga (cm/s/s) = 10^(2.506+0.02200*(M-7)-0.1254*(M-7)^2-0.4868*LOG(R/5)-0.005*(R-5))

for M <= 7

and

pga (cm/s/s) = 10^(2.506 -0.4868*LOG(R/5)-0.005*(R-5))

for M > 7, where

r = sqrt(rjb^2+3^2)

and stage1, stage2, and total sigmas of 2.104E-01, 1.116E-01, 0.24.

I conclude with a figure showing the stage1 event terms for pgv, pga, and psa for the four oscillator periods, with simulated magnitude scaling superposed (shifted vertically to more-or-less go through the cloud of points). Of particular interest for the Developer's meeting on July 27 is the plot for T = 3 sec, because that plot also includes the simulated scaling for T = 9 s. As expected, the scaling is the same for small enough earthquakes (less than about M = 7), because the response spectra are controlled by periods greater than the corner period of the earthquakes.