Epistemic Uncertainty Model for Use of PEER-NGA Ground Motion Models in National Hazard Mapping

Sources of Epistemic Uncertainty

1. Uncertainty in selecting the appropriate database
2. Uncertainty in selecting the appropriate model formulation
3. Uncertainty in estimating the population mean with a finite data set
4. Uncertainty in estimating the population mean due to uncertainty in the predictor variables in the dataset (magnitude, distance, V_{S30}, etc.)

Epistemic Uncertainty

- Components 1 and 2 addressed by multiple PEER-NGA models
  - Differences in selection of appropriate data
  - Differences in selection of appropriate model formulations

Epistemic Uncertainty

- Components 3 and 4 can be assessed for each model
  - Component 3 a function of selected data
  - Component 4 a function of data uncertainty

Uncertainty in Mean of a Sample

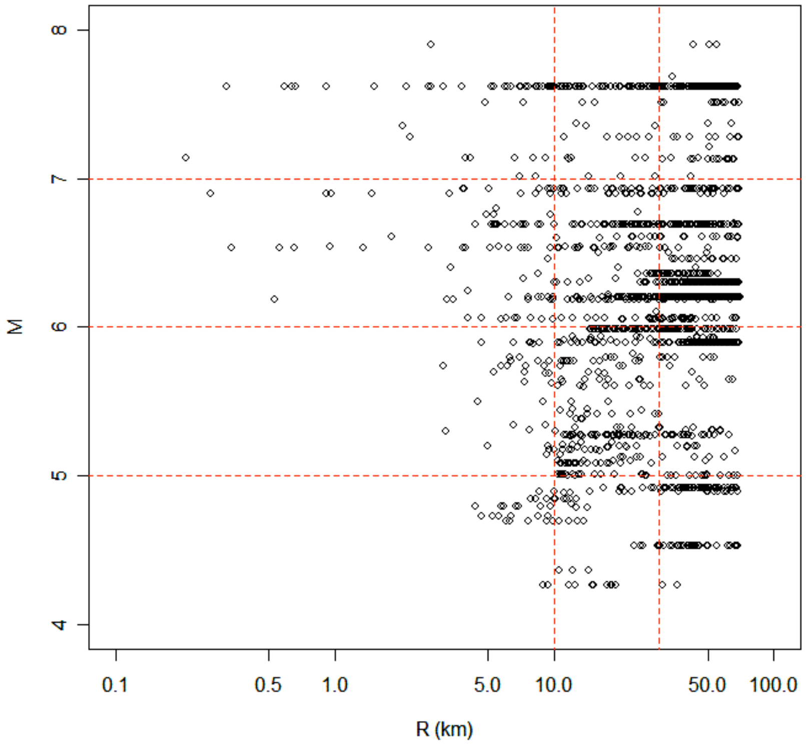
- Estimate of mean of a sample

$$\sigma_{\bar{x}} = \sqrt{\frac{\sigma_x^2}{n}}$$

- Extension to ground motion data with nested sampling

$$\sigma_{\overline{\ln[PGA]}} = \sqrt{\frac{\tau^2}{n_{Eq}} + \frac{\sigma^2}{n_{Sites}}}$$

Example from Chiou and Youngs (2006)



M and R_{RUP} Range	Average **M**	Median R_{RUP}	n_{Eq}	n_{Sites}	
M < 5 $R_{RUP} < 10$	4.7	7.3	6	25	0.156
M < 5 $10 \leq R_{RUP} < 30$	4.7	15.7	12	47	0.112
M < 5 $R_{RUP} \geq 30$	4.8	45.4	5	98	0.139
5 ≤ **M** < 6 $R_{RUP} < 10$	5.6	7.0	23	32	0.109
5 ≤ **M** < 6 $10 \leq R_{RUP} < 30$	5.5	16.9	50	263	0.052
5 ≤ **M** < 6 $R_{RUP} \geq 30$	5.7	46.5	26	241	0.065
6 ≤ **M** < 7 $R_{RUP} < 10$	6.5	4.4	24	78	0.083
6 ≤ **M** < 7 $10 \leq R_{RUP} < 30$	6.5	20.1	26	210	0.067
6 ≤ **M** < 7 $R_{RUP} \geq 30$	6.3	48.8	23	670	0.063
M ≥ 7 $R_{RUP} < 10$	7.5	4.1	7	45	0.133
M ≥ 7 $10 \leq R_{RUP} < 30$	7.5	17.9	8	71	0.119
M ≥ 7 $R_{RUP} \geq 30$	7.5	50.3	10	170	0.099

Results for PEER-NGA Models

M and R_{RUP} Range			
	Boore and Atkinson	Campbell and Bozorgnia	Chiou and Young s
M < 5 $R_{RUP} < 10$	0.313		0.156
M < 5 $10 \leq R_{RUP} < 30$	0.155	0.133	0.112
M < 5 $R_{RUP} \geq 30$	0.124	0.105	0.139
5 ≤ **M** < 6 $R_{RUP} < 10$	0.176	0.193	0.109
5 ≤ **M** < 6 $10 \leq R_{RUP} < 30$	0.084	0.073	0.052
5 ≤ **M** < 6 $R_{RUP} \geq 30$	0.078	0.073	0.065
6 ≤ **M** < 7 $R_{RUP} < 10$	0.097	0.077	0.083
6 ≤ **M** < 7 $10 \leq R_{RUP} < 30$	0.080	0.062	0.067
6 ≤ **M** < 7 $R_{RUP} \geq 30$	0.068	0.058	0.063
M ≥ 7 $R_{RUP} < 10$	0.124	$\sigma_{\overline{\ln[PGA(m,r)]}}$ 0.111	0.133
M ≥ 7 $10 \leq R_{RUP} < 30$	0.105	0.093	0.119
M ≥ 7 $R_{RUP} \geq 30$	0.086	0.065	0.099

Assessment of Constraints on Model Parameters

- Assessment of uncertainty in parameters is given by asymptotic standard errors and correlations
- The non-linearity in the model and the high degree of correlation between some parameters makes interpretation difficult
- Use a non-parametric simulation approach
  - Bootstrap standard errors of predictions

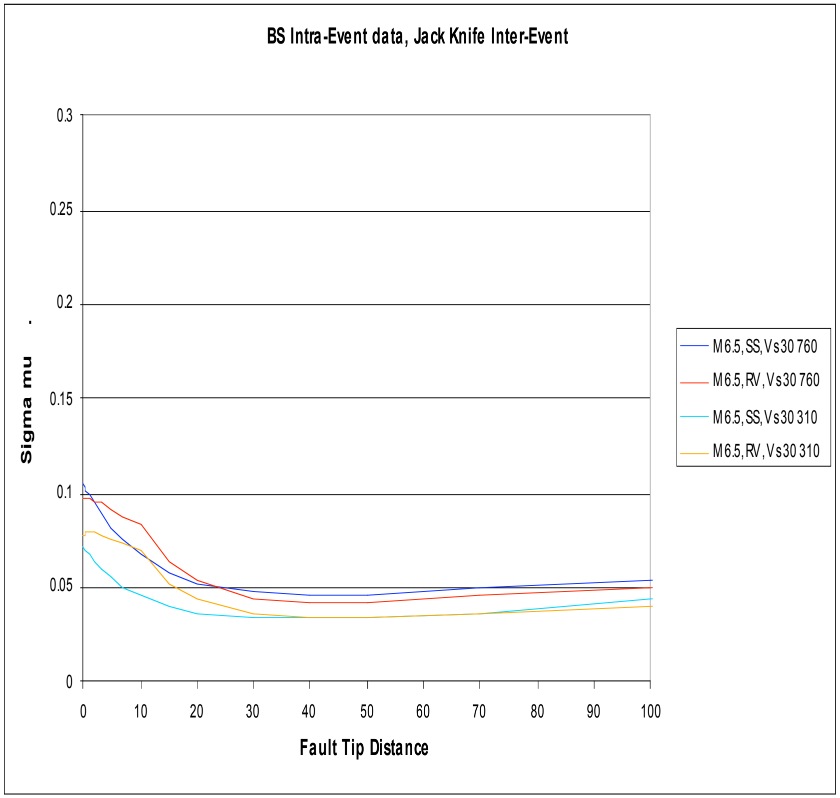
Bootstrap

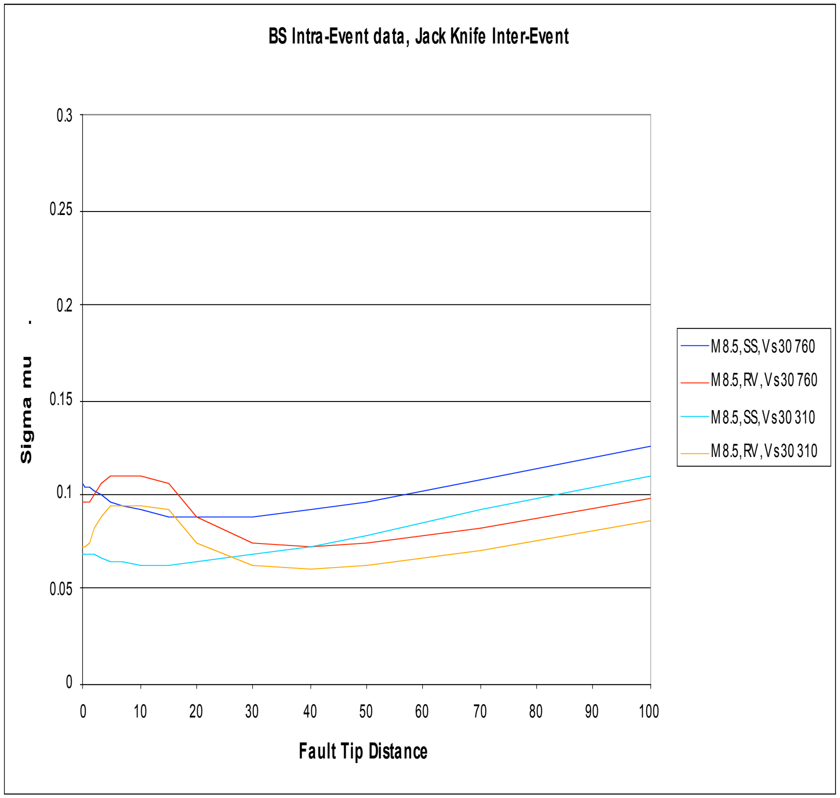
- Assume that observed data defined an empirical joint distribution for predictor variables and response
- Sample with replacement from this empirical distribution to generate new data set
- Fit this simulated data with functional form and obtain ground motion predictions for a range of M, R, Style of Faulting, etc.
- Repeat the process multiple times and compute the standard deviation of the predictions. This provides an estimate of σ_μ

Bootstrap

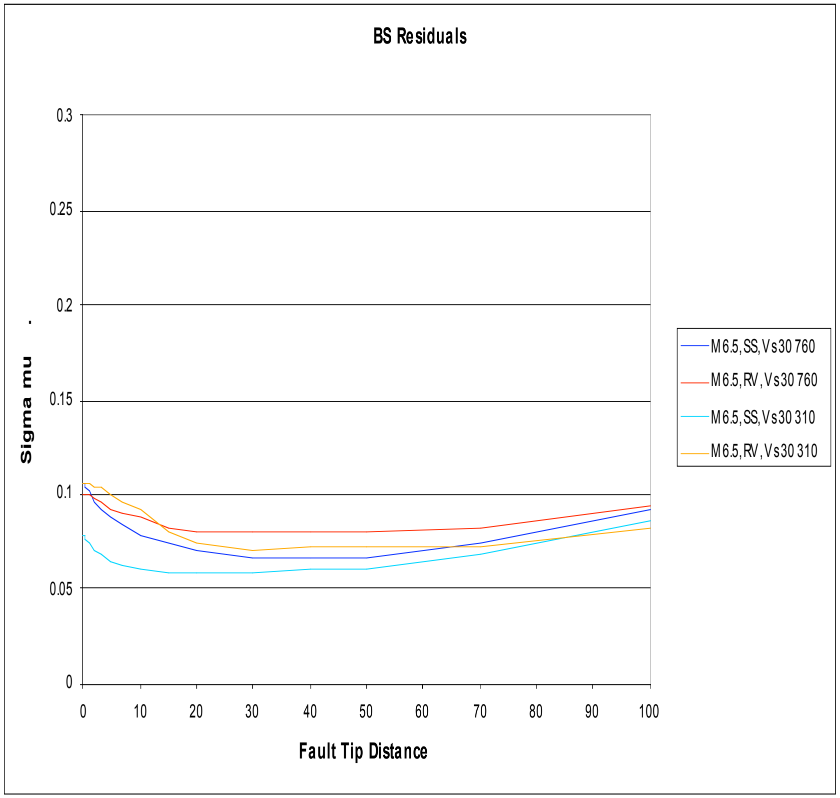
- Two approaches for regression analysis
  - Re-sampling data with replacement
  - Re-sampling residuals and adding these residuals to the model predictions (this breaks the tie between the predictor variables and the response)
- Approaches for sampling from nested data (i.e. correlated data) are not well developed

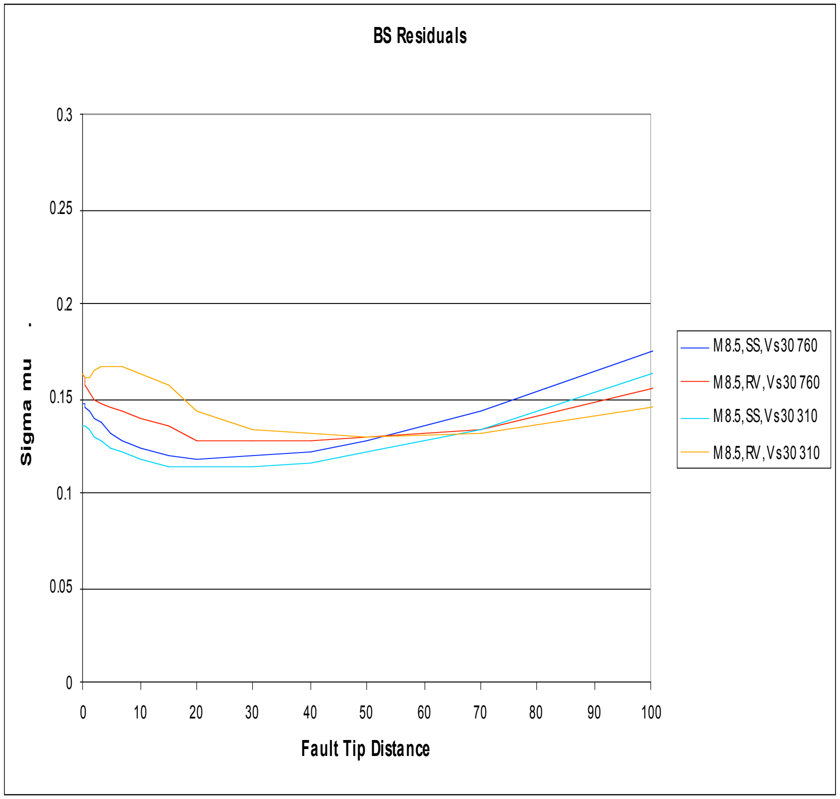
BS Intra-Event data, Jack Knife Inter-Event



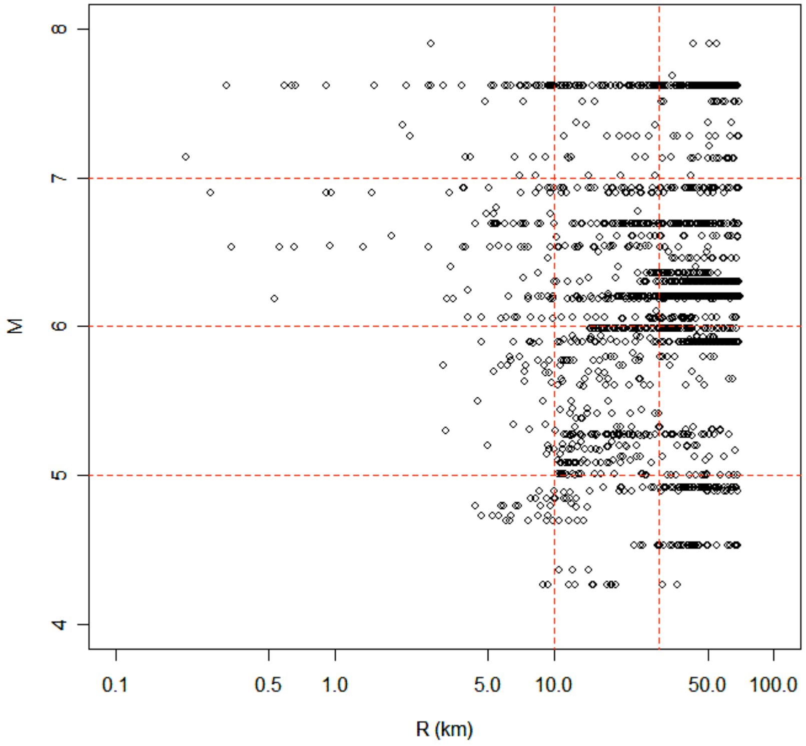


BS Residuals





Example from Chiou and Youngs (2006)



$\mathbf{M}$ and R_{RUP} Range	Average $\mathbf{M}$	Median R_{RUP}	n_{Eq}	n_{Sites}	
$\mathbf{M} < 5$ $R_{RUP} < 10$	4.7	7.3	6	25	0.156
$\mathbf{M} < 5$ $10 \le R_{RUP} < 30$	4.7	15.7	12	47	0.112
$\mathbf{M} < 5$ $R_{RUP} \ge 30$	4.8	45.4	5	98	0.139
$5 \le \mathbf{M} < 6$ $R_{RUP} < 10$	5.6	7.0	23	32	0.109
$5 \le \mathbf{M} < 6$ $10 \le R_{RUP} < 30$	5.5	16.9	50	263	0.052
$5 \le \mathbf{M} < 6$ $R_{RUP} \ge 30$	5.7	46.5	26	241	0.065
$6 \le \mathbf{M} < 7$ $R_{RUP} < 10$	6.5	4.4	24	78	0.083
$6 \le \mathbf{M} < 7$ $10 \le R_{RUP} < 30$	6.5	20.1	26	210	0.067
$6 \le \mathbf{M} < 7$ $R_{RUP} \ge 30$	6.3	48.8	23	670	0.063
$\mathbf{M} \ge 7$ $R_{RUP} < 10$	7.5	4.1	7	45	0.133
$\mathbf{M} \ge 7$ $10 \le R_{RUP} < 30$	7.5	17.9	8	71	0.119
$\mathbf{M} \ge 7$ $R_{RUP} \ge 30$	7.5	50.3	10	170	0.099

Effect of Predictor Variable Uncertainty

- Simulate new data sets with predictor variables estimated from standard errors reported in Flat File
- Fit this simulated data with functional form and obtain ground motion predictions for a range of M, R, Style of Faulting, etc.
- Repeat the process multiple times and compute the standard deviation of the predictions. This provides an estimate of σ_μ due to data uncertainty
- Combine this with bootstrap estimate (add variances or directly in combined simulation)

Epistemic Model for USGS Mapping

- Multiple PEER-NGA models provide partial estimate of epistemic uncertainty across models
- Bootstrap type analyses provide an assessment of epistemic uncertainty for a given model
  - Use to construct estimate of σ_μ(M, R, etc.)