Example: Ground Motion Attenuation

- **Problem:** Predict the probability distribution for Peak Ground Acceleration (PGA), the level of ground shaking caused by an earthquake
  - Earthquake records are used to update the predictive probability for PGA based on earthquake magnitude M, distance to the fault d, and local soil conditions $G=(G_B,G_C)$ (soil classified A, B or C)
  - Well-known attenuation relation developed by Boore, Joyner and Fumal (BJF, 1993). Let $R=\sqrt{d^2+h^2}$, h= fictitious depth:

$$\log_{10}(PGA) = b_1 + b_2(M-6) + b_3(M-6)^2 + b_4 R + b_5 \log_{10}(R) + b_6 G_B + b_7 G_C + \varepsilon$$

$$\triangleq q(M,d,G;b,h)+\varepsilon \qquad \text{(where } b=(b_1,\ldots,b_7)\text{)}$$

$\varepsilon =$ uncertain prediction error modelled as $N(0,\sigma)$ (max. entropy model)

Example: Ground Motion Attenuation

- **Goal**: Estimate $p(PGA \mid U, \mathrm{D_N}, \mathrm{M})$ for input $U = (M, d, G)$
  - Available set of data $\mathrm{D_N} = \{U_n, Y_n : n = 1, \ldots, N\}$ where $U_n = (M_n, d_n, G_n)$ (magnitude, distance, soil conditions), and corresponding $Y_n = (PGA)_n$ are data from earthquakes at various sites (we use N=271 ground motion records from 20 earthquakes)
  - Model class M is $p(PGA \mid U, \theta)$ from BJF model with specified prior PDF over the model parameters $\theta = (b, h, \sigma)$
- **Robust posterior predictive probability model:**

$$p\left(PGA \middle| U, \mathrm{D_N}, \mathrm{M}\right) = \int p\left(PGA \middle| U, \theta\right) p\left(\theta \middle| \mathrm{D_N}, \mathrm{M}\right) d\theta$$

Bayesian Updating

- **Bayes Theorem**: $p\left(\theta \middle| \mathrm{D}_N, \mathrm{M}\right) = c_N\, p\left(\theta \middle| \mathrm{M}\right) \prod_{n=1}^{N} p\left(Y_n \middle| U_n, \theta\right)$
- **Computing Optimal Posterior Predictive Model**
  - Find "optimal" (most probable values) of parameters $\hat{\theta}$ which maximize $p\left(\theta \middle| \mathrm{D}_N, \mathrm{M}\right)$, then to $O(\frac{1}{N})$ (Laplace's asymptotic approx.):

$$p\left(PGA \middle| U, \mathrm{D}, \mathrm{M}\right) \approx p\left(PGA \middle| U, \hat{\theta}\right)$$

  - Assumes M is globally identifiable on D_N ($\hat{\theta}$ is unique) and need a large amount of data N for an acceptable approximation (updated PDF will then have single sharp peak). Then no need to evaluate c_N and $\hat{\theta}$ is insensitive to the choice of prior PDF
  - Parameter estimation (i.e. using $\hat{\theta}$) is reasonable only under these conditions; otherwise, spurious reduction in uncertainty in predictions

Bayesian Updating (Continued)

- **Computing Robust Posterior Predictive Model**
  - Normalizing constant c_N, and $p\left(\theta \middle| \mathrm{D_N}, \mathrm{M}\right)$, is difficult to evaluate
  - However, if we can generate M samples $\left\{\tilde{\theta}_k, k=1,\ldots,M\right\}$ from $p\left(\theta \middle| \mathrm{D_N}, \mathrm{M}\right)$, then we can approximate the robust PDF by the corresponding sample mean:

$$p\left(PGA \middle| U, \mathrm{D_N}, \mathrm{M}\right) = \int p\left(PGA \middle| U, \theta\right) p\left(\theta \middle| \mathrm{D_N}, \mathrm{M}\right) d\theta$$

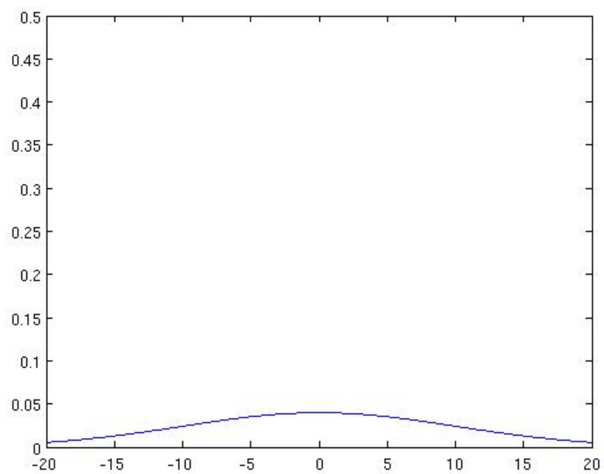
$$\approx \frac{1}{M} \sum_{k=1}^{M} p\left(PGA \middle| U, \tilde{\theta}_k\right)$$

  - These samples can be obtained using stochastic simulation methods, e.g. Gibbs sampler, Metropolis-Hastings (more later)

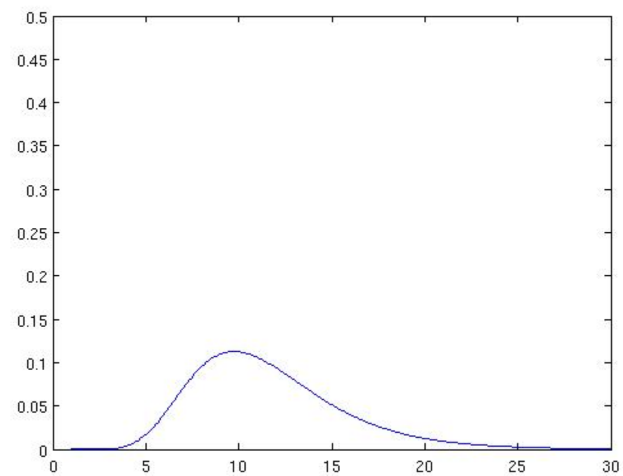
Sampling Posterior (Updated) PDF

- Samples generated using Markov Chain Monte Carlo
- Prior PDF chosen to reflect knowledge (or lack thereof)
  - For the regression coefficients b, take i.i.d. N(0,10), i.e. each has zero mean and standard deviation of 10 (very flat)
  - For the depth parameter h (km), lognormal distribution with mean and variance based on the depth of earthquakes in the data set

b_i's

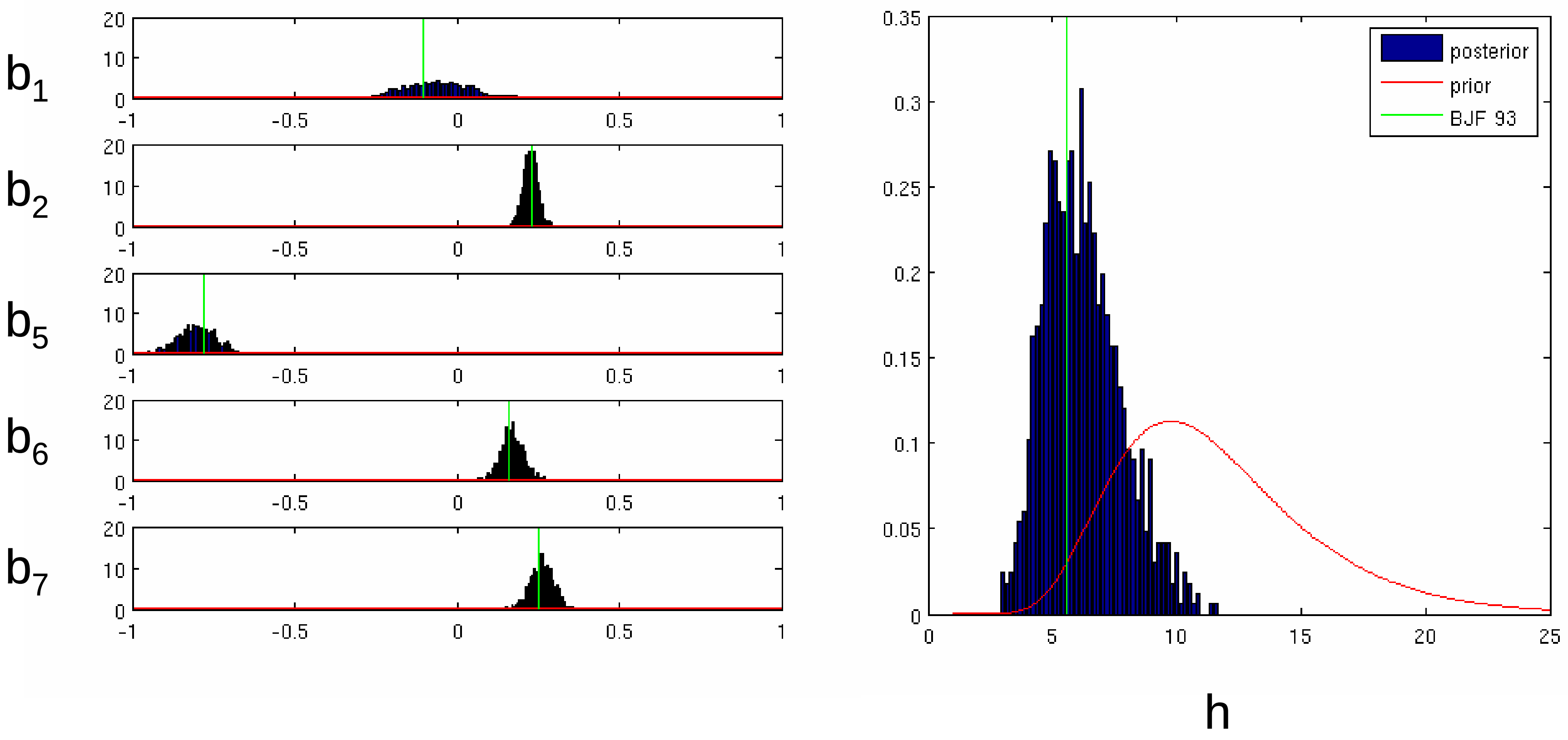


h

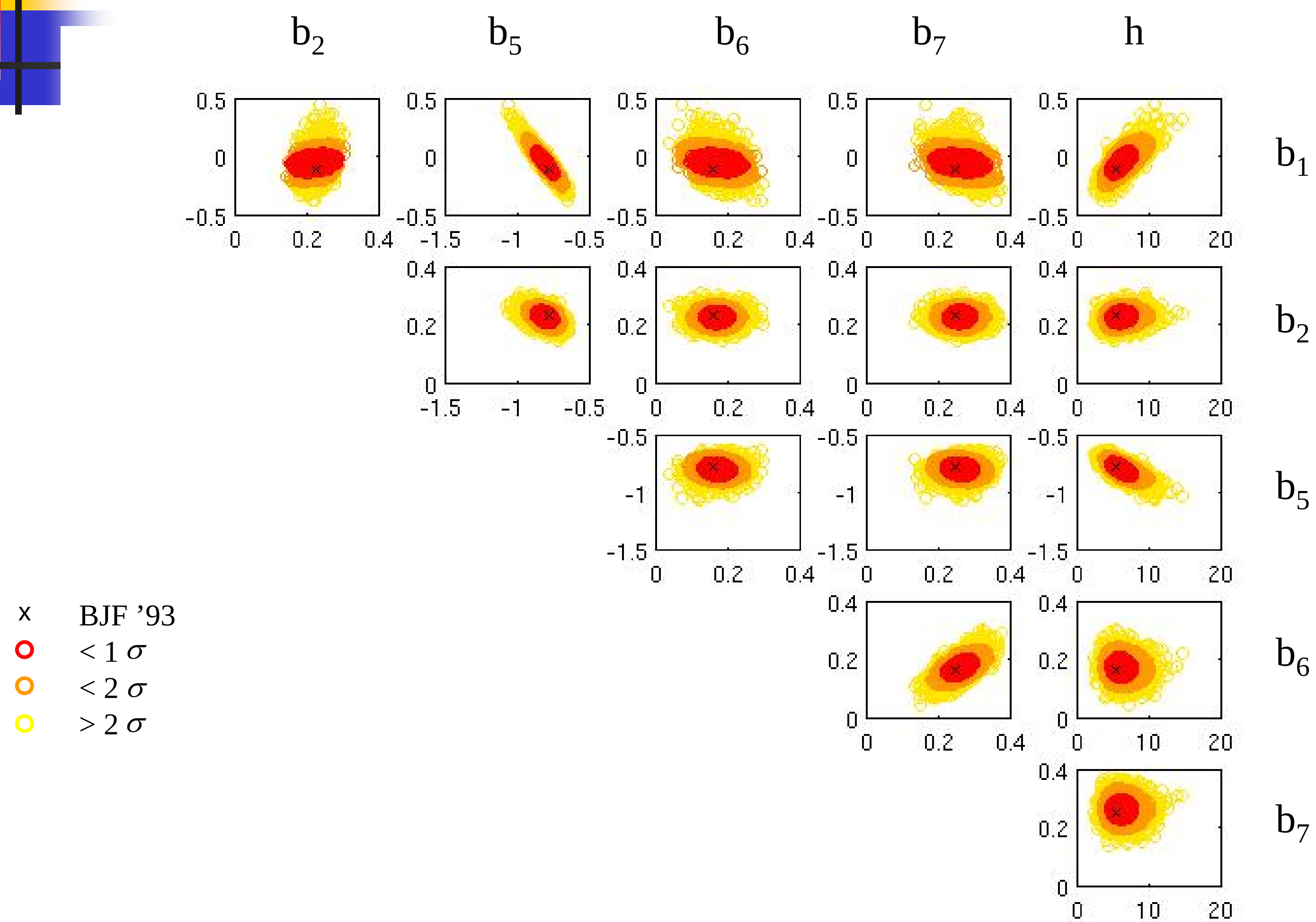


MCMC Samples from Posterior PDF

Model Class 4

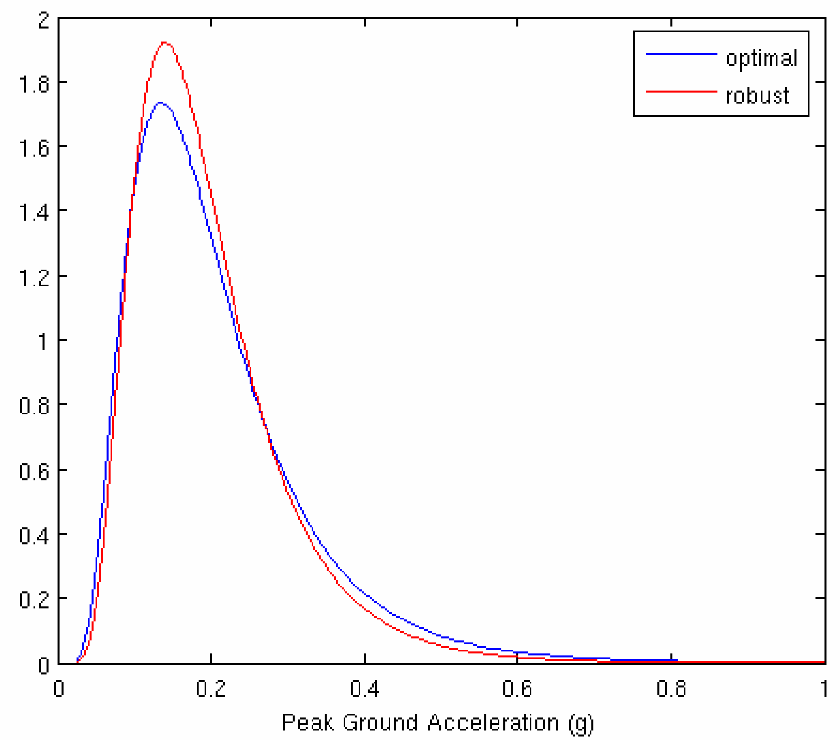


Posterior Samples: Model Class 4



Optimal vs. Robust Predictive Analysis

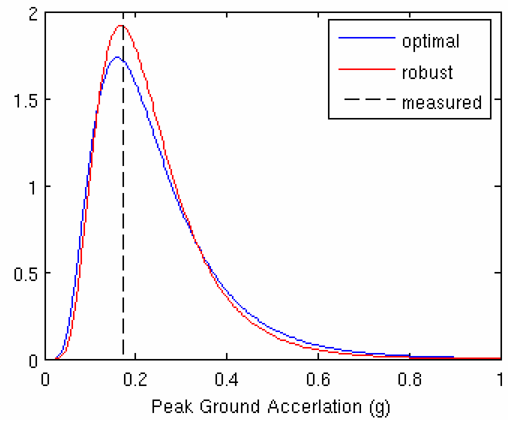
- Results using MCMC samples for robust PDF compared to results for optimal PDF computed from values reported in BJF 93
- User input U=(M, d, G)
  - Magnitude 7.0
  - Distance to fault 30km
  - Site geology is stiff soil



Posterior Predictive Analysis

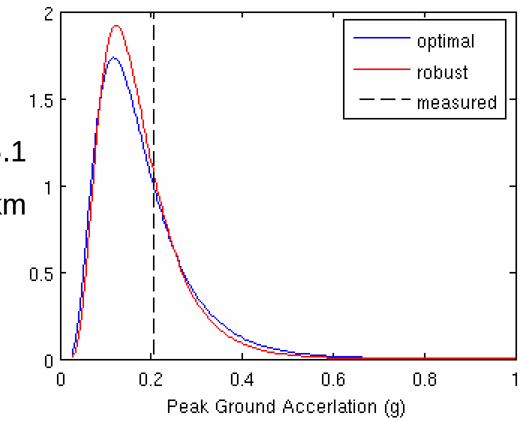
1971 San Fernando M 6.6

JPL Building 180 23.7 km



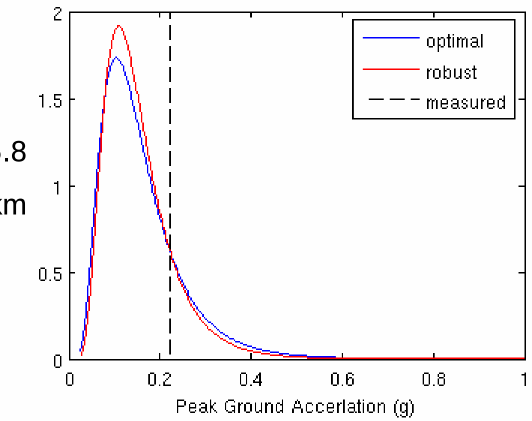
1987 Whittier Narrows M 6.1

1488 Old House Rd 19.0 km



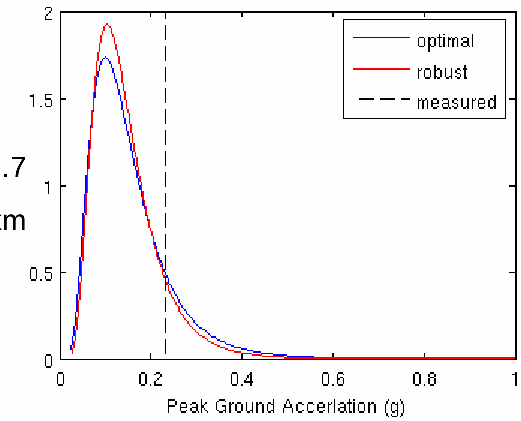
1991 Sierra Madre M 5.8

535 S Wilson Ave 18.1 km



1994 Northridge M 6.7

1150 N Sierra Madre Villa 36.2 km



Model Class Selection

- Bayesian model class selection (Beck and Yuen 2004) for set $\mathcal{M}$ of candidate models

$$P\left(\mathcal{M}_i \mid \mathcal{D}, \mathcal{M}\right) = \frac{p\left(\mathcal{D} \mid \mathcal{M}_i\right) P\left(\mathcal{M}_i \mid \mathcal{M}\right)}{\sum_{j=1}^{I} p\left(\mathcal{D} \mid \mathcal{M}_j\right) P\left(\mathcal{M}_j \mid \mathcal{M}\right)}, \quad \forall i = 1, \ldots, I$$

- The *evidence* for model class $\mathcal{M}_i$ is given by

$$p\left(\mathcal{D} \mid \mathcal{M}_i\right) = \int p\left(\mathcal{D} \mid \mathcal{M}_i, \theta_i\right) p\left(\theta_i \mid \mathcal{M}_i\right) d\theta_i$$

- Evaluating the evidence directly by stochastic simulation would require sampling from the prior, which is typically inefficient

Model Class Selection (Continued)

- The evidence may expressed as

$$\ln\left[p(\mathcal{D} \mid \mathcal{M}_i)\right]$$

$$= \int \ln\left[p(\mathcal{D} \mid \mathcal{M}_i, \theta_i)\, p(\theta_i \mid \mathcal{M}_i)\right] p(\theta_i \mid \mathcal{M}_i, \mathcal{D})\, d\theta_i + H\left[p(\theta_i \mid \mathcal{M}_i, \mathcal{D})\right]$$

- First term may be approximated from MCMC samples $\int \ln\left[p(\mathcal{D} \mid \mathcal{M}_i, \theta_i)\, p(\theta_i \mid \mathcal{M}_i)\right] p(\theta_i \mid \mathcal{M}_i, \mathcal{D})\, d\theta_i$

$$\approx \frac{1}{N} \sum_{k=1}^{N} \ln\left[p\left(\mathcal{D} \mid \mathcal{M}_i, \hat{\theta}_k\right) p\left(\hat{\theta}_k \mid \mathcal{M}_i\right)\right]$$

- Numerous methods for approximating information entropy from samples

Model Class Selection (Continued)

- New result generalizes to any model class the conclusion by Beck and Yuen (2004) made for a globally identifiable model class using an asymptotic approximation of the evidence for the model class:

$$
\begin{aligned}
&\ln\left[p\left(\mathcal{D} \mid \mathcal{M}_i\right)\right] \\
&= \int \ln\left[p\left(\mathcal{D} \mid \mathcal{M}_i, \theta_i\right) p\left(\theta_i \mid \mathcal{M}_i\right)\right] p\left(\theta_i \mid \mathcal{M}_i, \mathcal{D}\right) d\theta_i \\
&- \int \ln\left[p\left(\theta_i \mid \mathcal{M}_i, \mathcal{D}\right)\right] p\left(\theta_i \mid \mathcal{M}_i, \mathcal{D}\right) d\theta_i \\
&= \int \ln\left[p\left(\mathcal{D} \mid \mathcal{M}_i, \theta_i\right)\right] p\left(\theta_i \mid \mathcal{M}_i, \mathcal{D}\right) d\theta_i - \int \ln\left[\frac{p\left(\theta_i \mid \mathcal{M}_i, \mathcal{D}\right)}{p\left(\theta_i \mid \mathcal{M}_i\right)}\right] p\left(\theta_i \mid \mathcal{M}_i, \mathcal{D}\right) d\theta_i \\
&= \text{Data Fit - Expected Information Gained from Data}
\end{aligned}
$$

Model Class Selection

$$\log_{10}(PGA) = b_1 + b_2(M-6) + b_3(M-6)^2 + b_4 R + b_5 \log_{10}(R) + b_6 G_B + b_7 G_C + \varepsilon, \quad R = \sqrt{d^2 + h^2}$$

Model Class	b_1	b_2	b_3	b_4	b_5	b_6	b_7	h	σ	Prob. (%)
BJF '93	-0.105	0.229			-0.778	0.162	0.251	5.570	0.230	
Model 1	0.312 (0.327)	0.274 (0.034)	-0.082 (0.039)	0.003 (0.002)	-1.079 (0.252)	0.169 (0.036)	0.242 (0.038)	8.668 (2.631)	0.206	0.01
Model 2	0.276 (0.326)	0.223 (0.023)		0.003 (0.002)	-1.079 (0.254)	0.174 (0.036)	0.264 (0.037)	8.424 (2.534)	0.207	0.18
Model 3	-0.076 (0.103)	0.274 (0.034)	-0.076 (0.040)		-0.759 (0.064)	0.165 (0.036)	0.241 (0.038)	6.031 (1.510)	0.206	0.07
Model 4	-0.056 (0.108)	0.226 (0.023)			-0.803 (0.063)	0.172 (0.036)	0.262 (0.036)	6.336 (1.162)	0.207	2.12
Model 5	-0.896 (0.043)	0.222 (0.026)		-0.008 (0.001)		0.143 (0.039)	0.264 (0.041)	1.936 (1.664)	0.230	0.00
Model 6	0.255 (0.114)	0.227 (0.025)			-0.888 (0.072)			7.026 (1.720)	0.225	0.00
Model 7		0.230 (0.022)			-0.834 (0.021)	0.163 (0.033)	0.251 (0.034)	6.834 (1.041)	0.207	97.62

Comparison by Mean Data Fit Only

$$\log_{10}(PGA) = b_1 + b_2(M-6) + b_3(M-6)^2 + b_4 R + b_5 \log_{10}(R) + b_6 G_B + b_7 G_C + \varepsilon, \quad R = \sqrt{d^2 + h^2}$$

Model Class	b_1	b_2	b_3	b_4	b_5	b_6	b_7	h	σ	Data Fit (%)
BJF '93	-0.105	0.229			-0.778	0.162	0.251	5.570	0.230	
Model 1	0.312 (0.327)	0.274 (0.034)	-0.082 (0.039)	0.003 (0.002)	-1.079 (0.252)	0.169 (0.036)	0.242 (0.038)	8.668 (2.631)	0.206	40.67
Model 2	0.276 (0.326)	0.223 (0.023)		0.003 (0.002)	-1.079 (0.254)	0.174 (0.036)	0.264 (0.037)	8.424 (2.534)	0.207	6.62
Model 3	-0.076 (0.103)	0.274 (0.034)	-0.076 (0.040)		-0.759 (0.064)	0.165 (0.036)	0.241 (0.038)	6.031 (1.510)	0.206	35.61
Model 4	-0.056 (0.108)	0.226 (0.023)			-0.803 (0.063)	0.172 (0.036)	0.262 (0.036)	6.336 (1.162)	0.207	7.58
Model 5	-0.896 (0.043)	0.222 (0.026)		-0.008 (0.001)		0.143 (0.039)	0.264 (0.041)	1.936 (1.664)	0.230	0.00
Model 6	0.255 (0.114)	0.227 (0.025)			-0.888 (0.072)			7.026 (1.720)	0.225	0.00
Model 7		0.230 (0.022)			-0.834 (0.021)	0.163 (0.033)	0.251 (0.034)	6.834 (1.041)	0.207	9.52

Posterior Samples : Model Class 7