Epistemic Uncertainty in Median Ground Motions
SWUS project approach

Nicolas Kuehn\(^1\), SWUS TI Team

\(^1\)Pacific Earthquake Engineering Research Center

15 July 2014 - NGA East Workshop

Motivation

**SSHAC Guidelines**
To account for (epistemic) uncertainties in *probabilistic seismic hazard analysis*, one needs to capture the **center**, **body** and **range** of *technical defensible interpretations* (the **CBR** of the **TDI**).

- For a discrete set of GMPEs, this is difficult.
  \(\Rightarrow\) go to a continuous distribution of GMPEs
    - Assessment of models \(\Rightarrow\) Visualization tools
Epistemic Uncertainties
Ground Motions

To account for epistemic uncertainties in (median) ground motions
- Select GMPEs based on certain criteria, populate the logic tree
- give subjective weights
  - Problem: GMPEs are non-exclusive and not collectively exhaustive
  - Problem: GMPEs are not independent (NGA project)
- increase epistemic uncertainty by selecting a backbone model and scaling it up and down
  - misses epistemic uncertainty in magnitude/distance scaling

GMPE Model Space
GMPE Distribution

- Each GMPE is just a (discrete) point in modelspace
  - What about unpublished models?
  - Hypothetical additional developers would come up with different models.
- Center/body/range should be based on all models.
- Assumption: GMPEs are sampled from a distribution of models
Enhancing GMPE Model Space

GMPE distribution

- Each GMPE $f_i$ is a sample from a distribution $\mathcal{F}$:
  \[ f \sim \mathcal{F} \]

  - $\mathcal{F}$ can be thought of as a distribution over physically plausible functions.
  - In general, $\mathcal{F}$ will be something complicated.
    - Simplification: all models have the same functional form, with coefficients $\theta_i$:
      \[ \mu_i(\ln Y) = f(\theta_i; M, R, \ldots) \]

- Then the distribution over GMPEs is a distribution over coefficients $\theta$:
  \[ \theta \sim \mathcal{D} \]
  \[ \theta \sim \mathcal{N}(\theta; \mu_\theta, \Sigma_\theta) \]

Enhancing GMPE Model Space

Common Functional Form

Is a common functional form justified?

- GMPEs are generally similar
  - monotonically increasing with magnitude
  - monotonically decreasing with distance
  - restrictions on slopes with magnitudes/distances

- For some applications, only parts of the model are relevant.
GMPE Distribution

Estimation

- common functional form for median prediction
  \[ \mu(\ln Y) = f(\theta; M, R, \ldots) \]
- for each GMPE \( i \), the set of coefficients \( \theta_i \) can be estimated
  - generating “data” (median predictions) for several scenarios, fit
- From the \( \theta_i \), calculate mean \( \mu_\theta \) and covariance \( \Sigma_\theta \).
- Given \( \mu_\theta \) and \( \Sigma_\theta \), sample new sets of coefficients \( \theta \) and thus generate new models.

Enhancing GMPE Model Space

Approach

Broadening the model space based of existing GMPEs
- Take selected GMPEs
- fit to common functional form
- estimate joint parameter distribution
- generate new models (sample from joint distribution)
- evaluate new models (visualization)
- reduce model space
GMPE Distribution
Application

- Diablo Canyon Power Plant (SWUS project)
  - SSHAC Level 3 project: Hazard estimation for DCPP (California)
  - Ground-motion characterization (source characterization is different project)
  - Project output: hazard for rock ($V_{S30} = 760m/s$)
  - controlling sources for DCPP are nearby faults (Los Osos, Hosgri, Shoreline)
    - controlling sources nearby strike-slip and reverse events

For this application, only specific parts of a GMPE are relevant.

GMPE Distribution
Application

- Diablo Canyon Power Plant (SWUS project)

Approach

- selected GMPEs: ASK 14, ASB13, BSSA 14, CaBo 14, ChYo 14, Id 14, ZhLu 11, Zh+ 06
- generate synthetic “data”
  - $M_w = 5., 5.2, \ldots , 8.$
  - $R_{JB} = 1., 2., \ldots , 30., 35., \ldots , 70.$ km, $R_{rup}$ calculated from geometry footwall
  - $V_{S30} = 760m/s$
  - different values of $Z_{tor}$ and hypocentral depth for each magnitude
  - 17 response spectral periods between $T = 0.01, \ldots , 3s$
Each GMPE is fitted to the following functional form:

\[ f(M_W, R, Z_{tor}, F, T) = \]
\[ \theta_1(T) - \exp(\theta_6(T))R + \exp(\theta_9(T))Z_{tor} + \exp(\theta_{10}(T))F + \]
\[ (\theta_5(T) + \theta_6(T)(M_W - 5.))\log \sqrt{R^2 + \theta_7(T)^2} + \]
\[ \begin{cases} 
\theta_2(T)(5.5 - 6.5) + \theta_3(T)(M_W - 5.5), & M_w < 5.5 \\
\theta_2(T)(M_W - 6.5), & 5.5 \leq M_W \leq 6.5 \\
\theta_3(T)(M_W - 6.5), & 6.5 \leq M_W 
\end{cases} \]
Common Functional Form

Fit

- fit to $R_{RUP}$ and $R_{JB}$
- different $Z_{tor}$ values for same magnitude

Common Functional Form

- fit models to common functional form using both $R_{RUP}$ and $R_{JB}$
  - behave differently on the hanging wall
  - capture characteristics of both types
  - reduce misfit

- From fitted sets of coefficients, calculate $\mu_\theta$ and $\Sigma_\theta$
  $\Rightarrow$ two coefficient distributions, one for $R_{RUP}$, one for $R_{JB}$
generate “data” for each GMPE

fit common functional form to each GMPE
GMPE Distribution

Approach

- generate “data” for each GMPE
- fit common functional form to each GMPE
- sample new GMPEs (1000 new models are shown)

uncertainty model of Al-Atik and Youngs (2014)

\[ \ln Y = f(M, R, \ldots) \pm \alpha \sigma_{AY14}(T, M, F), \]

with \( \alpha = \{-2, 0, 2\} \)
GMPE Distribution
Evaluation

- In total, 15000 new models are sampled
  - 7500 from the $R_{RUP}$-distribution
  - 7500 from the $R_{JB}$-distribution
- 15000 new models too many for application
- need to evaluate with respect to center/body/range
  ⇒ visualize models in 2D (Scherbaum et al., 2010)

Visualization

- A GMPE is a function
- A function is an infinitely-dimensional vector (sort of)
  - discretize function → GMPE is a point in high-dimensional space
- Project points from HD to lower (2D, 3D) dimensions
Visualization

Example

- 1000 sampled models, evaluated at $M_W = 5, 6, 7$
  → each model is a point in 3D-ground-motion space
- models lie on lower dimensional manifold

Visualization

Method

- Go from 3D to higher dimensions ⇒ calculate model predictions at many different $M, R$ scenarios
- perform principal component analysis (PCA)
- use output from PCA as input to Sammon’s mapping
  - minimizes differences between HD and 2D Euclidean distances
  - $d_E(x, y) = \sqrt{\sum_i^N (x_i - y_i)}$
GMPE Distribution
Visualization

- Visualization of models in 2D
  - $M_W = 5., 5.5, \ldots 8.$
  - $R_{JB} = 1., 2., \ldots, 10., 15., 20., 30.$
  - $F = 0, 1$
  - $T = 0.01s$
GMPE Distribution
Visualization

- Contour plot of Euclidean Distance (HD) to mean model

![Contour plot](image)

- The map can be used to select the center, body and range of median predictions.
  - 2D visualization of ground-motion space.
  - Need to evaluate the models.

Visualization of Model Space
Evaluation

To evaluate the models, different information is available:

- GMPEs
  - All GMPEs are published, good models → use GMPE distribution as is.

- Data
  - Select data that is relevant for the application, calculate likelihood, residuals.

- Simulations
  - Do finite-fault simulations, compare with models.
Visualization of Models
Comparison with Data

- data set underlying ASK 14, $R_{JB} < 70$ km, $M_W > 5$, number of recordings per event > 3, $V_{S30} > 250$ m/s corrected to $V_{S30} = 760$ m/s, on footwall

For each of the 15000 new models
- Residuals are calculated.
  - These are split into between-event and within-event residuals, $\delta B_i$ and $\delta W_{ij}$.
- The log-likelihood is calculated (eq. (7) of Abrahamson and Youngs, 1992).
- Between and within-event variability $\tau$ and $\phi$ are fixed at the values of the BSSA 14 model.
Evaluation of Models
Residuals – Data

- Mean between residual $\approx$ mean bias
- most models are approximately centered

Evaluation of Models
Visualization – Data

- Contour plot of mean between-event residual
Evaluation of Models
Visualization – Data

- Contour plot of log-likelihood

Evaluation of Models
Simulations

- Finite-Fault Simulations (Hosgri, Shoreline)
- Different scenarios: $M_W = 5.5, 6., 6.5, 6.6, 7.2$
- Three methods: ExSIM, GP, SDSU

- Treat as data: calculate residuals to all models

J. Bayless, K. Woodell
Evaluation of Models
Visualization – Data

Contour plot of mean between-event residual

Evaluation of Models
Visualization – Data and Residuals

Contour plot of mean between-event residual
Selection of models:
- calculate ellipse fitting the convex hull of all GMPEs ± uncertainty model (Al-Atik and Youngs, 2014)
- scale ellipse (times 0.5, 1.5)
- calculate intersection points of ellipse and contours of mean between-event residual (-0.3, -0.15, 0, 0.15, 0.3)
- select $R_{RUP}$ and $R_{JB}$ models closest to these points
- evaluate
Selection of Models

Scaling

- magnitude/distance scaling of original GMPEs and selected models

Selection of Models

Spectrum

- Selection is done for each period separately.
- Different number of selected models per period.
Models represent a large ground-motion space, but need to be weighted.

- Models that are more consistent with data (smaller absolute residual, larger likelihood) should receive higher weight.
- Models that are far from other models should receive higher weight (contain more information).

Weights are calculated by calculating some mean statistic (likelihood, residual) over the area of a selected model.
Selection of Models

Weights

Selected statistics:

- \( \frac{1}{\mu(\delta B)} \): 1/mean bias
  - NGA, simulations
- likelihood: variability
  - NGA
- “prior” – value of PDF of GMPE distribution (uninformativ)

\[ w_{\text{Total}} = 0.6 \left( 0.6 \times w_{\text{Residual}_{\text{NGA}}} + 0.4 \times w_{\text{LL}_{\text{NGA}}} \right) + 0.2 \times w_{\text{Residual}_{\text{SIM}}} + 0.2 \times w_{\text{Prior}} \]
Weighted Models

**Distribution**

![Graph showing weighted models with parameters: $M_W = 7.5$, $R_x = -10$, sof = 1, $T = 0.01$.]

- **CDF**
- **Ln(HPSAL)**
- **$M_W = 7.5$**
- **$R_x = -10.$**
- **sof = 1**
- **$T = 0.01$**

- **Legend**:
  - wTotal
  - wResidual_NGA
  - wLL_NGA
  - wPrior
  - wResidual_SIM
  - GMPEs

- **Axes**: $M_W$ vs. Ln(PSA)
- **Range**: $M_W$ from 5.0 to 8.0
- **Ln(PSA)** from -5.0 to 0.0

---

**Weighted Models**

**Distribution**

![Graphs showing different distributions with parameters: $M = 6.5$, $R_x = -10$, sof = 0, $T = 0.01$.]

- **Graph 1**: $M_W$ vs. Ln(PSA) with $M = 6.5$, $R_x = -10$, sof = 0, $T = 0.01$
- **Graph 2**: $-R_x$ vs. Ln(PSA) with $M = 6.5$, sof = 0, $T = 0.01$
- **Graph 3**: Ln(PSA) vs. $T$ with $M = 6.5$, $R_x = -10$, sof = 0, $T$ range from 0.01 to 2.00
NGA East

- Calculation of PSA using different methods (empirical FAS, stochastic simulations)
  - Distribution of Models (posterior distribution?)
  - Do they overlap?
- Evaluation of models
  - Data
  - Simulations

Acknowledgements

- SWUS PPRP
  - Steven Day
  - Brian Chiou
  - Ken Campbell
  - Tom Rockwell
- Nick Gregor
- Gabriel Toro
- Linda Al-Atik
References


K. W. Campbell and Y. Bozorgnia (2014) NGA-West2 Ground Motion Model for the Average Horizontal Components of PGA, PGV, and 5%-Damped Linear Acceleration Response Spectra, *Earthquake Spectra*


References
