NGA-East Final Standard Deviation Models

WS#3C June 17, 2015 TI Team

Outline

- Summary of CENA data, results and updates on analysis
- For each of Tau, PhiSS, and PhiS2S:
 - Candidate models
 - Logic trees: weights and justification
 - Comparison to existing models
- Sigma and SigmaSS logic trees and models
- Additional Considerations:
 - Recommendation for PIE
 - Recommendation for Gulf region
 - Mixture model

CENA Data Distribution

Min of 3 recs per eqk: 1,544 recs (548 PIE) - 61 eqks (9 PIE) - 418 stations



- CENA results cannot inform models that extrapolate reliably to large M
 - CENA results are reliable for T between ~0.075 and 1.5 sec

CENA Ground Motion Model (PSA)

$$\frac{R_{rup} < 50km}{\ln(Y) = c_1 + c_2M + c_3M^2 + c_4\ln\left(\sqrt{R_{rup}^2 + c_6^2}\right) + c_7\sqrt{R_{rup}^2 + c_6^2} + c_8\ln(V_{s30}) + c_9\left(\frac{H_{dep}}{20}\right)$$

$$\frac{R_{rup} > = 50km}{\ln(Y) = c_1 + c_2M + c_3M^2 + (c_4 - c_{4h})\ln\left(\sqrt{50^2 + c_6^2}\right) + c_4\ln\left(\sqrt{R_{rup}^2 + c_6^2}\right) + c_7\sqrt{R_{rup}^2 + c_6^2} + c_8\ln(V_{s30}) + c_9\left(\frac{H_{dep}}{20}\right)$$

 $c_{\rm 4h} = Period \ dependent$, determined based on SMSIM Results

$$c_6 = \begin{cases} 1 & for \quad M \leq 4.0 \\ 3.5M - 13 & for \quad 4.0 < M \leq 5.0 \\ 4.5 & for \quad 5.0 < M \end{cases}$$

 $c_{4} = \begin{vmatrix} c_{4.R8} & for & Oklahoma - Arkansas \\ c_{4.N8} & for & Non & Oklahoma - Arkansas \end{vmatrix}$

 $c_{7} = \begin{bmatrix} c_{7.R8} & for & Oklahoma - Arkansas \\ c_{7.N8} & for & Non & Oklahoma - Arkansas \end{bmatrix}$

- Regression done for distances up to 500 km.
- Does not include Gulf region.

CENA Data Distribution (cont'd)

Min of 3 recs per eqk and station: 1,290 recs (482 PIE) – 60 eqks (9 PIE) – 231 stations



- CENA results cannot inform models that extrapolate reliably to large M
 - CENA results are reliable for T between ~0.075 and 1.5 sec





CENA Phi



CENA PhiSS



CENA PhiS2S



At 4 Hz: 39 stations with meas. V_{S30} ; 236 stations with inferred V_{S30}

9

Updates on CENA PhiS2S



Results using 3 recs/event for Tau and Phi And 3 recs/station for PhiSS and PhiS2S CENA PhiS2S is large. Potential issues investigated:

- Regression approach
- Regional impacts and potential trade-offs between Tau and PhiS2S
- PIE/tectonic events

Regression Approach

- Two-regression approach:
 - ME regression to estimate δB_{e} and δW_{es} by Justin
 - Another ME regession to estimate $\delta S2S_{s}$ and δWS_{es}

- Other approaches investigated for solving for event terms and site terms simultaneously by Justin:
 - Iterative approach
 - Bayesian regression in STAN

Regression Approach (cont'd)

- Iterative approach:
 - 1. Event terms and model coefficients are solved for.
 - 2. Site terms are computed from the within-event residuals.
 - 3. Event terms and model coefficients are recomputed from data that has the site term from the previous step removed.
 - 4. Site terms are recomputed.
 - 5. Steps 3 and 4 are repeated until the change in the likelihood of the fit falls below some threshold.
- Bayesian regression:
 - Estimates event terms and site terms simultaneously

Regression Approach (cont'd)



Iterative Approach: Nmin 3 for events and Nmin 1 for sites and fixed coefficients

Regional Effects and Trade-offs

Event Terms at F = 4 Hz

Site Terms at F = 4 Hz



Regional Effects and Trade-offs (cont'd)

At F = 4 Hz:



Area 1: 692 recordings, 35 earthquakes, 153 stations Area 2: 246 recordings, 20 earthquakes, and 58 stations

Regional Effects and Trade-offs (cont'd)



O All Tectonic Eqks O Area 1 X Area 2



PhiS2S: PIE/Tectonic Events



At F = 4 Hz, 275 stations using all eqks, 144 stations for tectonic only, 141 stations for PIE only

PhiS2S Conclusions

- Large CENA PhiS2S values do not appear to be the result of errors in V_{S30} estimates; regression method; trade-offs between Tau and PhiS2S
- PhiS2S computed from PIE events are smaller than those computed from tectonic events; likely due to sampling a smaller region
- CENA PhiS2S values have similar characteristics to PhiS2S for Japan

Issues for Model Building

- CENA dataset cannot inform Tau and PhiSS models that extrapolate reliably to large M
 - M-dependence in Tau and PhiSS observed in datasets that cover a wide M range
- CENA dataset is limited to the period range of 0.1 to 1 sec
 - Cannot inform models for Tau, PhiSS, and PhiS2S outside this period range

Approach

- Build PhiS2S model based on CENA data (tectonic + PIE)
 - Extrapolation outside of 0.1 to 1.0 sec is informed by Japanese scaling of PhiS2S
- Tau and PhiSS Models:
 - Adopt global models based on average NGA-W2 PhiSS and Tau
 - Evaluate CENA-based models (using tectonic data with $M \ge 3.0$):
 - Constant and M-dependent
 - Extrapolation outside of 0.1 to 1.0 sec
- SigmaSS and ergodic sigma models combining the components of aleatory variability

Tau

Elements of Tau Logic Tree

- Three Candidate Models:
 - Global model based on average Tau for 4 NGA-W2 models (M-dependent; period-independent)
 - Models based on CENA data:
 - Constant (homoscedastic) and period-independent
 - M-dependent and period-independent with ratios at M5.5 to M5.0 and M6.5 to M5.0 controlled by global model
- Statistical uncertainty

Statistical Uncertainty

- Sample variance of a normal distribution follows a scaled chi-square distribution (ex, Tau², PhiSS², and PhiS2S²)
 - Can be represented by 3 discrete values at 5th, 50th, and 95th percentile with weights of 0.185, 0.63, and 0.185 (Keefer and Bodily, 1983)
 - Models are characterized by mean and SD[Var]

$$\tau_{Central} = \sqrt{c\chi_{2,k}^{-1}(0.5)} \qquad \tau_{High} = \sqrt{c\chi_{2,k}^{-1}(0.95)} \qquad \tau_{Low} = \sqrt{c\chi_{2,k}^{-1}(0.05)}$$
$$c = \frac{\left(SD(\tau^2)\right)^2}{2\tau^2} \qquad k = \frac{2\tau^4}{\left(SD(\tau^2)\right)^2}$$

Global Tau Model



Average of 4 NGA-W2 Models M breaks: 4.5, 5.0, 5.5 and 6.5

Global Tau Model (cont'd)



 $\tau = \begin{cases} \tau_1 & \text{for } M \le 4.5 \\ \tau_1 + (\tau_2 - \tau_1) * \frac{(M - 4.5)}{0.5} & \text{for } M \le 5.0 \\ \tau_2 + (\tau_3 - \tau_2) * \frac{(M - 5.0)}{0.5} & \text{for } M \le 5.5 \\ \tau_3 + (\tau_4 - \tau_3) * \frac{(M - 5.5)}{1.0} & \text{for } M \le 6.5 \\ \tau_4 & \text{for } M > 6.5 \end{cases}$

Period-Independent Model

Uncertainty in Global Tau Model



Uncertainty in Global Tau Model (cont'd)



CENA Constant Tau Model

0.12

0.1

0.08

0.06

0.04

0.02

0

0.1

SD[Tau²]





SD[Tau²] replaced by SD[Tau²] for the global model at M5.0

Frequency (Hz)

1

Statistical Uncertainty

100

10

Smoothed

CENA M-Dependent Tau Model

$$\tau = \begin{cases} \tau_1 & \text{for } M \le 5.0 \\ \tau_1 + \frac{(\tau_2 - \tau_1)}{0.5} (M - 5.0) & \text{for } M < 5.5 \\ \tau_2 + (\tau_3 - \tau_2) (M - 5.5) & \text{for } M < 6.5 \\ \tau_3 & \text{for } M \ge 6.5 \end{cases}$$

<u>where:</u> $\tau_2 = \tau_1^*$ Ratio2 and Ratio2 is obtained from the global model

 $\tau_3 = \tau_1^*$ Ratio3 and Ratio3 is obtained from the global model

- Solved for τ_1 using mle
- Subset of data with M>= 3 and no PIE

CENA M-Dependent Tau Model (cont'd)





SD[Tau²] replaced by SD[Tau²] for the global model at M5.0, 5.5, and 6.5

Tau Models



Evaluation of Tau Models

- TI Team strongly favored global model over the 2 CENA models:
 - Global model built using a large uniformly-processed global dataset with the CENA dataset is significantly smaller
 - CENA dataset is limited to M<5.5 and does not extrapolate reliably to large M
 - CENA dataset is limited to F range od 1 to 10 Hz

F-Test of Equality of Variances (Tau²)

- Test against null hypothesis of equal variance of event terms
 - $H_0: Tau_{CENA}^2 = Tau_{NGAW2}^2$
 - H_a: Tau_{CENA}² ≠ Tau_{NGAW2}²
- Test statistic: $F = Tau_{CENA}^2/Tau_{NGAW2}^2$
 - Under the null hypothesis: F has a F-distribution with numerator dof of N1 – 1 and denominator dof of N2 – 1, where N1 and N2 are the sample sizes of the two data sets
 - F, N1, and N2 are needed for the test
- Calculate p-value and compare to significance level (0.05)
- If p-value > 0.05, we fail to reject H_0

F-Test of Equality of Variances (Tau²)

- CENA:
 - Tau: values obtained from ME regression with M-dependent Tau
 - N: number of events used in the ME regression
- NGA-W2 Models:
 - Option1:
 - Tau: based on event terms with M between 3.0 and 5.0. *Underestimates the models proposed Tau at M5.0*
 - N: number of events with M between 3.0 and 5.0 for each model
 - <u>Option2:</u>
 - Tau: average of proposed Tau for each model for M between 3.0 and 5.0
 - N: number of events with M between 3.0 and 5.0 for each model

Equality of Tau²

 Tau for NGA-W2 models calculated based on event terms with M between 3.0 and 5.0

	T0.10	T0.15	T0.20	T0.25	T0.30	T0.40	T0.50	T0.75	T1.00	PGV
	F10.00	F6.67	F5.00	F4.00	F3.33	F2.50	F2.00	F1.33	F1.00	
ASK14	0.012	0.520	0.882	0.701	0.332	0.503	0.998	0.259	0.141	0.954
CB14	0.180	0.352	0.054	0.075	0.031	0.089	0.299	0.759	0.379	0.043
CY14	0.505	0.899	0.690	0.467	0.636	0.662	0.653	0.596	0.734	NA
BSSA14	0.000	0.077	0.365	0.687	0.776	0.763	0.777	0.089	0.022	0.894

 Tau for NGA-W2 models is average of proposed Tau for M between 3.0 to 5.0

	T0.10	T0.15	T0.20	T0.25	T0.30	T0.40	T0.50	T0.75	T1.00	PGV
	F10.00	F6.67	F5.00	F4.00	F3.33	F2.50	F2.00	F1.33	F1.00	
ASK14	0.016	0.072	0.046	0.037	0.075	0.097	0.106	0.106	0.186	0.489
CB14	0.051	0.893	0.497	0.546	0.316	0.388	0.702	0.461	0.260	0.366
CY14	0.111	0.292	0.183	0.138	0.230	0.257	0.254	0.223	0.332	0.372
BSSA14	0.115	0.675	0.529	0.627	0.516	0.647	0.888	0.281	0.144	0.351

Tau Logic Tree


Comparison to SWUS Model



Comparison to Hanford Model



38



Elements of PhiSS Logic Tree

- Three Candidate Models:
 - Global model based on average PhiSS for 4 NGA-W2 models (M-dependent; period-dependent)
 - Models based on CENA data:
 - Constant (homoscedastic) and period-independent
 - M-dependent with ratios at M6.5 to M5.0 controlled by global model
- Statistical uncertainty: 5th, 50th, and 95th percentile of scaled chi-square distribution with wgts of 0.185, 0.63, 0.185

Global PhiSS Model



M breaks: 5.0 and 6.5 M-independent PhiSS for T >= 2 sec

Global PhiSS Model (cont'd)



$$\Phi_{SS} = \begin{cases} a & for \ M \le 5.0 \\ a + (b - a) * \frac{(M - 5.0)}{1.5} & for \ M \le 6.5 \\ b & for \ M > 6.5 \end{cases}$$

Uncertainty in PhiSS Models

- Measures the station-to-station variability in PhiSS
 - Due to azimuthal dependency, topographic effects and other unknown factors
 - Estimated using SD[PhiSS,S] for stations with a lot of recordings to correct for sampling error (<u>NGA-W2 data</u>)
- Add error in model fit to data and statistical uncertainty in PhiSS estimates

Station-to-Station Variability (NGA-W2)

- Simulating 3000 recordings per station for 147 stations such that the COV(PhiSS,S) = 0, 0.05, 0.10, 0.15 assuming normal distribution.
- Sampling N recordings per station and calculating the COV(PhiSS,S) for each N



Uncertainty in Global PhiSS Model



CENA Constant PhiSS Model



CENA M-Dependent PhiSS Model



CENA M-Dependent PhiSS Model (cont'd)



PhiSS Models



49

PhiSS Models (cont'd)



Evaluation of PhiSS Models

- TI Team strongly favored global model over the 2 CENA models:
 - Global model built using a large uniformly-processed global dataset with the CENA dataset (M>3.0, R_{rup}<300 km, and tectonic events only) is significantly smaller
 - CENA dataset is limited to M<5.5 and does not extrapolate reliably to large M
 - CENA dataset is limited to F range od 1 to 10 Hz

Evaluation of PhiSS Models (cont'd)

- F-Test: Test against null hypothesis of equal variance of CENA PhiSS² and NGA-W2 PhiSS² for M between 3.0 and 5.0
 - H_0 : PhiSS_{CENA}² = PhiSS_{NGAW2}²
 - $H_a: PhiSS_{CENA}^2 \neq PhiSS_{NGAW2}^2$

	T0.10	T0.15	T0.20	T0.25	T0.30	T0.40	T0.50	T0.75	T1.00	PGV
	F10.00	F6.67	F5.00	F4.00	F3.33	F2.50	F2.00	F1.33	F1.00	
ASK14	0.004	0.111	0.129	0.999	0.976	0.098	0.745	0.061	0.008	0.002
CB14	0.008	0.125	0.132	0.940	0.789	0.032	0.225	0.000	0.000	0.005
CY14	0.011	0.185	0.294	0.751	0.642	0.034	0.455	0.006	0.000	NA
BSSA14	0.002	0.074	0.125	0.966	0.904	0.096	0.769	0.038	0.002	0.000

Logic Tree for PhiSS



Comparison to Hanford Model



Comparison to SWUS Model





PhiS2S Logic Tree

- Model based on CENA data (PIE + Tectonic events)
 - Extrapolated for F outside of 1 to 10 Hz using the scaling of PhiS2S versus F for Japanese data
- PhiS2S models from other regions not adopted because the variability in site terms is not constant across regions



CENA PhiS2S Model



PhiS2S Model



Single-Station Sigma

Single-Station Sigma

• Combining 3 PhiSS models and 1 Tau model:

 $\sigma_{SS}^2 = \phi_{SS}^2 + \tau^2 \qquad SD[\sigma_{SS}^2] = \sqrt{(SD[\phi_{SS}^2])^2 + (SD[\tau^2])^2}$

- Three SigSS models all M-dependent with 4 M breaks at 4.5, 5.0, 5.5, and 6.5:
 - Global Model: Global PhiSS and global Tau
 - CENA Model-1: CENA constant PhiSS and global Tau
 - CENA Model-2: CENA M-dependent PhiSS and global Tau

Single-Station Sigma Logic Tree



Single-Station Sigma Models



63

Composite Single-Station Sigma Model

- Develop CDFs for the 3 continuous SigSS models at the M breaks (SigSS² follows scaled Chi-square distribution)
- Develop weighted composite CDF using the CDFs of the 3 models and their weights
- Represent the composite distribution by 3 discrete points at 5th, 50th, and 95th percentiles
- Weights of 0.185, 0.63, and 0.185 on the 3 branches produce the mean and SD of the continuous composite distribution

Composite SigSS Model (cont'd)



Comparison to Hanford Model



Total Ergodic Sigma

Total Ergodic Sigma

• Combining 3 PhiSS models, 1 PhiS2S model and 1 Tau model:

 $\sigma^{2} = \phi_{SS}^{2} + \phi_{S2S}^{2} + \tau^{2} \qquad SD[\sigma^{2}] = \sqrt{(SD[\phi_{SS}^{2}])^{2} + (SD[\phi_{S2S}^{2}])^{2} + (SD[\tau^{2}])^{2}}$

- Three Sigma models all M-dependent with 4 M breaks at 4.5, 5.0, 5.5, and 6.5:
 - Global Model: Global PhiSS, CENA PhiS2S, and global Tau
 - CENA Model-1: CENA constant PhiSS, CENA PhiS2S, and global Tau
 - CENA Model-2: CENA M-dependent PhiSS, CENA PhiS2S, and global Tau

Total Ergodic Sigma Logic Tree



Total Ergodic Sigma Models



70

Composite Ergodic Sigma Model



Comparison to NGA-W2 Models


Comparison to EPRI 2013



73

Application to PIE

- CENA PhiS2S model developed using tectonic events and PIE
- CENA Tau and PhiSS models using tectonic events only:
 - Applicability of these models to PIE in CENA
 - F-Test of equality of PhiSS² for tectonic and PIE events
 - F-Test of equality of Tau² for tectonic and PIE events

Application to PIE (cont'd)

	T0.100	T0.150	T0.200	T0.250	T0.300	T0.400	T0.500	T0.750	T1.00
	F10.00	F6.67	F5.00	F4.00	F3.33	F2.50	F2.00	F1.33	F1.00
Tectonic vs PIE	0.72	0.73	0.63	0.80	0.85	0.87	0.56	0.48	0.28



All
Tectonic Only
PIE Only

All
Tectonic Only
PIE Only

	T0.100	T0.150	T0.200	T0.250	T0.300	T0.400	T0.500	T0.750	T1.00
	F10.00	F6.67	F5.00	F4.00	F3.33	F2.50	F2.00	F1.33	F1.00
Tectonic vs PIE	0.54	0.71	0.94	0.61	0.62	0.83	0.64	0.61	0.06

Application to Gulf Region

- Analyze the standard deviations of the residuals of the Gulf data with respect to 2 PEER-developed models for the Gulf
 - Gulf data: 8 tectonic events and 1 PIE with M 3.4 to 4.7
 - Compare components of variability for the Gulf region to CENA models
 - Plots
 - F-Tests of equality of variance between CENA and Gulf

Application to Gulf (cont'd)



77

Application to Gulf (cont'd)

P-values from the F-test of equality of Tau² for tectonic events in the Gulf region and CENA

	T0.10	T0.15	T0.20	T0.25	T0.30	T0.40	T0.50	T0.75	T1.00
	F10.00	F6.67	F5.00	F4.00	F3.33	F2.50	F2.00	F1.33	F1.00
CENA vs GULF_Model1	0.54	0.26	0.08	0.00	0.01	0.58	0.63	0.79	0.89
CENA vs GULF_Model2	0.48	0.61	0.63	0.09	0.25	0.58	0.88	0.92	0.77

P-values from the F-test of equality of PhiSS² for tectonic events in the Gulf region and CENA

	T0.10	T0.15	T0.20	T0.25	T0.30	T0.40	T0.50	T0.75	T1.00
	F10.00	F6.67	F5.00	F4.00	F3.33	F2.50	F2.00	F1.33	F1.00
CENA vs GULF_Model1	0.25	0.47	0.30	0.10	0.56	0.17	0.13	0.13	0.13
CENA vs GULF_Model2	0.43	0.26	0.21	0.06	0.48	0.12	0.09	0.05	0.09

P-values from the F-test of equality of PhiS2S² for the Gulf region and CENA

	T0.10	T0.15	T0.20	T0.25	T0.30	T0.40	T0.50	T0.75	T1.00
	F10.00	F6.67	F5.00	F4.00	F3.33	F2.50	F2.00	F1.33	F1.00
CENA vs GULF_Model1	0.00	0.00	0.09	0.51	0.28	0.21	0.35	0.38	0.16
CENA vs GULF_Model2	0.00	0.01	0.14	0.68	0.40	0.28	0.60	0.97	0.41

Mixture Model

- SWUS project observed using NGA-W2 withinevent residuals that GM deviate from lognormal distribution at upper tails
- Two alternatives for representing the shape of GM distribution:
 - Traditional lognormal with wgt of 0.2
 - Mixture of 2 lognormal distributions with wgt of 0.8
 - Equally weighted mixture of 0.8*normal SigSS and 1.2*normal SigSS

Single-Station Sigma Logic Tree



Ergodic Sigma Logic Tree



Thank you!