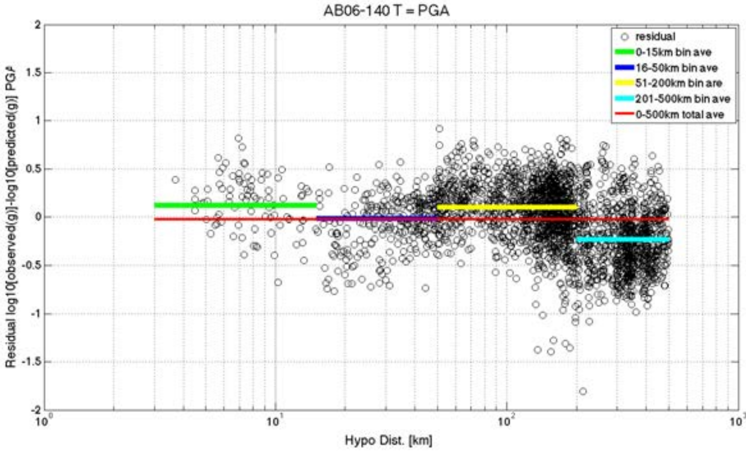
USGS
science for a changing world

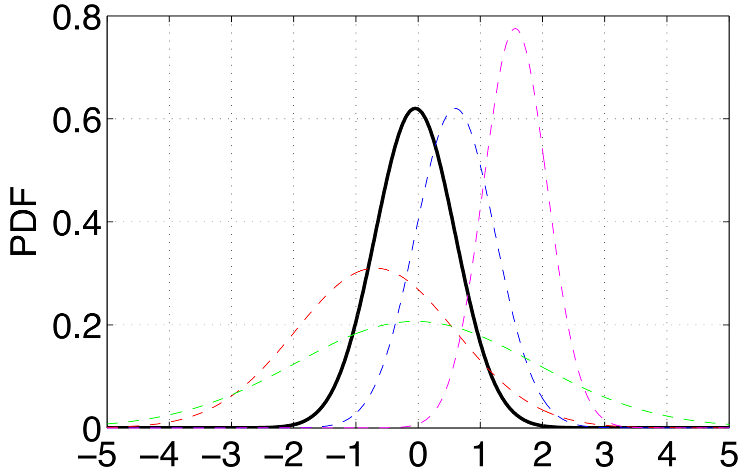
CEUS GMM Residual Study

Presented by Dan McNamara
USGS, Golden, CO
USGS 2018 NSHM Update Workshop
Wednesday, March 7th, 2018
RMS Headquarters, Newark, CA

U.S. Department of the Interior
U.S. Geological Survey

CEUS GMM Residual Study





Objective:

Evaluate CEUS GMMs for the 2018 hazard model.

GM Data

- NGAE DB
- OK, KS induced DB
- simulations M6-8

GMMs

- 2014 CEUS + A15 GMMs
- NGAW2 GMMs
- NGAE-seed GMMs
- NGAE-usgs GMMs
- logic tree weighted GMMs

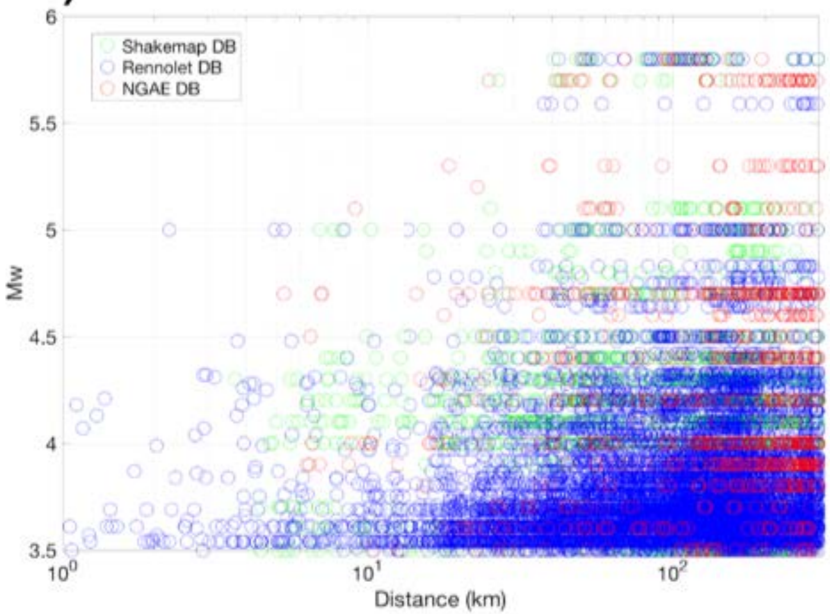
Methods

- LLH – Scherbaum et al., (2009)
- multivariate LLH – Mak and Schorlemmer (2017)

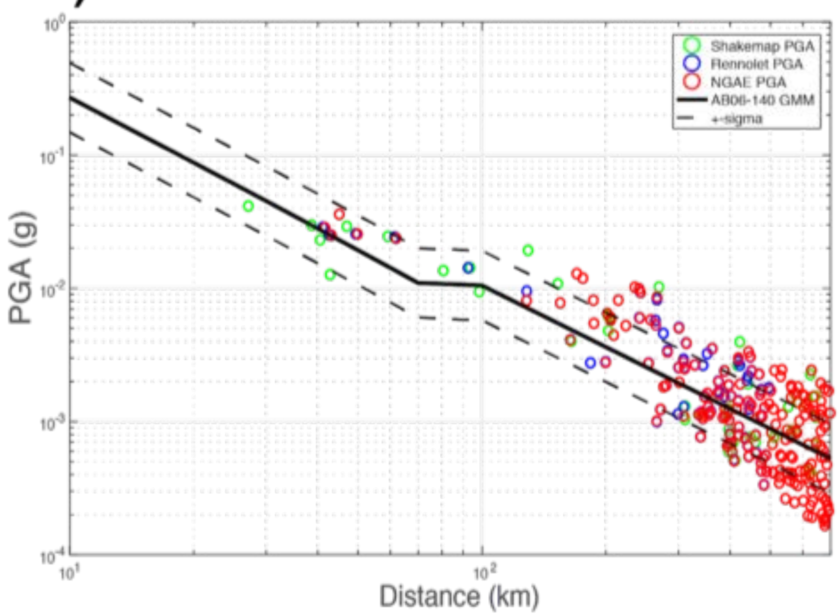
Observations

GM Data

a)



b)



PGA, PSA1s, PSA0.2s, M4-5.8

OK, KS induced. Renolet et al., (2017)
NGAE tectonic. Goulet et al., (2015)
Shakemap for more recent EQs

Also consider SM simulations M6-8

Sun and S. Rezaeian
A. Frankel
S. Hartzell

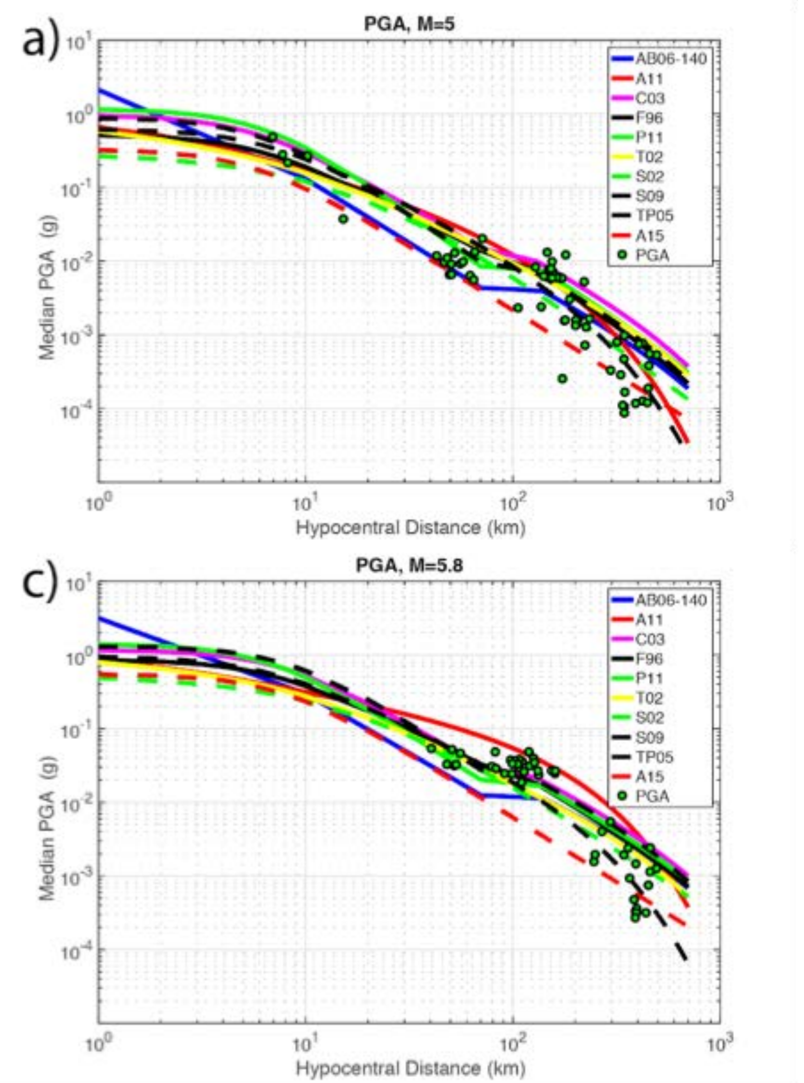
GMMs Evaluated

GMMs recently implemented in nshmp-haz software system by Peter Powers.

- CEUS GMMs used in 2014, 2016 (**A15**) hazard models
- NGAW2
- NGAE seed and updates (**SP16, G16, G17**)
- NGAE usgs
- Weighted mean GMMs
  - 2014, 2016, 2017, 2018 hazard models

Site amplification:
NGAE site amplifications applied.

- Vs30 - Topo. Slope proxy used
- measured Vs30 where available

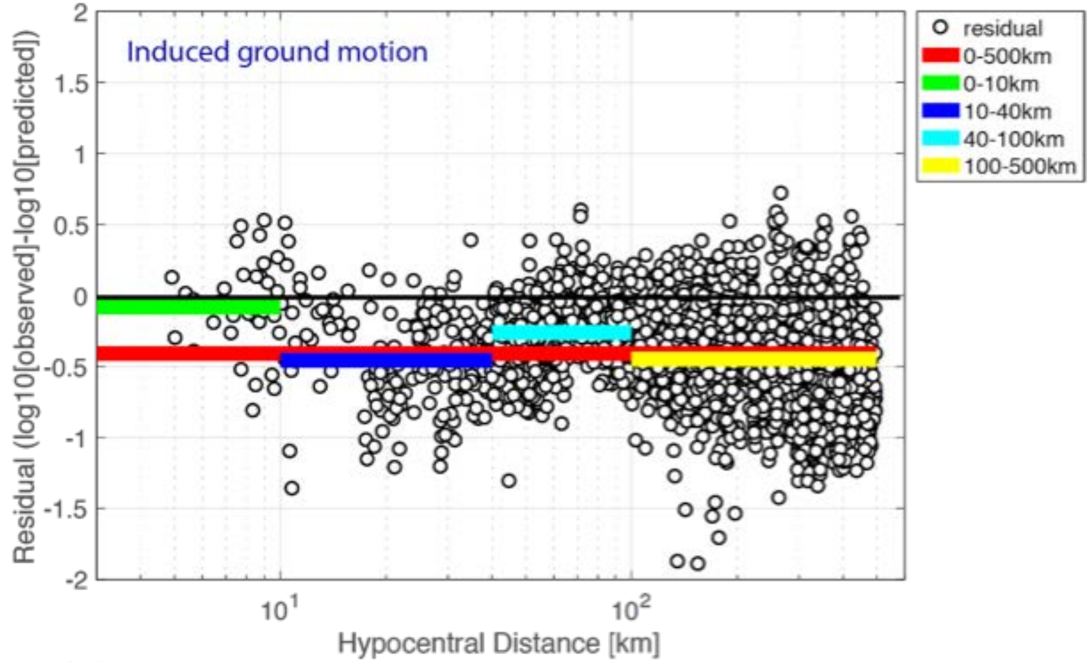


Step 1: compute residuals

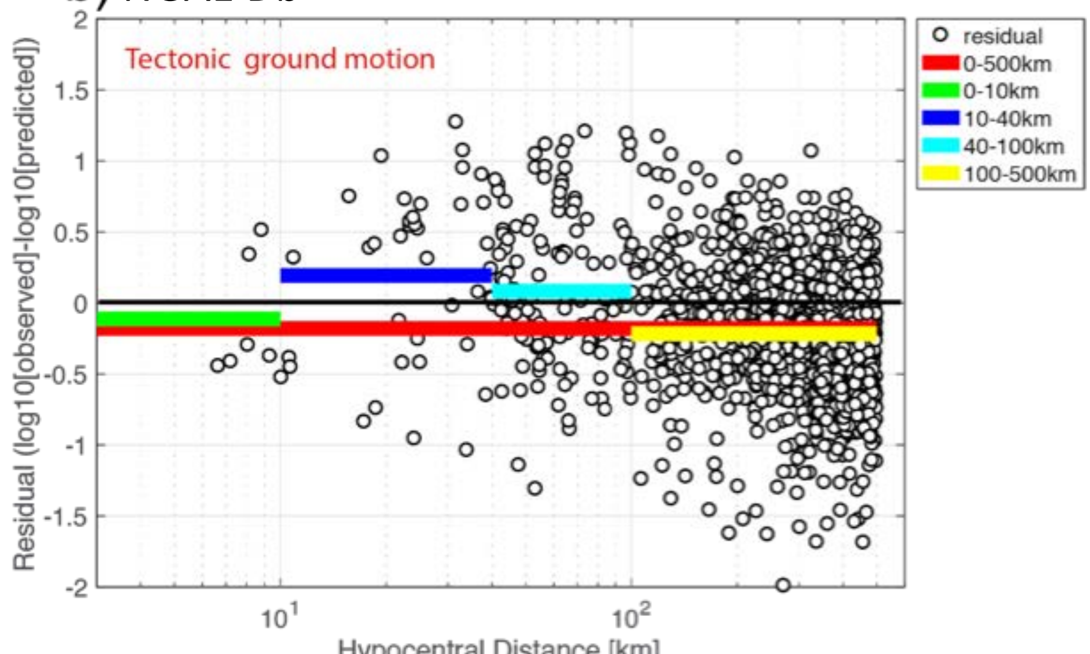
Ex. using AB06-140 GMM

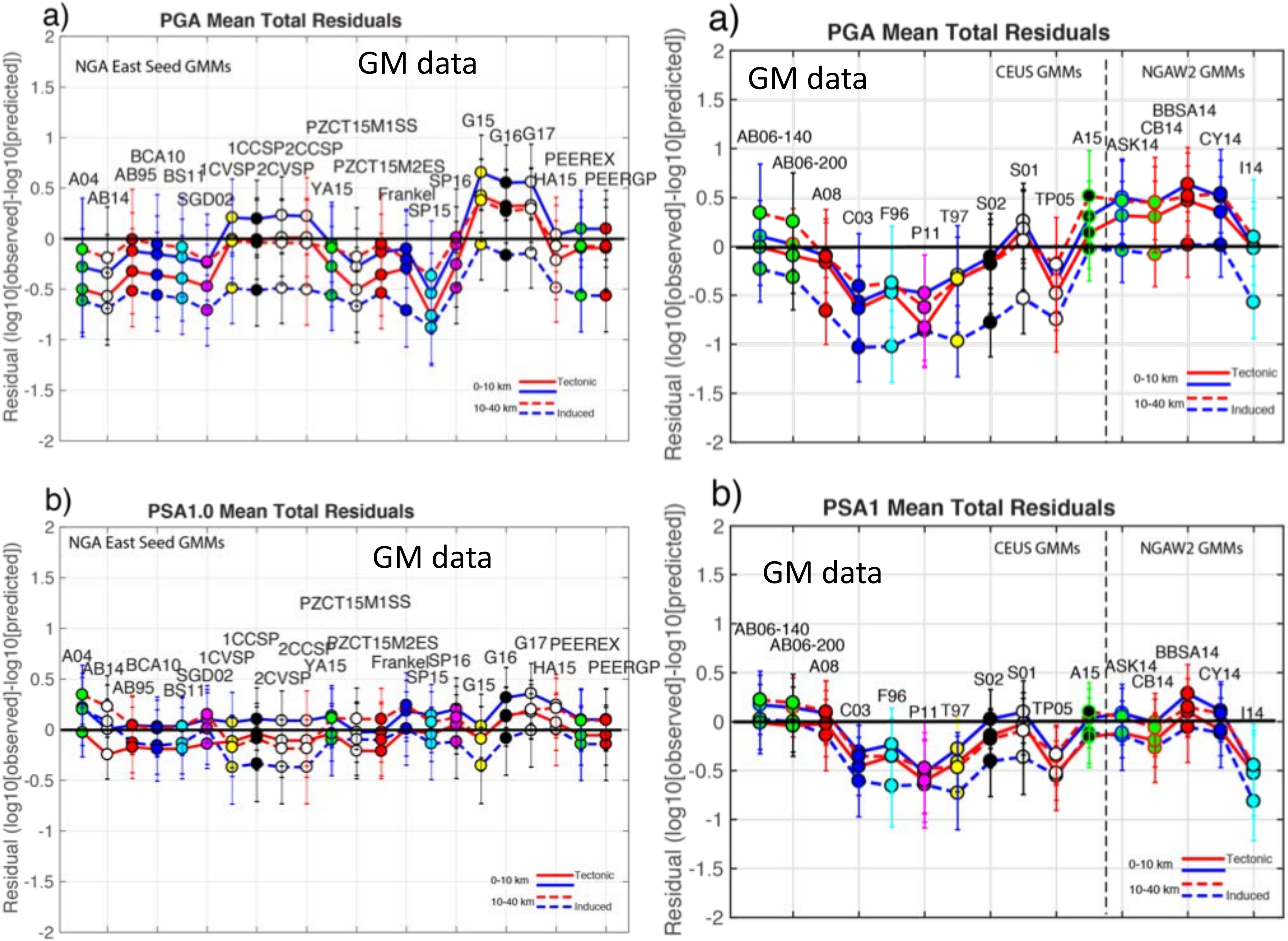
Focus on 0-40 km

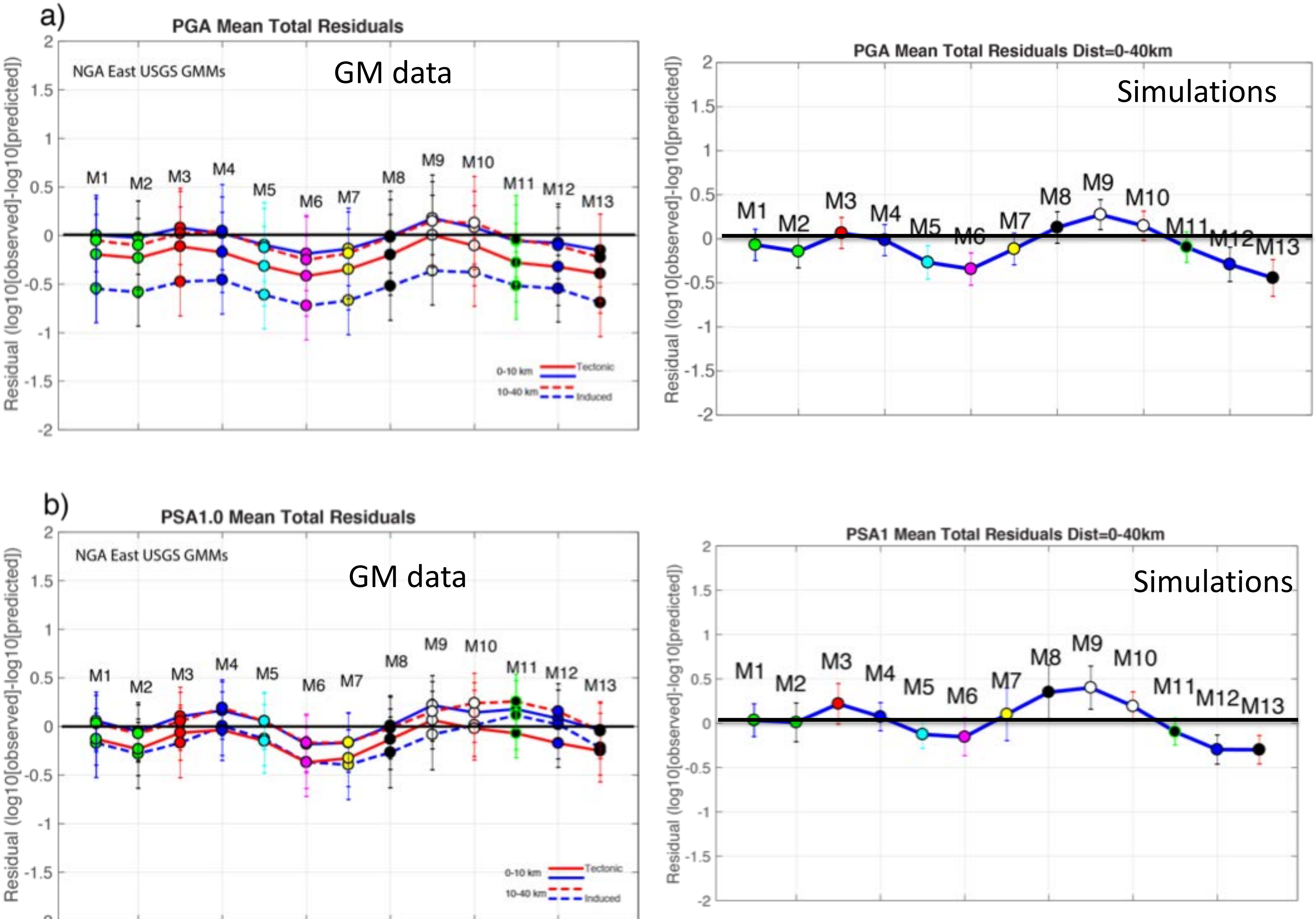
a) Renolet Db

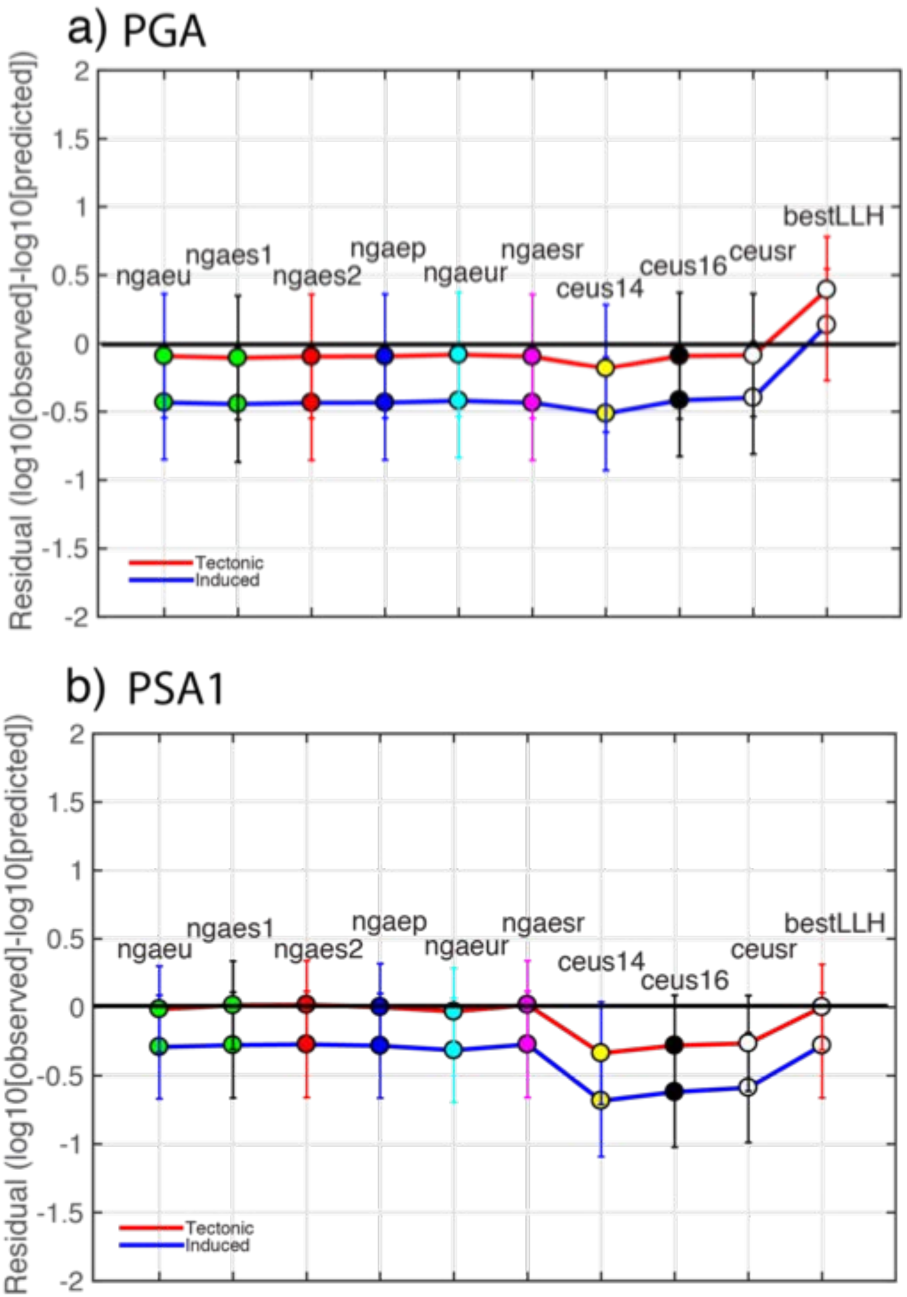


b) NGAE Db




a)
PGA Mean Total Residuals
NGA East Seed GMMs
GM data
A04 AB14 AB95 BCA10 BS11 SGD02 1CVSP 1CCSP 2CCSP 2CVSP YA15 PZCT15M1SS PZCT15M2ES Frankel SP15 SP16 G15 G16 G17 HA15 PEEREX PEERGP
Residual (log10[observed]-log10[predicted])
0-10 km
10-40 km
Tectonic
Induced
a)
PGA Mean Total Residuals
GM data
CEUS GMMs
NGAW2 GMMs
AB06-140 AB06-200 A08 C03 F96 P11 T97 S02 S01 TP05 A15 ASK14 CB14 BBSA14 CY14 I14
Residual (log10[observed]-log10[predicted])
0-10 km
10-40 km
Tectonic
Induced
b)
PSA1.0 Mean Total Residuals
NGA East Seed GMMs
GM data
A04 AB14 AB95 BCA10 BS11 SGD02 1CVSP 1CCSP 2CCSP 2CVSP YA15 PZCT15M1SS PZCT15M2ES Frankel SP15 SP16 G15 G16 G17 HA15 PEEREX PEERGP
Residual (log10[observed]-log10[predicted])
0-10 km
10-40 km
Tectonic
Induced
b)
PSA1 Mean Total Residuals
GM data
CEUS GMMs
NGAW2 GMMs
AB06-140 AB06-200 A08 C03 F96 P11 T97 S02 S01 TP05 A15 ASK14 CB14 BBSA14 CY14 I14
Residual (log10[observed]-log10[predicted])
0-10 km
10-40 km
Tectonic
Induced


a)
PGA Mean Total Residuals
NGA East USGS GMMs
GM data
Residual (log10[observed]-log10[predicted])
M1
M2
M3
M4
M5
M6
M7
M8
M9
M10
M11
M12
M13
0-10 km
10-40 km
Tectonic
Induced
PGA Mean Total Residuals Dist=0-40km
Simulations
b)
PSA1.0 Mean Total Residuals
NGA East USGS GMMs
GM data
PSA1 Mean Total Residuals Dist=0-40km
Simulations



Weighted GMMs

(ngaeu) 13 NGA-East USGS GMMs Sammon's map weighted GMM (**Goulet et al., 2017**),

(ngaes1) 19 NGA-East SEED GMMs

(ngaes2) 19 NGA-East SEED GMMs

(ngaep) - ngaes1 and ngaes2 equally weighted (**Petersen et al., 2018**),

(ngaeur) NGA-East USGS GMMs weighted based on LLH results from this study,

(ngaesr) NGA-East SEED GMMs weighted based on results from this study,

(ceus14) CEUS GMMs

(ceus16) CEUS GMMs + A15

(ceusr) CEUS GMMs using weights determined in this study

(bestLLH) 3 best LLH results for PGA GMMs using the induced earthquake ground motions (**A15**,**ASK14**,**G17**) equally weighted.

LLH Method

The negative average log-likelihood LLH reflects the fit between the data and model:

$$\mathrm{LLH} = -\frac{1}{N}\sum_{i=1}^{N}\log_2(g(x_i)),$$

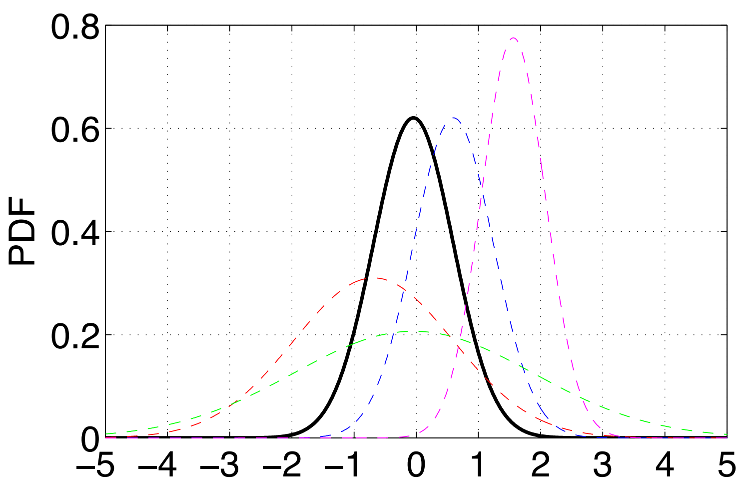
N = the number of observations xi , and g the probability
density function (PDF) predicted by the GMPE (normal distribution)

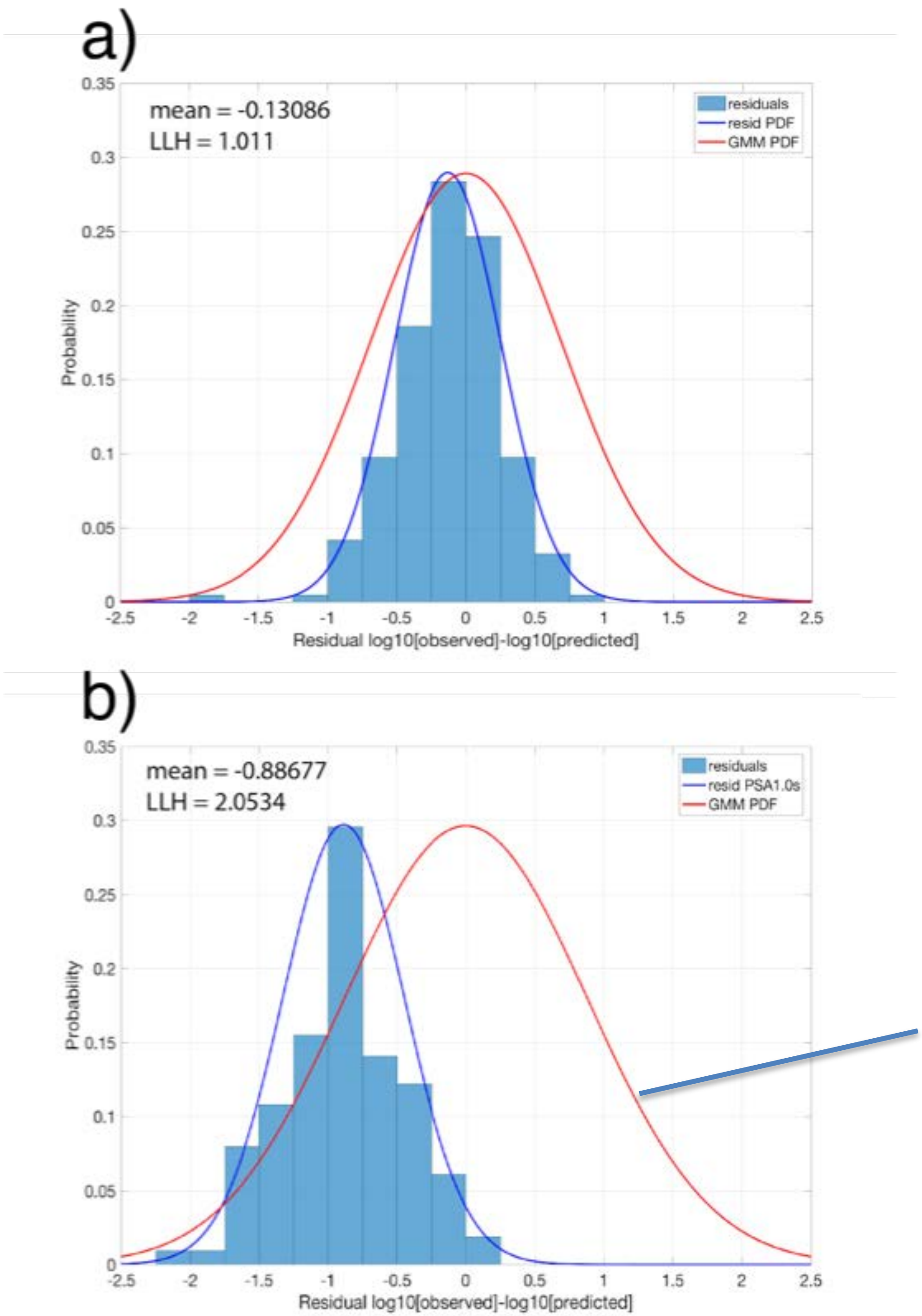
Bulletin of the Seismological Society of America, Vol. 99, No. 6, pp. 3234–3247, December 2009, doi: 10.1785/0120080347

Model Selection in Seismic Hazard Analysis: An Information-Theoretic Perspective

by Frank Scherbaum, Elise Delavaud, and Carsten Riggelsen

Abstract Although the methodological framework of probabilistic seismic hazard analysis is well established, the selection of models to predict the ground motion at the sites of interest remains a major challenge. Information theory provides a powerful theoretical framework that can guide this selection process in a consistent way. From an information-theoretic perspective, the appropriateness of models can be expressed in terms of their relative information loss (Kullback–Leibler distance) and hence in physically meaningful units (bits). In contrast to hypothesis testing, information-theoretic model selection does not require *ad hoc* decisions regarding significance levels nor does it require the models to be mutually exclusive and collectively exhaustive. The key ingredient, the Kullback–Leibler distance, can be estimated from the statistical expectation of log-likelihoods of observations for the models under consideration. In the present study, data-driven ground-motion model selection based on Kullback–Leibler-distance differences is illustrated for a set of simulated observations of response spectra and macroseismic intensities. Information theory allows for a unified treatment of both quantities. The application of Kullback–Leibler-distance based model selection to real data using the model generating data set for the Abrahamson and Silva (1997) ground-motion model demonstrates the superior performance of the information-theoretic perspective in comparison to earlier attempts at data-driven model selection (e.g., Scherbaum *et al.*, 2004).





LLH score examples

Lower LLH score indicates better fit

LLH penalizes large sigma

best LLH scores when GMM sigma
Similar to data sigma.

Defined by published GMM
Mean and sigma

MLLH Method

Expands LLH to take advantage of inter (within) and intra (between) event residual uncertainty (phi, tau) such that

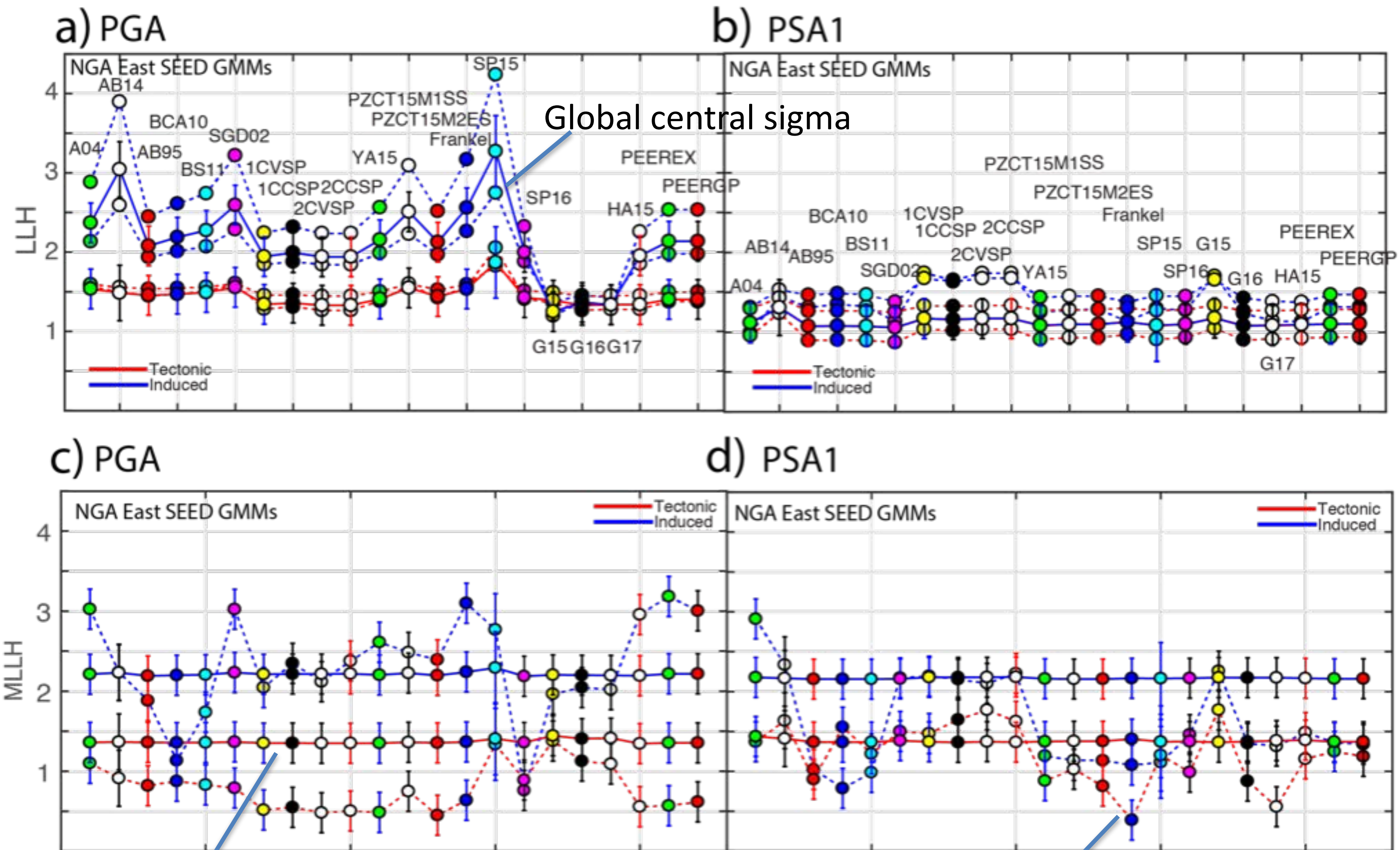
total Sigma = sqrt (phi^2 + tau^2)

Not all GMMs in the study have published phi and tau
NGAE models have uniform phi and tau

used mixed effects regression (LME) to determine GMM phi and taus for MLLH scoring

Mak and Schorlemmer (2017)

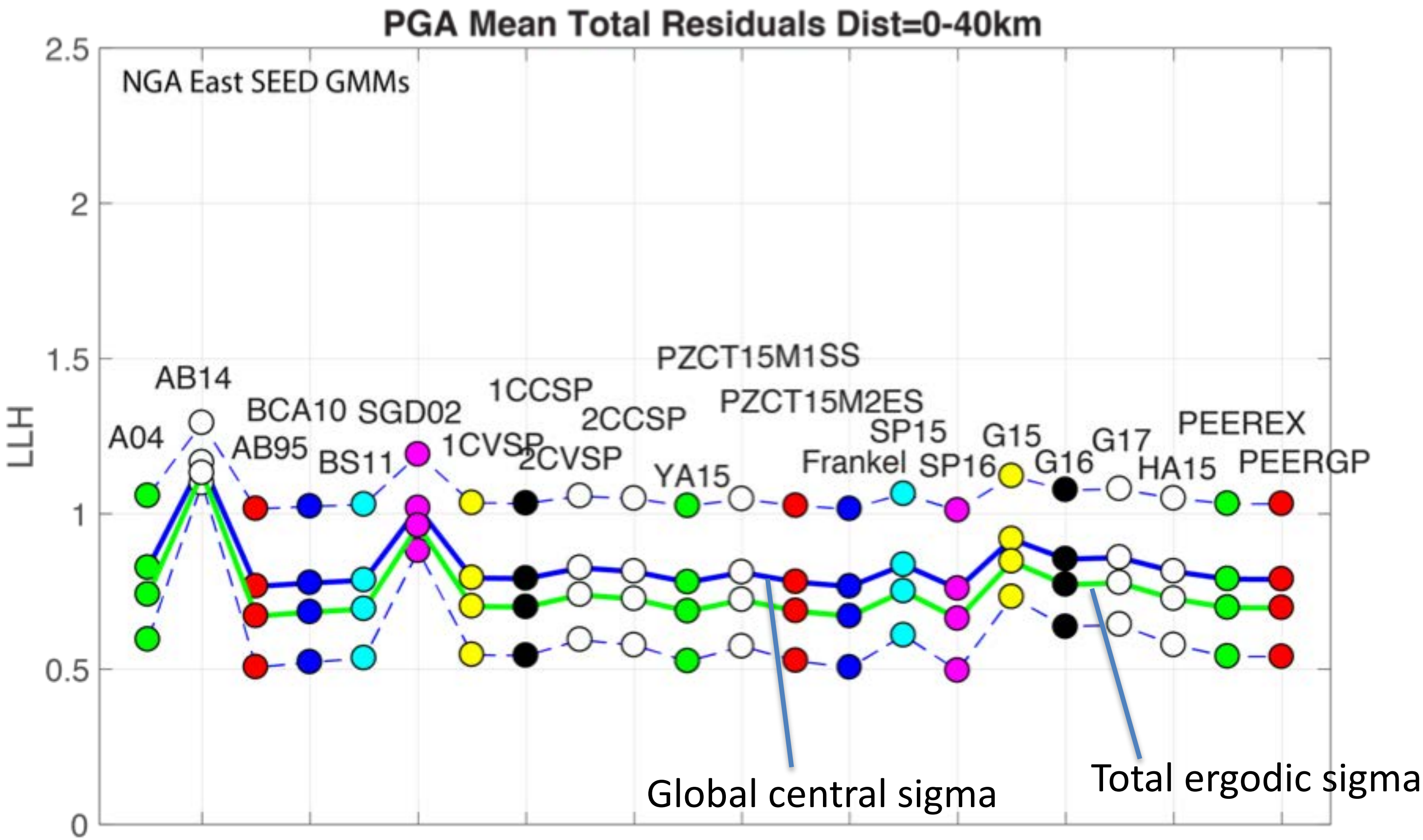
LLH/MLLH GM data, NGAE seed GMMs



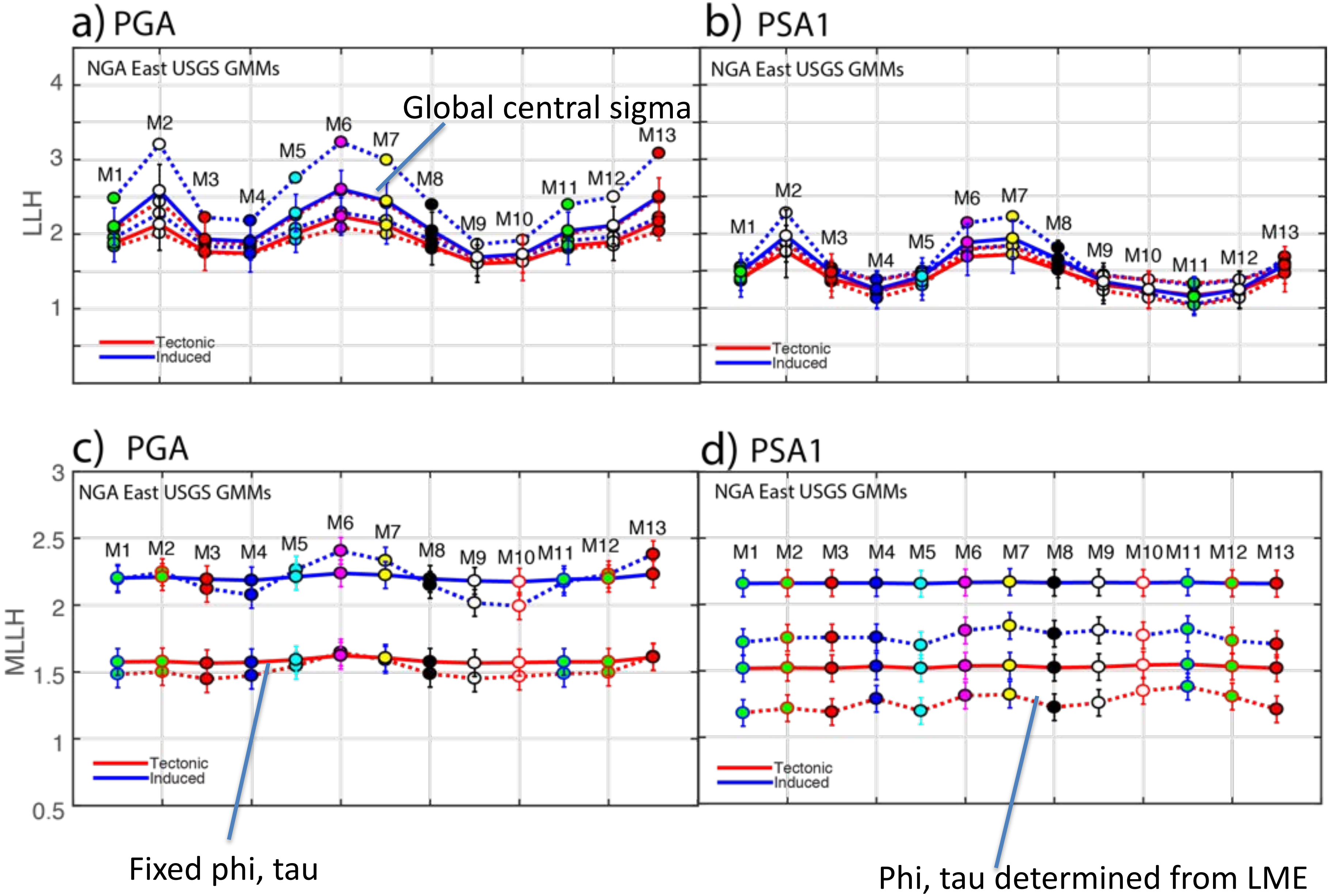
Fixed phi, tau

Phi, tau determined from LME

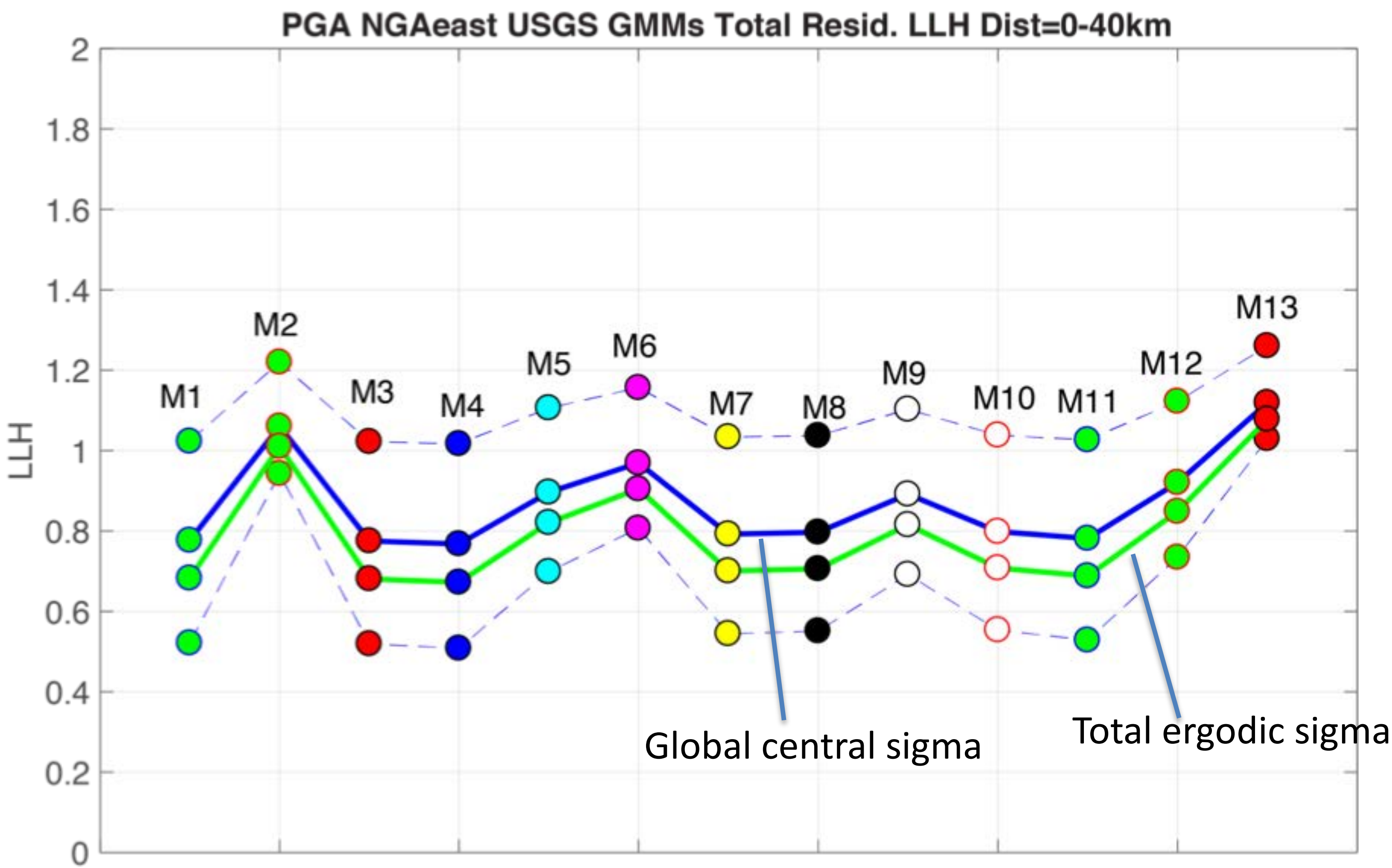
NGAE seed GMMs Simulation LLH



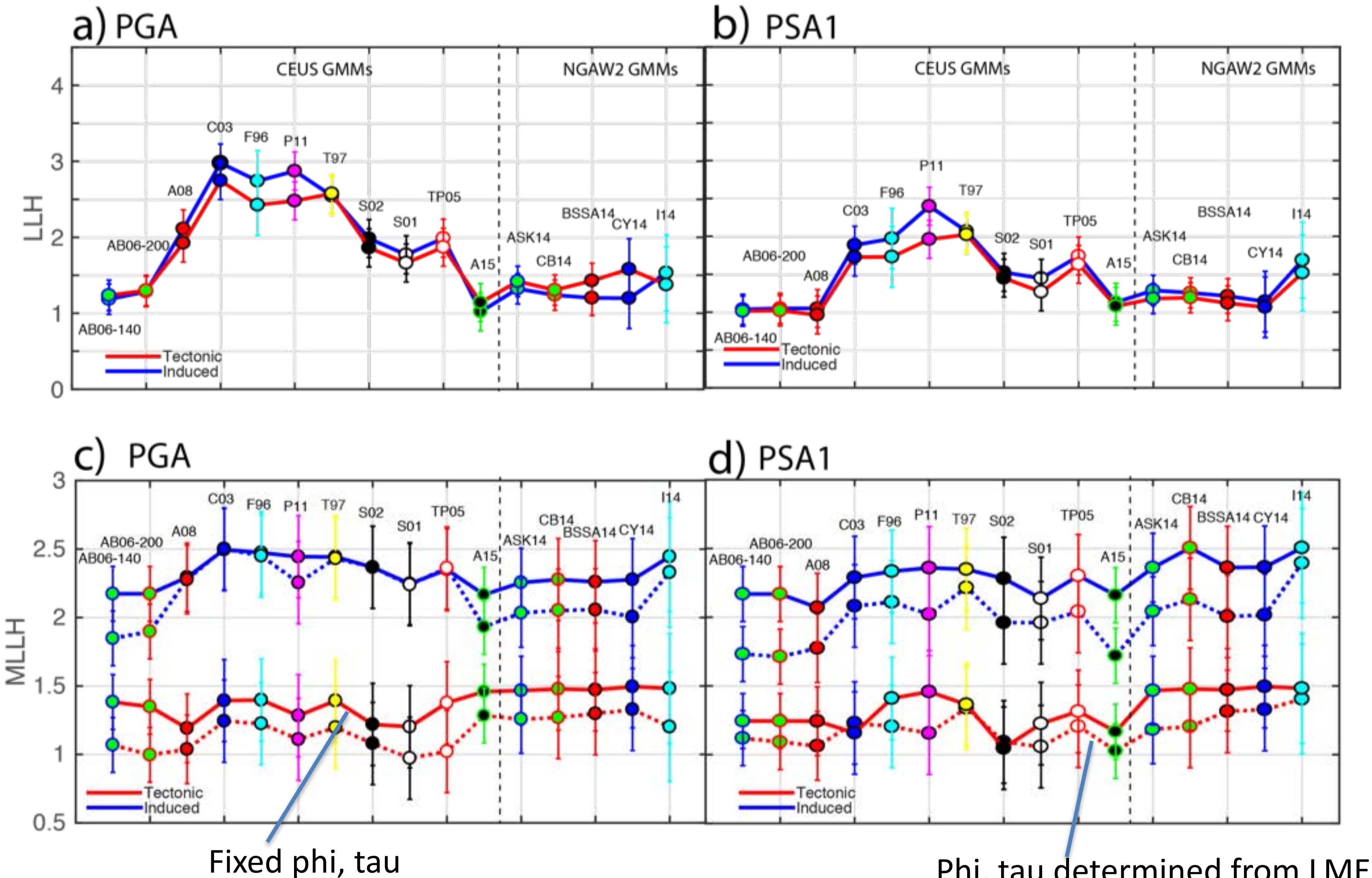
NGAE usgs LLH/MLLH GM data



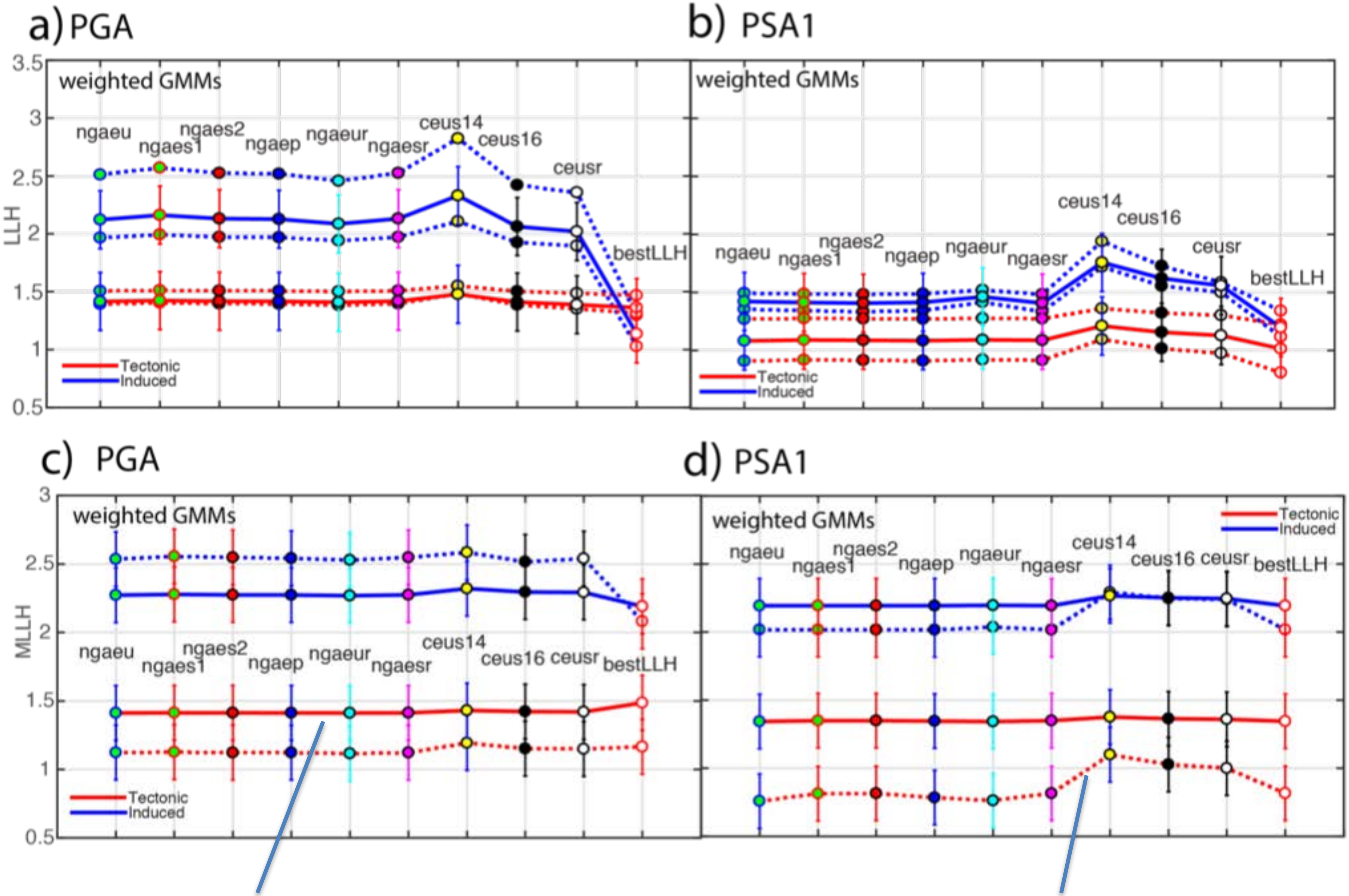
NGAE usgs GMMs Simulation LLH



LLH/MLLH GM data



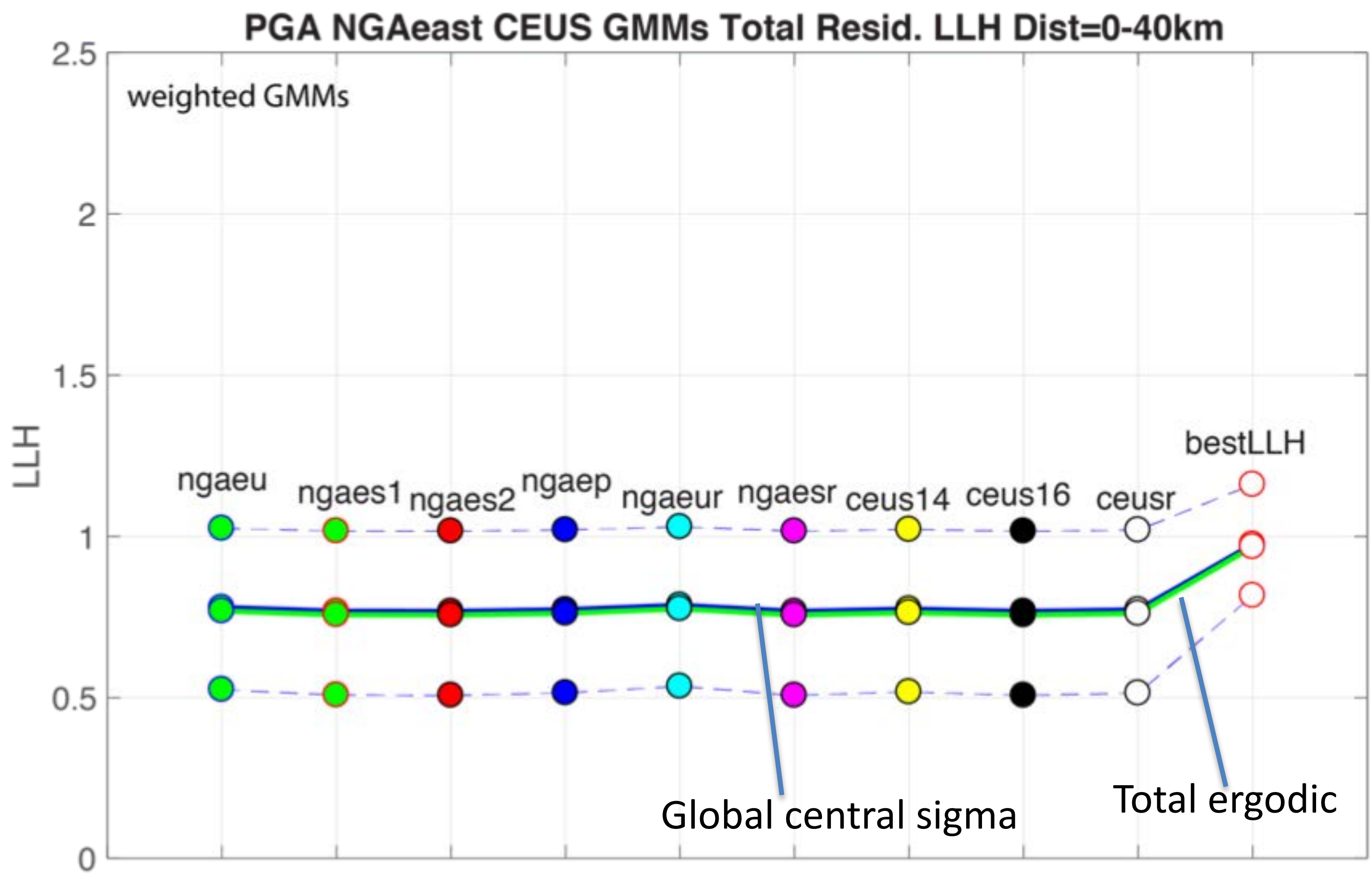
LLH/MLLH GM data weighted GMMS



Fixed phi, tau

Phi, tau determined from LME

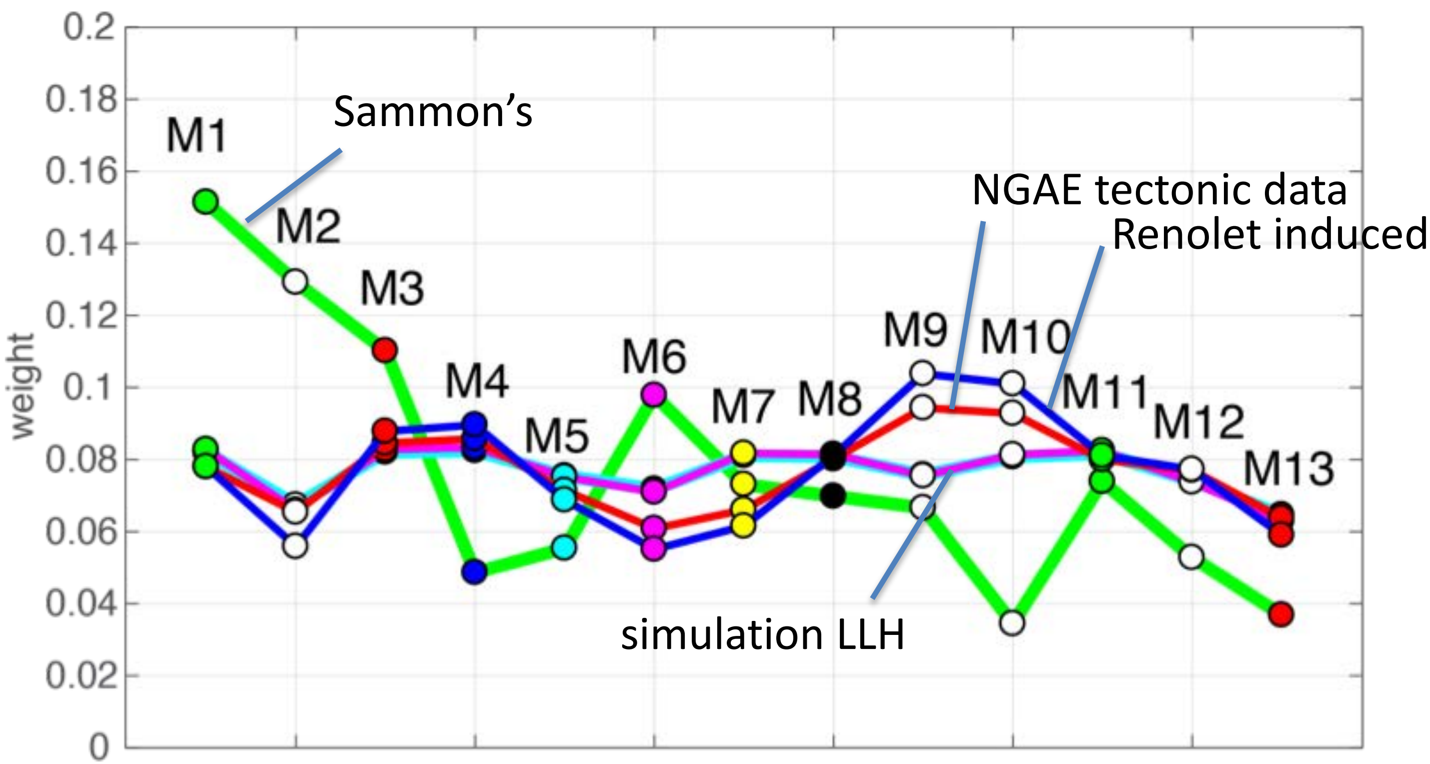
LLH Simulations weighted GMMs



Scherbaum LLH weighting

$$w_l = \frac{b^{-\langle \log_b(\mathcal{L}(g_l|\mathbf{x}))\rangle}}{\sum_{k=1}^{K} b^{-\langle \log_b(\mathcal{L}(g_k|\mathbf{x}))\rangle}}.$$

NGAE usgs GMMs PGA example



Observations

Older CEUS GMMs with large sigma have low LLH/MLLH scores

Most CEUS GMMs fit tectonic ground motions better than induced

All GMMs over-predict Induced GMs in 10-40km distance range

MLLH scores show GMM performance variation with individual phi, tau vs. fixed

LLH weights different than Sammons map weights

Not enough GM data to resolve significant differences in NGAE weighted GMMs

NGAE usgs GMMs generally perform better than older CEUS GMMs